

Article

Cognitive Networks and Text Analysis Identify Anxiety as a Key Dimension of Distress in Genuine Suicide Notes

Massimo Stella ^{1,*} , Trevor James Swanson ² , Andreia Sofia Teixeira ^{3,4}, Brianne N. Richson ², Ying Li ⁵, Thomas T. Hills ⁶, Kelsie T. Forbush ², David Watson ⁷ 

¹ CogNosco Lab, Department of Psychology and Cognitive Science, University of Trento, 38121 Trento, Italy

² Department of Psychology, University of Kansas, Lawrence, KS 66045, USA; trevorswanon222@gmail.com (T.J.S.); brichson@ku.edu (B.N.R.); kforbush@ku.edu (K.T.F.)

³ LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal; sofia.teixeira@nulondon.ac.uk

⁴ Network Science Institute, Northeastern University London, London E1W 1LP, UK

⁵ State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China; liying@psych.ac.cn

⁶ Department of Psychology, University of Warwick, Coventry CV4 7AL, UK; t.t.hills@warwick.ac.uk

⁷ Department of Psychology, University of Notre Dame, Notre Dame, IN 46556, USA; db.watson@nd.edu

* Correspondence: massimo.stella-1@unitn.it

Abstract

Understanding the mindset of people who die by suicide remains a key research challenge. We map conceptual and emotional word–word co-occurrences in 139 genuine suicide notes and in reference word lists, an Emotional Recall Task, from 200 individuals grouped by high/low depression, anxiety, and stress levels on DASS-21. Positive words cover most of the suicide notes' vocabulary; however, co-occurrences in suicide notes overlap mostly with those produced by individuals with low anxiety (Jaccard index of 0.42 for valence and 0.38 for arousal). We introduce a “words not said” method: It removes every word that corpus A shares with a comparison corpus B and then checks the emotions of “residual” words in $A - B$. With no leftover emotions, A and B are similar in expressing the same emotions. Simulations indicate this method can classify high/low levels of depression, anxiety and stress with 80% accuracy in a balanced task. After subtracting suicide note words, only the high-anxiety corpus displays no significant residual emotions. Our findings thus pin anxiety as a key latent feature of suicidal psychology and offer an interpretable language-based marker for suicide risk detection.

Keywords: complex networks; text analysis; emotional profiling; cognitive network science; suicide behavior; psychological distress



Academic Editor: Fabrizio Messina

Received: 4 February 2025

Revised: 5 June 2025

Accepted: 11 June 2025

Published: 27 June 2025

Citation: Stella, M.; Swanson, T.J.; Teixeira, A.S.; Richson, B.N.; Li, Y.; Hills, T.T.; Forbush, K.T.; Watson, D. Cognitive Networks and Text Analysis Identify Anxiety as a Key Dimension of Distress in Genuine Suicide Notes. *Big Data Cogn. Comput.* **2025**, *9*, 171. <https://doi.org/10.3390/bdcc9070171>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Approximately 800,000 people die by suicide every year—that is about one suicide every 40 s [1], and there is no evidence indicating a decline in these rates over time. Suicide often arises from complex, distressed emotional processing that includes distorted perceptions about the self and others, as well the world in general [2,3]. Given the variety and complexity of factors that may lead an individual to contemplate or complete suicide, it is important to advance research into the psychological conditions surrounding such a tragic event. Suicide notes—that is, letters written immediately prior to the author's suicide—represent one potential window into the mindset of individuals who complete suicide [4,5]. By analyzing the contents and language of suicide notes, we can gain unique

insights into the common features of individual experiences and obtain a greater understanding of the cognitive components that characterize suicidal ideation [6].

Previous research on suicide notes has focused on identifying their contents, as well as explicating what features differentiate them from other types of texts [7,8]. Al-Mosaiwi and Johnstone [9] found that the vocabulary of individuals at risk for suicide contained more absolutist words (e.g., “always”, “totally”, “entire”) than what was observed among those diagnosed with depression or anxiety. Understanding the emotional contents of suicide notes has also been a prominent research goal, wherein sentiment analysis has been employed to classify the emotions in such texts and train learning algorithms to distinguish between genuine and simulated suicide notes [10,11]. While these automated text-analytic techniques can be powerful for identifying written signs of suicidal ideation, they can be difficult to interpret and often use a “bag-of-words” approach that ignores the relationships between words and how they are used. Cognitive networks, however, offer a more readily-interpretable quantitative framework wherein the emotional contents of suicide notes can be analyzed while taking into account their associative structure—that is, networks allow for the study of both the *contents* and *contexts* of words used in suicide notes [12].

Recent studies of suicide notes with computational methods have provided unique insights into the structure of their emotional contents. For instance, Teixeira et al. [13] and Stella et al. [14] found, with different methodologies, that suicide notes exhibited a compartmentalized structure, such that positively- versus negatively-valenced words were often connected with one another [15]. Moreover, Stella et al. [14] found that the narratives of suicide notes contained a higher degree of narrative complexity, i.e., associations of contrastive elements, than notes written by a non-suicidal control group. Together, these findings reflect that suicide notes may contain a unique emotional and lexical footprint.

However, an important unresolved issue is whether traces of that lexical footprint [16] can be observed in notes written by individuals who exhibit high levels of emotional distress, such as stress, depression, and anxiety. Addressing this question represents the primary goal of the present research. Comparing the emotional structure of suicide notes with that of notes written by individuals reporting symptoms of internalizing disorders [17] provides a step forward in using cognitive networks to identify signs of suicidal ideation (and thereby, possible suicide risk) in written texts. Our motivation for this work lies in the potential of these insights to have clinical implications for improving prevention efforts. Specifically, by supplementing previous machine learning research [7,18,19], we can create network models that are more informative to clinicians and may be used to detect nuanced features of texts that flag groups who are at risk for suicidal behavior.

Several studies have shown that higher levels of stress, depression, and anxiety can be risk factors for suicidal ideation and behavior [20–23]. Research has shown that among adults who reported a lifetime suicide attempt, up to 70% of them had an anxiety disorder [24]. Similarly, approximately 60% of individuals who completed suicide were diagnosed with major depressive disorder [25]. Thus, to identify signs of suicidal ideation among individuals who report symptoms of internalizing disorders, it is an important goal to assess how the mindset of individuals who complete suicide might differ from the mindsets of those who exhibit psychopathological symptoms but did not attempt suicide.

Manuscript Aims and Research Questions

In this paper, we employ a cognitive network science framework to investigate the emotional contents and structure of suicide notes. Similar to past research that used networks to study suicide notes [13,14], we construct co-occurrence networks to map the associative structure of words used in such notes and analyze the organization of their

emotional contents. We create networks from a sample of recalls produced by individuals who reported high/low levels of internalizing symptoms (i.e., stress, depression, and anxiety), and aim to assess the structural similarities and differences between suicide notes and these other recalls.

Our main research questions are:

- RQ1: Using cognitive network science, do people with high/low anxiety or stress or depression tend to perform the same emotional associations compared to authors of genuine suicide notes?
- RQ2: Are there other negative emotional trends that are being obfuscated by a dominance of positive words and associations in suicide letters?

We explore RQ1 by building emotional networks from texts of suicide notes and from data available from past clinical studies [26]. To investigate RQ2, we outline a novel analysis of residual emotions in texts.

To the best of our knowledge, this paper represents the first attempt to directly compare networks created from suicide notes with those created from notes produced by individuals experiencing high/low levels of emotional distress. The primary goal behind this research is to compare and contrast the mindsets of individuals experiencing these negative psychological conditions with those of individuals who completed suicide to inform future research on suicide prevention.

2. Transparency and Openness

No studies within this paper were preregistered. The suicide note data were obtained via personal communication with external researchers, and so are not made publicly available by the current authors. Access to the de-identified ERT data collected by [26] is available at the following repository (accessed on 23 June 2024): <https://osf.io/b9h2t>. Access to all code used to conduct the analyses described in this paper are also available at the following repository (accessed on 23 June 2024): <https://osf.io/vxznr>. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. Our sample size of suicide notes was determined by the preexisting data that were obtained by Schoene and Dethlefs [11]; the same is true of the sample size for the ERT data obtained by Li et al. [26]. Ethical approval was granted from the HSSREC IRB panel at the University of Warwick to work with the suicide notes and the emotional recall data (HSSREC 63/21-22).

3. Methods

This section briefly outlines the datasets and network structures investigated in this study. An outline of the Methods adopted in this work is presented in Figure 1.

3.1. Genuine Suicide Notes and ERT Data

This work used 139 genuine suicide notes curated by Schoene and Dethlefs [11] and investigated also in previous works, through different methodologies [13,14,18]. These letters were written in English by individuals who completed suicide. A suicide letter included an average of 120 words. No additional contextual information (e.g., demographics) was available from the dataset. Notes were processed with names and sensitive information anonymized by Schoene and Dethlefs [11], thus filtering out any sensitive information from the analysis. Text was tokenized (i.e., split into words) and stemmed (e.g., regularize different forms to a common morphological root) through Mathematica 11.3. Punctuation was discarded and we considered only stemmed words appearing in the Emotional Recall Task (ERT). The corpus was gathered by other researchers, Schoene and Dethlefs [11], from fact-checked newspaper articles and other sources like books or

diaries collected by clinical psychologists. For privacy reasons, no demographic data about authors was mentioned in the original dataset gathered by Schoene and Dethlefs [11]. This corpus is mostly relative to native English speakers from the US and UK was collected over a span of 60 years, between 1958 and 2016.

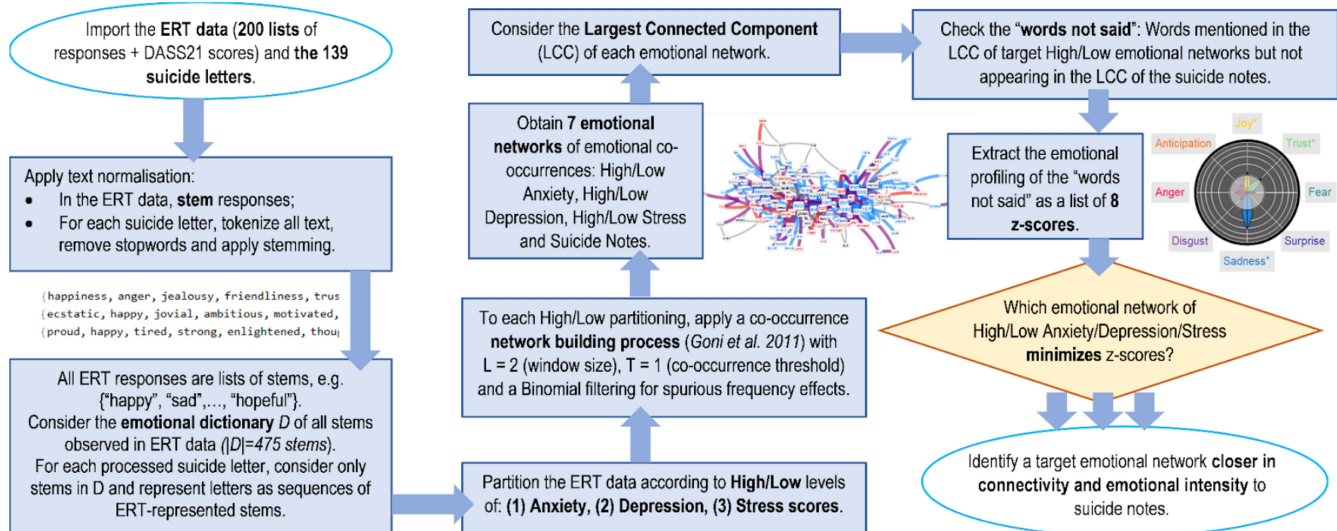


Figure 1. Flowchart of data processing within the “words not said” analysis. The aim of this data pipeline is to identify whether a given set of emotional responses (i.e., the ones mentioned in suicide notes) is similar in structure and emotional content to patterns observed in recalls produced by individuals with high or low levels of anxiety, depression and stress.

The Emotional Recall Task was introduced by Li et al. [26]. $N = 200$ individuals recruited from Amazon Mechanical Turk (MTurk) recalled 10 emotional states while reporting how they felt in their last month through a fluency task. The overall dataset included 475 different stems of emotional words (e.g., happy, sad, etc.) in English. We considered each recall as a list of words and stemmed it with Mathematica 11.3. Each word list/recall was relative to a numeric score from the Depression Anxiety Stress Scales (DASS; [17]). We partitioned ERT recalls as produced by people with higher-than-median and lower-than-median anxiety. We did the same for stress and depression, obtaining six different collections of ERT recalls.

For text analysis of the suicide notes, we used BERTopic version 0.17.0 in Python 3.13.3 to perform probabilistic topic modelling Boyd [27].

3.2. Emotional Co-Occurrence Networks

To identify the structure of emotional co-occurrences across populations experiencing depression, anxiety, stress and suicide ideation, we used a cognitive network science framework [28]. We used the sequences of emotional words from suicide notes and 200 responses from the ERT data to build networks of emotion co-occurrence in the narratives and recollections of our target populations (see also Supporting Information). As in previous studies dealing with fluency data, we used the approach by Goñi et al. [28] for building co-occurrence networks out of word lists. We used an $l = 2$ window, i.e., we considered the two words that immediately preceded and followed each word in the texts as co-occurring together. We applied this window to all words in all sentences, and recorded all co-occurrences as connections between pairs of concepts. We then discarded all idiosyncratic co-occurrences, that is, those appearing only once. We additionally selected links

by considering a Binomial filtering based on word frequency in the corpora. All corpora featured a total of $|D| = 475$ unique stems.

In any list, either those coming from ERT data or from texts, in order for any two words to appear in a co-occurrence link they must:

1. appear in the same list;
2. be at a distance equal to or lower than L in the list.

Considering f_i and f_j as the frequencies for words i and j in the considered list of lists, then the probability for those words to appear in the same list is:

$$P_{ij} = \frac{f_i f_j}{M^2}, \quad (1)$$

where M is the number of lists available. This equation assumes that words appear in lists at random, proportionally to their own frequencies and independently of other words, i.e., words' appearance follows a Bernoulli process. According to this assumption, it is relatively easy to identify the probability for any two words to appear in the same list and at up to distance $L = 2$ being:

$$P_{ij}^c = \frac{f_i f_j M^2}{2|U|(|U| - 1) \left(L|U| - \frac{L(L+1)}{2} \right)}, \quad (2)$$

where $|U|$ is the number of unique words across the considered lists (in our case U is a subset of D). For a detailed mathematical description of the method, we invite the interested reader to see Goñi et al. [28]. The probabilities P_{ij}^c can be used as a Binomial filter to discard all co-occurrences occurring a certain amount of times in the generative phase. Following Goñi et al. [28], we reconstructed a confidence interval at a significance of 0.05 from each empirically counted co-occurrence words i and j , using a Clopper–Pearson exact method. Our null hypothesis was that the observed co-occurrence was due to random frequency effects. We failed to reject the null hypothesis for co-occurrences where the left bound of the empirically retrieved confidence interval was lower than the random null probability P_{ij}^c . Otherwise, we rejected the null hypothesis and considered for the observed co-occurrence to arise not by chance (due to frequency effects) but rather to cognitive efforts in linking words together.

Those nodes selected according to the above process, within the set D , were added to all networks, and then connections for every emotional network were added according to the procedure by Goñi et al. [28]. This led to seven *emotional networks* where nodes represent emotional states and links indicate stronger-than-expected emotional co-occurrences in clinical populations, i.e., people experiencing suicidal ideation and low/high depression, anxiety and stress. These networks were: high anxiety (HA), low anxiety (LA), high depression (HD), low depression (LD), high stress (HS), low stress (LS) and suicide notes (SN).

As reported in the flowchart in Figure 1, we used the above data sources to identify whether the structure of emotional associations expressed in suicide notes resemble those produced by individuals with high or low levels of emotional distress, across the dimensions of anxiety, depression and stress.

Notice that emotional networks encode emotional associations as provided by individuals with higher or lower emotional distress and by authors of genuine suicide letters. The topology of these networks encapsulate relevant information for understanding how these groups of individuals structured and associated emotional words.

3.3. Emotional Auras and Jaccard Similarity

Words in each emotional network were enriched with valence/arousal labels from the valence–arousal–dominance dataset by Mohammad [29]. Labels for valence were “positive” (upper quartile), “negative” (lower quartile) or neutral (otherwise). Labels for arousal were “exciting” (upper quartile), “inhibitive” (lower quartile) or neutral (otherwise). These labels were used to detect the type of negative/positive or boring/exciting associations between emotional states. As in Stella et al. [14], we used the metric of *emotional auras* to combine valence/arousal labels and network connectivity. For each word, we detected its neighborhood, i.e., the set of all words co-occurring with it. Example neighborhoods for “happi” are reported in Figure 2a. We then counted the most frequent valence–arousal polarities populating each neighborhood; this is the *emotional aura* attributed to a word and depending on its neighborhood of associations. The aura of each word does not depend on its valence–arousal labels, but rather on the affective labels of its neighbors/associates. For instance, “happi” (stem for “happiness” and “happy”) is always labeled as a positive/arousing concept. However, it inherits a positive aura in the network produced by people with low depression, and a negative aura in the network produced by people with high depression, as represented in Figure 2a. Disconnected nodes received neutral auras. By considering an ordered list of nodes $\{D_1, D_2, \dots, D_{475}\}$, we represented each emotional network \mathcal{N} as a vector:

$$V_a(\mathcal{N}) = \{A_1, A_2, \dots, A_{475}\}, \quad (3)$$

of emotional auras for each corresponding node. By computing Jaccard similarity between vectors, we measured how similar emotional networks were in associating concepts with analogous emotional auras, mixing valence/arousal and emotion co-occurrence.

Emotional profiles were built by using the NRC Emotion Lexicon [30]. The latter is a behavioral mapping between words and the emotions elicited by them [30].

3.4. Mathematical Characterization of the “Words Not Said” Analysis

Counting how many emotions are present in a given set powered our “words not said” analysis. We measured residual emotions present in the complement C of two lists of words with degree $k \geq 1$, one list coming from a network of low/high negative states and the other list coming from suicide notes. For example, this procedure makes it possible to consider the emotions persisting in words that were mentioned by people with high depression but not mentioned/associated in suicide notes.

From a mathematical perspective, we denote with $V_{LCC}(\mathcal{N})$ the vertex set of all vertices/words being featured in the largest connected component of a network \mathcal{N} . The “words not said” focuses on a set subtraction operation between the vertex set of the largest connected component of the original emotional network based on suicide nodes, namely $LCC(SN)$, and each of the high/low emotional networks, respectively. We thus compute residual vertex sets:

$$R(s) = V_{LCC}(s) - (V_{LCC}(SN) \cap V_{LCC}(s)), \quad (4)$$

for $s \in \{HA, LA, HD, LD, HS, LS\}$.

The intersection is relevant as to not remove words that are not present in $LCC(s)$. The residual vertex sets R feature words that were mentioned and associated by individuals of the same level and type of emotional distress (e.g., high anxiety levels) but “not said” by authors of suicide letters. By design, this subtractive operation should remove any overlap between SN and a reference emotional network. Since SN features an abundance of positive associations, then the subtractive operation should remove most positive associations in SN and potentially highlight other emotional patterns. Notice also that the

subtractive operation is mathematically induced by network connectivity: Only words that are mentioned and associated in the reference emotional network (e.g., *HA*) can remain in $R(s)$. As reported in Table 1, most reference emotional networks do not coincide with their largest connected component. This means that the residual sets feature and depend on words of relevance for the connectivity of the overall emotional network.

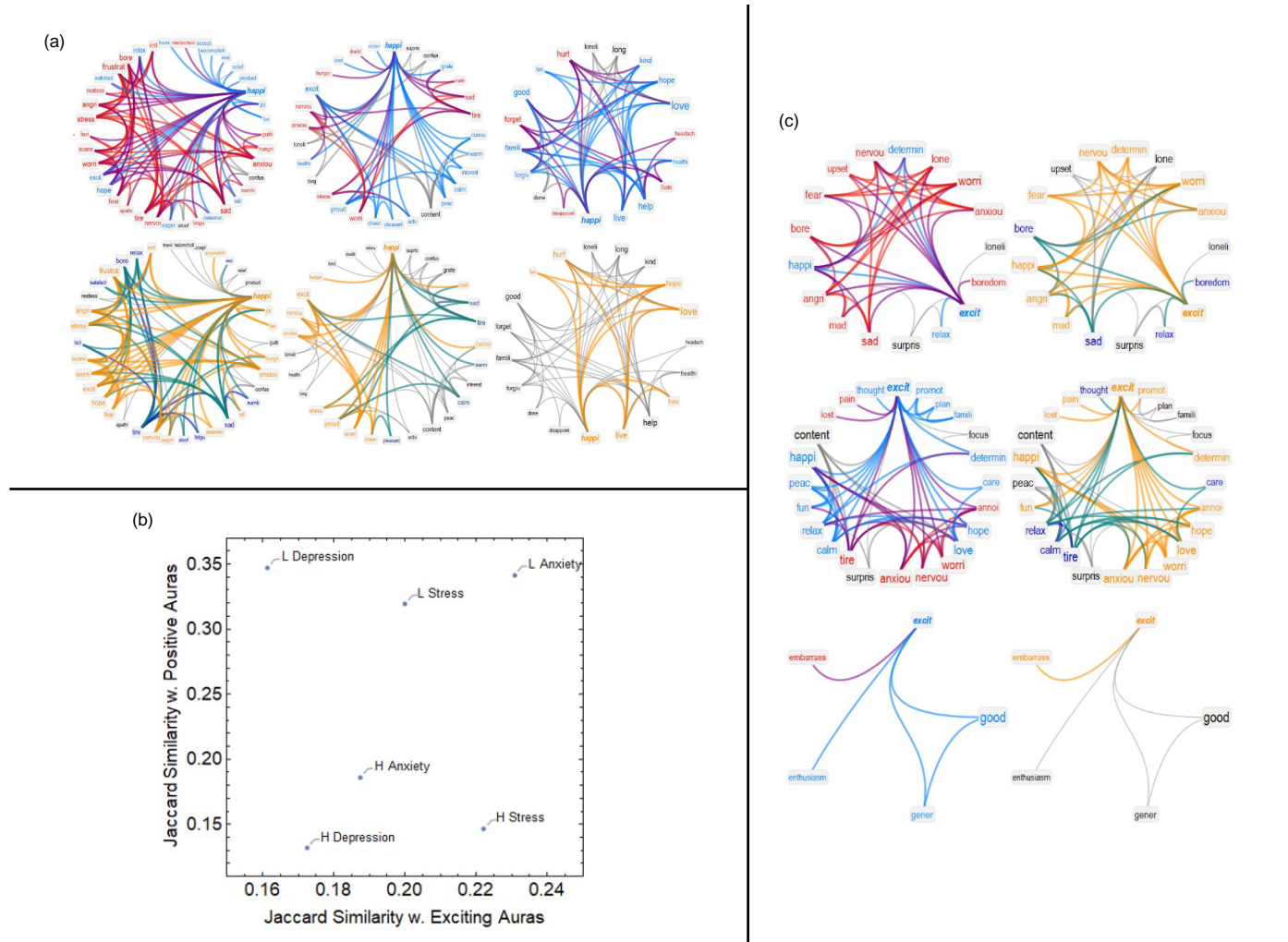


Figure 2. (a) Semantic frame for “happi” (stem for “happiness”, “happy”, etc.) in the emotional networks by people with high depression (left), low depression (center) and suicide ideation (right), respectively. On the top (bottom) row, colors are based on valence (arousal). Positive (exciting) concepts are highlighted in cyan (orange) while negative (inhibiting) concepts are highlighted in red (teal). Neutral words are in black. We see that “happi” has a strongly negative emotional aura in one context (high depression), and a positive emotional aura in another (low depression). This contextual effect exemplifies how the emotional aura of a word is determined by its associative structure, rather than its valence–arousal label. (b) Jaccard similarities between networks represented as vectors of emotional auras. (c) Semantic frame for “excit” (stem for “excitement”, “exciting”, etc.) in the emotional networks by people with high anxiety (top), low anxiety (center) and suicide ideation (bottom), respectively. On the left (right) column, colors are based on valence (arousal). Positive (exciting) concepts are highlighted in cyan (orange) while negative (inhibiting) concepts are highlighted in red (teal). Neutral words are in black.

These residual sets $R(s)$ feature emotional words mentioned by individuals with certain distress levels but not said in suicide letters. Intuitively, the closest similarity between the SN and the reference emotional network should happen when the residual sets feature emotionally neutral words, i.e., words that do not bring strong emotional

content. This is an important point, where similarity with the *SN* network originates from the fact that its emotional associations, when removed as an *LCC*, do not leave other strong emotional patterns in the *LCC* of another emotional network.

Table 1. Number of active vertices $|V|$, number of vertices in the largest connected component (LCC), number of links $|E|$, average clustering coefficient c and mean network distance d between any two words in each network. Notice that active vertices are by definition those vertices with at least one connection in a network.

	$ V $	$ V $ (LCC)	$ E $ (LCC)	c (LCC)	d (LCC)
HS	185	185	385	0.139	3.890
LS	218	204	455	0.194	3.591
HD	196	176	400	0.196	3.462
LD	217	203	420	0.151	3.676
HA	209	189	398	0.152	3.834
LA	227	193	436	0.179	3.709
SN	120	120	303	0.393	3.034

For any complement list C , we reconstructed its affective content as an emotional flower [16], where petals indicate z-scores of eight different emotional states based on how many words elicit a given state.

We operationalize the detection of strong emotional patterns through the emotional profile analysis, i.e., a statistical comparison between the observed counts of words that elicit specific emotions against a random null model assembling words at random from a model accounting for different amounts of words eliciting different emotions. Given a set of words W , our emotional profiling consists in computing pseudo-z-scores for every emotion e as:

$$z_e = \frac{n_e(W) - \langle n_e(r) \rangle}{\sigma_e(r)}, \quad (5)$$

where $n_e(W)$ counts how many words in W elicit emotion e (according [30]), $\langle n_e(r) \rangle$ is the average number of words found to elicit emotion e over 1000 iterations of a sampling process where the same amount a of words eliciting for any emotion in W is drawn uniformly at random from the whole emotional dataset, and $\sigma_e(r)$ is the standard error for $\langle n_e(r) \rangle$. The index e ran over emotions like: anger, fear, surprise, trust, anticipation, joy, disgust, and sadness.

The above statistical comparison indicates how rich a text is in terms of words eliciting emotions when compared to random assemblies of words. Importantly, these random null models take into account that the distribution of words across emotions is uneven, e.g., there are more words in language coding for anger than for surprise. The statistical appropriateness of this approach was extensively tested in [14]. We operationalized “strong” emotional intensities as being relative to $z_e > 1.96$.

For suicide notes, antonyms of words linked with meaning negations (e.g., “not”) were added to the count (antonyms were defined via WordNet 3.0; [31]). The observed profile was matched against 1000 random emotional profiles built by sampling uniformly at random from the NRC Emotion Lexicon as many words as those that elicited at least one emotion in the $\{r_i\}_i$ s. We computed eight z-scores, one per emotional state, and plotted them in a sector bar chart inspired by Plutchik’s wheel of emotions [32]. The rejection area $z < 1.96$ was plotted as a semi-transparent area and concentric circles indicated units of z-scores higher than 2. See the Supplementary Information for additional details on how z-scores were computed.

Importantly, all of these analyses rely on network structural properties for their conclusions. Emotional auras are determined by local patterns of connectivity within the

networks, as described in Figure 2a, while the “words not said” analysis evaluates the global connectivity of the networks in order to focus only on the largest connected component (LCC) for creating the network complement C. Further details about this analysis and our methodology can be found in the Supplementary Information.

3.5. Robustness of the “Words Not Said” Analysis

The “words not said” analysis is ultimately a psychometric measurement of the degree of emotional similarity between the SN connectivity (captured through connectedness) and the connectivity of other reference emotional networks. In psychometrics, it is common practice to establish the appropriateness and performance of measurement through artificial simulations (cf. Golino and Epskamp [33]). We follow this numerical simulation approach in view of a simple classification task.

Focusing on the six reference emotional high–low networks, we consider a binary, balanced classification task where we generate a simulated network as the mix of lists of high and low responses from the ERT data and then check whether the “words not said” analysis can correctly classify the network as being most similar to the classification where most lists come from. Let us unpack this statement. We generate an artificial network (AN) by adopting the same methodology described in Supplementary Materials Section S1, i.e., the co-occurrence protocol introduced by Goñi et al. [28]. We use the same parameters defined in Supplementary Information Section S1. However, an AN comes from a mixture of emotional recalls. We consider two cases: (1) a moderate noise scenario and (2) a high noise scenario.

Each of the reference high/low emotional networks is built over data partitioned through a median, so that all reference emotional networks are based on half the recalls in the ERT (which considers 200 recalls). To follow the same sample size, in both our simulated scenarios, we build each AN by using 100 recalls sampled uniformly from both the high and low partitions of a given emotional distress, e.g., depression. We build an AN by using predominantly recalls from the high (low) partition and then reach the 100 quota by sampling from the remaining low (high) partition:

- In the moderate noise scenario for simulating 100 high distress networks (e.g., 100 high depression networks), we sample 80% of recalls from high and 20% from low when building AN that should be classified as high;
- In the moderate noise scenario for simulating 100 low distress networks (e.g., 100 low depression networks), we sample 80% of recalls from low and 20% from high when building AN that should be classified as low;
- In the high noise scenario for simulating 100 high distress networks (e.g., 100 high anxiety networks), we sample 60% of recalls from high and 40% from low when building AN that should be classified as high;
- In the high noise scenario for simulating 100 low distress networks (e.g., 100 low anxiety networks), we sample 60% of recalls from low and 40% from high when building AN that should be classified as low.

For each dimension of emotional distress (anxiety, stress, depression), we build 100 ANs that should be categorized as high (containing mostly high-labeled recalls) and 100 ANs that should be categorized as low (containing mostly high-labeled recalls). We then perform a “words not said” analysis and note the estimated categorization. Measuring accuracy and the F1 score, we can quantify how accurately the “words not said” analysis performs in estimating high/low levels of emotional distress, for each distress dimension. Notice that this becomes a perfectly balanced, binary classification task, where either a random classifier or a majority-rule classifier would both achieve an accuracy and F1 score of 50%.

4. Results

This Section highlights our quantitative results relative to the syntactic, semantic and emotional organization of concepts within the considered sample of genuine suicide notes.

4.1. Topics and Emotional Content of Suicide Notes

Applying BERTopic to the considered suicide letters produced eight coherent thematic clusters, as reported in Table 2. The five most prevalent clusters characterized by high-probability lemmas, e.g., love, sorry, want, money; life, hope, God; and help, people, together encompassed 75% of the corpus. Qualitatively, these clusters capture (i) interpersonal affection or apology, (ii) financial/practical worries, (iii) existential or spiritual rumination, and (iv) pleas for understanding or assistance. These findings underscore how relational regret and existential distress dominate the narrative space of final communications.

Table 2. Topics, count of mentions across suicide notes and keywords for a topic analysis of suicide notes with BERTopic.

Topic	Count	Keywords
1	29	love, sorry, time, life, way, darling, think, always
2	23	love, know, good, sorry, just, like, life, time
3	18	love, way, good, things, make, happy, man, much
4	18	want, love, know, money, good, make, life
5	17	life, hope, god, people, help, sorry, father, love, dear
6	12	park, paris, mansfield, 10, 2000, matter, dear, son
7	12	got, like, january, told, loved, feel, time, good, years
8	10	life, going, way, like, people, friends, just, love

Before focusing on network construction, let us quantify the emotional content of every single suicide letter in the dataset. Results are reported in Figure 3. The figure displays the distribution of z-scores for the eight emotions considered in EmoAtlas. Each histogram represents how frequently specific z-scores occurred for a given emotion in suicide notes, with bars colored according to their respective emotion (consistent with the EmoAtlas palette). Vertical dashed red lines at ± 1.96 demarcate the statistical rejection region ($p < 0.05$), highlighting emotionally relevant content with stronger colors. Notably, joy, trust, and anticipation show a right-skewed distribution with a substantial number of texts exceeding the $+1.96$ threshold, suggesting these emotions were unusually elevated in certain notes. Conversely, anger, fear, and sadness exhibit broader distributions, with several texts falling into the significant range, though often in the negative direction. The gray-toned bars within the central region ($|z| < 1.96$) indicate emotionally neutral or statistically unremarkable content. This visualization reveals that, contrary to expectations, positive emotions may coexist with or even dominate certain suicide narratives, warranting further psychological interpretation. This finding motivates a closer look at the content and structure of the considered suicide notes, using our network-based approach.

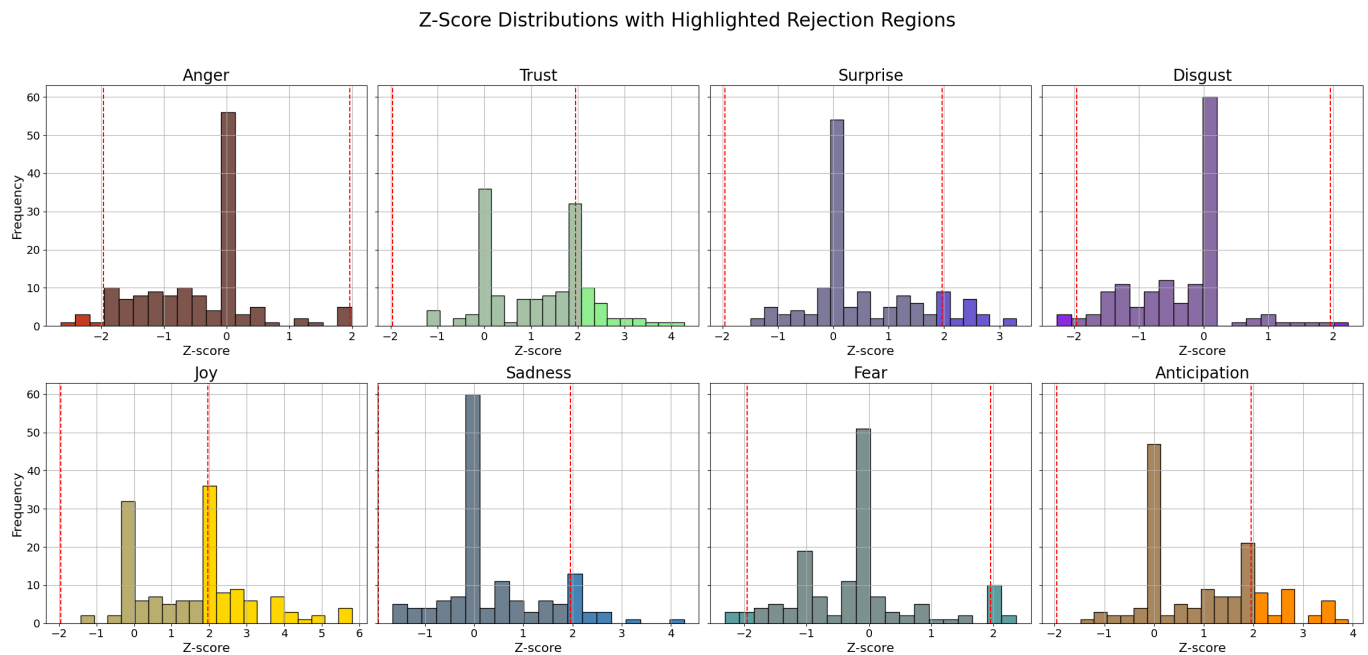


Figure 3. Emotional z-scores coming from each one of the considered suicide notes. The different colors highlight the rejection region ($|z| < 1.96$) while the overall color scheme follows the same of emotional flowers.

4.2. Topology of Emotional Networks for Anxiety, Stress, Depression and Suicide Notes

Table 1 reports the number of active vertices (i.e., words with at least one co-occurrence), the size of the largest connected set of words, the number of co-occurrences, the mean local clustering coefficient and the average network distance between any two words in all emotional networks.

As evident from Table 1 and from the cumulative degree distribution in Supplementary Materials Figure S1, all emotional networks share very similar topological structures, i.e., heavy-tailed degree distribution with short average network distance and relatively high clustering coefficient, all indications of a small-world structure. Using these features does not help us with the exploration of our research question and we need to go deeper, harnessing the specific semantic and emotional content of such networks in order to understand their similarities.

4.3. Suicide Notes' Semantic Frames

Suicide notes' semantic frames most resemble positive and arousing associations expressed by individuals with low anxiety.

First, as described by Stella et al. [14] and Teixeira et al. [13], we observed that the suicide notes contained an overwhelming majority of words with positive auras. Specifically, only three words were observed to have negative auras. Given the paucity of words with negative auras in the suicide notes as compared with the ERT data, we were not able to construct appropriate comparisons for words with negative auras across the two datasets. Thus, we limited the analysis of Jaccard similarity to words with positive auras. In addition to assessing the similarity of words with positive auras, we also included exciting auras as a second dimension for computing the Jaccard similarity of each ERT network against the suicide notes network.

The results for this analysis are displayed in Figure 2b, with the Jaccard similarity of words with positive auras plotted on the y-axis, and that for words with exciting auras plotted on the x-axis. Each point on the plot represents the Jaccard similarity between each ERT network and the suicide notes network on each of the two dimensions. The first

pattern that emerges is that low stress, anxiety, and depression networks have higher similarity values for words with positive auras than the corresponding high stress, anxiety, and depression networks. Notably, we see that the low anxiety network has the highest levels of Jaccard similarity on both dimensions. This means that words with positive auras and exciting auras in both the suicide notes and low anxiety networks show the most similar patterns of connectivity with other words, as compared to what we observe for the other ERT networks.

4.4. “Words Not Said”, Suicide Letters and High Anxiety

Given that the ERT networks contained a variety of words with negative auras that were not able to be included in the analysis (due to the dearth of words with negative auras in suicide notes), this begs the question of what emotional contents exist in those networks that were not expressed in the suicide notes. That is, our initial analysis was limited by restricting the ERT networks to only evaluate words that were also included in suicide notes. Thus, we wanted to evaluate the residual emotional contents expressed by the “words not said” in suicide notes, but *were said* by the individuals who completed the ERT. To do this, we constructed residual networks for the ERT data, wherein for each subsample (low/high stress, anxiety, and depression) we removed all words that were present in suicide notes and analyzed the emotional profiles for the remaining words in each network.

Figure 4 shows the emotional flowers for each of the six ERT residual networks, which reflect the “words not said” in suicide notes. In each emotional flower the petals reflect the prevalence, and statistical significance, of each individual emotion among the words in the corresponding residual network. Petals that extend beyond the semi-transparent circles indicate that the z-score for that emotion is greater than 2, or that it appeared with a significantly greater frequency than chance, given an alpha level of 0.05. The only ERT network that did not reveal any emotions occurring more frequently than chance was the high anxiety network.

4.5. The “Words Not Said” Analysis: Residual Emotional Levels in Suicide Notes

Low and high anxiety are important constructs to understand emotions (not) expressed in suicide notes. Our results reveal an interesting relationship between the emotional contents of suicide notes and those from emotional-recall data provided from a sample of healthy individuals. Specifically, we observed that the patterns of connectivity among emotion words with positive and exciting auras in suicide notes showed the greatest similarity with those in texts written by relatively low-anxiety individuals. This alone does not tell the whole story. We find that, after removing words used in suicide notes from networks generated by individuals high/low in stress, anxiety, and depression, all networks—except those representing written texts from high-anxiety individuals—contained additional emotional contents not present in suicide notes. Notice that the change of threshold from 25% to 33% in labeling words as “positive”, “negative” or “neutral” did not alter the layout of Jaccard similarities. This check indicates that our results are robust to the specific threshold adopted for partitioning words according to their valence and arousal.

Table 3 reports the results of the classification tasks across different dimensions of emotional distress once aggregated for each scenario.

Overall, our simulations indicate that even in the presence of high levels of noise in the artificial data, the “words not said” analysis performs consistently better than random classification (expected accuracy and F1 score of 50%). In the presence of high levels of noise, the accuracy and F1 scores are on average 14 points higher than random expectation. As expected, performance increases consistently in the presence of lower noise levels,

where the “words not said” analysis achieves an average accuracy centered around 80%. In this scenario, our “words not said” approach is able to accurately predict roughly 80% of the times the expected high/low categorization of an artificial network (see also Supplementary Materials).

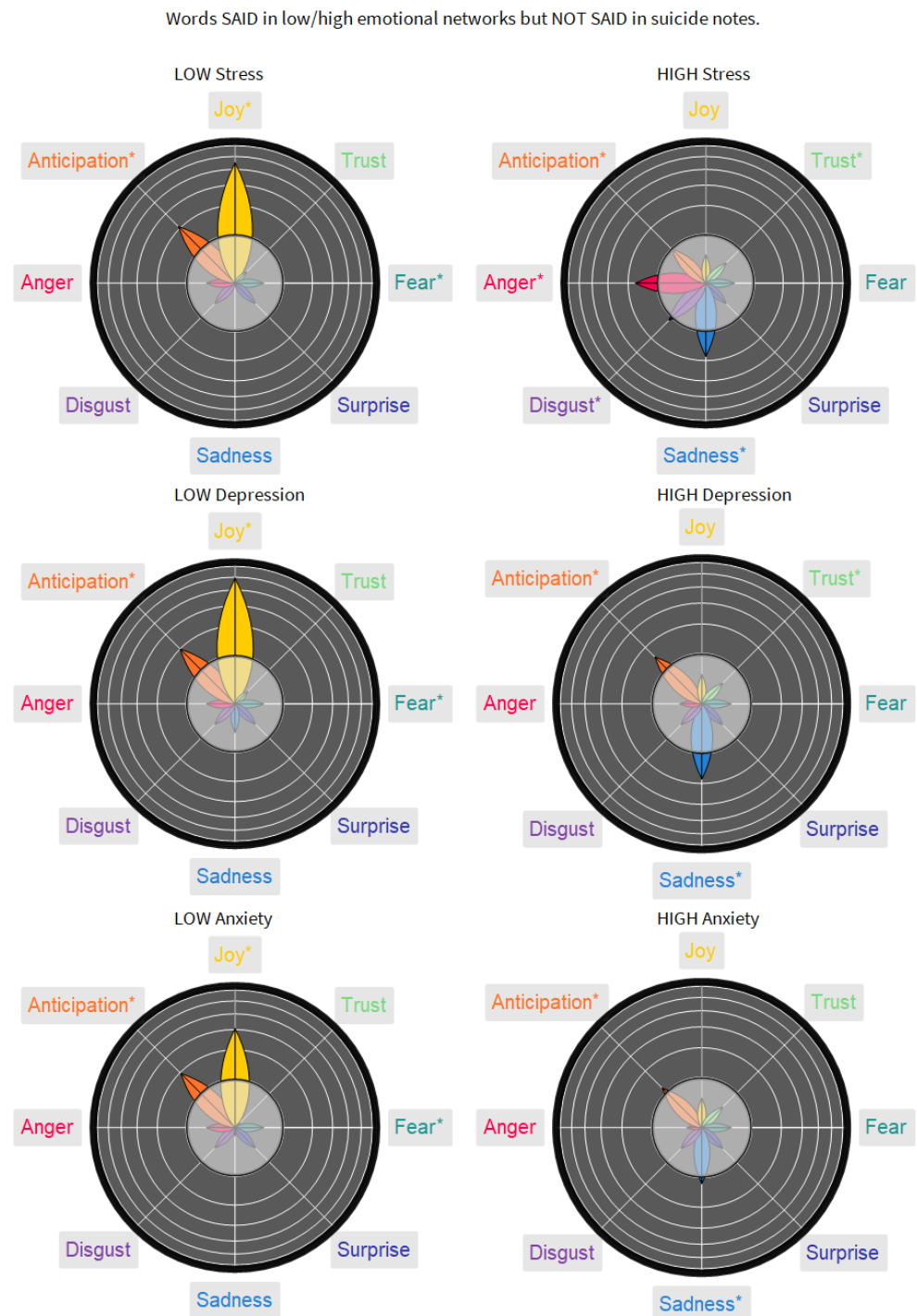


Figure 4. Emotional flowers for all words not mentioned in the suicide notes’ network but recalled by people experiencing high (left) and low (right) levels of anxiety (top), depression (center) and stress (bottom). On each row, the emotional flower in the middle is relative to suicide notes for an easier comparison. Petals falling outside of the semi-transparent circle indicate a stronger emotional richness of words eliciting a given emotion beyond random expectation (p -value < 0.05 , marked also with a star in the figure).

Table 3. Accuracy and F1 scores for each dimension of emotional distress. Each score is relative to 200 simulations with artificial networks and a “words not said” analysis. Values indicate means and are presented with standard errors.

	Moderate Noise (20%)		High Noise (40%)	
	Accuracy (%)	F1 (%)	Accuracy (%)	F1 (%)
Anxiety	80.1 ± 0.9	68.7 ± 0.9	66 ± 1	63.7 ± 0.9
Depression	79.5 ± 0.9	67.1 ± 0.9	63.9 ± 0.8	61.2 ± 0.9
Stress	80 ± 1	69.3 ± 0.7	67.2 ± 0.8	62.1 ± 0.9

The above results indicate that the network connectivity and the emotional pseudo-z-scores implemented in the “words not said” analysis are more accurate than random expectation in identifying levels of emotional distress. Hence, the outcome of these simulations is that the “words not said” analysis can consistently detect patterns of emotional distress and could thus be applied to the investigation of other networks, like the SN one. Notice that each of the above binary classification tasks are not performed perfectly with the “words not said” analysis, i.e., the algorithm never achieves an accuracy of 100%. This might be due to either some limitations of the data being processed as a network or rather to the presence of redundancy and correlations across the considered dimensions of emotional distress, i.e., some co-occurrences produced by individuals with low depression might be present also in the network produced by individuals with high depression. This overlap evidently limits the classification power of the approach, which should thus be interpreted in view of relevant psychological literature.

5. Discussion

The purpose of the current study was to compare the emotional content of words in 139 genuine suicide notes curated by Schoene and Dethlefs [11] to those from 200 healthy individuals who completed the Emotional Recall Task in which they recalled 10 emotional states while reporting how they felt in their last month through a fluency task. Our secondary goal was to identify how many emotions were present in a given set of “words not said” by measuring residual emotions from a network of low/high emotion states and one list coming from suicide notes. We observed that patterns of connectivity among emotional words in suicide notes were most similar to those in texts written by low-anxiety individuals, e.g., potentially healthy controls exhibiting low levels of anxiety in a DASS-21 scale. At the same time, after removing the bulk of positive associations, the remaining collection of emotions in suicide notes was most similar to states expressed by high-anxiety individuals.

Our findings reveal that the mindset of individuals who complete suicide has a complex association with the mindset of individuals experiencing anxiety. At first glance, it appears that the organization of emotion words in suicide notes is most similar to that in notes written by individuals with low levels of anxiety. However, upon closer investigation, it seems that individuals with low anxiety also frequently express other types of emotions within their written texts, particularly a high level of joy and anticipation. However, heightened levels of these additional emotions appear to be absent among individuals who reported higher levels of anxiety, indicating that the collection of emotional contents they express matches more closely to that observed in suicide notes. An interesting conclusion of this research is that understanding how the emotional contents in suicide notes relate to those expressed by individuals with different levels of internalizing symptoms may have less to do with what is actually written, and more to do with the “words not said”.

It is important to emphasize that these results are direct products of analyzing the connectivity of the networks studied in this paper. Where our study of emotional content

relies on the networks' local connectivity, and the "words not said" analysis relies on their global connectivity. Moreover, upon further analysis of other structural properties of the networks—including degree distribution, average clustering coefficient, and average network distance—we see that they all have a similar global topology (i.e., heavy-tailed degree distribution with short average network distance and relatively high clustering coefficient), highlighting that local connectivity is really the key feature driving the present results. See the Supplementary Information for more details about these additional analyses.

Suicidal ideation (a suicide risk factor) is associated with several anxiety disorders and related or secondary disorders (e.g., post-traumatic stress disorder, borderline personality disorder; [23,34]). One meta-analysis indicated that the presence of an anxiety disorder distinguishes individuals who think about suicide from those who attempt suicide, whereas depression was similarly prevalent in both groups [35]. By definition, anxiety is a psychological construct linked to a perception of real or perceived danger, eliciting a range of typical responses (e.g., fight-or-flight response; [36]), associated with high emotional and physiological arousal. High-anxiety individuals may have an external locus of control (i.e., feeling one's life is controlled by external factors); perceived lack of control may facilitate feelings of hopelessness that, in theory, exacerbate suicidal ideation [37,38]. Indeed, having an external locus of control has been associated with suicide attempts [39]. It is also possible that 'high anxiety' is a proxy for overarousal and/or sensitivity to the experience of anxiety, both of which have been suggested as risk factors for suicide [40,41]. More generally, neuroticism (a personality construct strongly related to anxiety) is likely negatively associated with positive affective experiences like joy [42].

The above provides some context to help explain why we observed a relationship between the emotions expressed by high-anxiety individuals and those expressed in suicide notes. However, the fact that patterns of connectivity among positively-valenced and high-arousal words in the suicide notes were most similar to those observed in low-anxiety individuals indicates that there is more to the story.

Multiple theories about suicide—e.g., the Interpersonal Theory of Suicide [38,43], the Three-Step Theory [44], and the Integrated Motivational-Volitional Model [45]—posit that the psychological pathway leading to suicidal ideation may differ from the pathway that fosters the capability for lethal self-injury. In these frameworks, interpersonal factors such as perceived burdensomeness and diminished belongingness can precipitate suicidal thoughts. By contrast, the capability to enact suicide is hypothesized to stem from diminished defensive responding—phenomenologically experienced as fearlessness or chronically low physiological anxiety—alongside higher pain tolerance [46,47]. Our "words not said" results importantly complement this view: We show that, once the overtly positive vocabulary of suicide notes is subtracted, the residual lexical network aligns most closely with high-anxiety fluency lists, suggesting that an anxious conceptual substrate co-exists with the ostensibly calm wording.

Our results capture an apparent contradiction: Suicide notes resemble language of individuals with low physiological anxiety yet contain hidden traces of high emotional anxiety—this echoes the dual-factor pattern reported in psychobiological studies. Higher self-reported emotional anxiety has been linked to suicidal ideation [23,47], whereas attenuated startle or skin-conductance responses (indices of low physiological anxiety) predict engagement in lethal or near-lethal behaviors; see the work by Smith and colleagues [48]. The residual-network method empirically captures both poles: positive surface wording maps onto low-anxiety fluency lists, and the "ghost" words that remain align with high-anxiety contexts.

This distinction becomes clearer when one inspects the content validity of the DASS-21 anxiety subscale, which we used to operationalize anxiety groups. The subscale emphasizes

somatic and panic-related symptoms (e.g., “I experienced difficulty breathing”), while containing relatively fewer items that tap cognitive worry or diffuse apprehension (e.g., “I felt scared without any good reason”). Consequently, a low DASS anxiety score may principally index low physiological arousal rather than the absence of anxious cognition, which links with the positive-surface/negative-residue split we observed. Moreover, expressive-writing research indicates that constructing a coherent narrative about impending self-injury can temporarily down-regulate autonomic arousal [49,50]. Such an acute calm-after-writing effect could partially explain the linguistic similarity between suicide notes and low-anxiety fluency lists, although longitudinal data would be required to test this mechanism directly.

It is important to note some limitations of this study. First, there was no demographic information available for the authors of the suicide notes we analyzed in this study. The sample of individuals who completed the ERT was collected from MTurk (a population that has been extensively studied in the social science literature; [51]), but it is unclear to what extent these individuals represented an appropriate comparison group for the sample of suicide-note authors. Future research would also benefit from analyzing a larger collection of suicide notes, such as the corpus of 1319 suicide notes described by Pestian et al. [52] which includes relevant demographic information such as gender, age, and education level. This would likely help to address another limitation we encountered, which was the dearth of negatively-valenced words in the suicide notes.

Some strengths of this study lie in its potential implications for informing research on predicting suicidal behavior from written texts. Previous research has used machine learning techniques to identify texts that may signal suicidal ideation [7,19], but such methods often produce results that are difficult to understand and thus may not be very helpful to clinicians for identifying the underlying features of a text that may signal risk for suicidal behavior [53]. Conversely, while qualitative methods for analyzing texts lend themselves better to interpretation, these may have limited reliability and be more difficult to apply on a broader scale (e.g., with larger corpora). Using cognitive network science, our approach retains the value of quantitative analysis while preserving the human dimension of the data. Network models [16] allow us to analyze the patterns of associative knowledge that represent the mindsets of individuals who may be at risk for suicide. This affords interpretable, automated emotion detection [16] which may be used to enhance clinical psychological investigations of an individual’s risk for suicide.

Future research should clarify if results similar to what we observed here emerge in studies wherein depression, stress, and anxiety are modeled together, as we recognize that these constructs are not mutually exclusive and are instead highly related [54]. Computational models [14] might tackle such future challenge. It will be important to test networks of suicide notes from other samples of individuals as well. For example, research is needed in clinical samples that are considered ‘at-risk’ for suicidal ideation and suicidal behavior, such as in individuals with depression, eating disorders, and PTSD [55,56]. The methods described in this paper could also be used to address other research questions relevant to mental health. For instance, another important application of these methods could be to analyze texts written by military service members diagnosed with PTSD, who are at a significantly elevated risk for death by suicide due to numerous risk factors [57,58]. In the presence of future datasets including also demographics, it would be interesting to explore whether the observed findings can change according to gender or other psychological dimensions. Evidence from large-scale lexical-semantic work indicates that affective ratings for stemmed words remain remarkably stable across historical periods [59]. Hence, the principal effect reported here—the disappearance of high-anxiety vocabulary once positively valenced words are subtracted—should be largely insensitive to sociocultural drift. In the absence of larger public datasets, and given the diachronic robustness documented

above, we regard the present corpus and analyses as a meaningful contribution despite their scale-related limitations. Last but not least, given that the risk of developing PTSD is high among military personnel [60,61], using cognitive networks to uncover traces of its ‘lexical footprint’ in texts written by active military members could offer a potential means of identifying those groups who may be most at-risk for death by suicide.

6. Conclusions

In this paper, we have shown how cognitive networks can provide unique insights into the mindsets of individuals who completed suicide, and how those insights can be used to identify similarities between the emotional contents of suicide notes and those expressed by individuals reporting high/low levels of emotional distress. Results preliminarily suggest that text communication platforms could be monitored for textual-connectivity patterns, or a lack of key positive emotional language, to inform suicide prevention.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/bdcc9070171/s1>, Figure S1: cumulative degree distributions; Section S1: Data processing, partitioning and cleaning; Section S2: Further notes on the mathematical characterization of the “words not said” analysis; Section S3: Measuring the correctness and performance of the “words not said” analysis [62–64].

Author Contributions: Conceptualization: T.J.S., A.S.T., Y.L., T.T.H. and M.S.; formal analysis: M.S.; methodology: M.S., T.J.S., A.S.T., Y.L. and T.T.H.; investigation: All authors; data curation: M.S., Y.L. and T.T.H.; writing—original draft preparation: All authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data was generated within this study.

Acknowledgments: A.S.T. acknowledges support by FCT—Fundação para a Ciência e Tecnologia—through the LASIGE Research Unit, ref. UID/000408/2025. The authors acknowledge Luciana Ciringione for useful feedback.

Conflicts of Interest: The authors declare that there were no conflicts of interest with respect to the authorship or the publication of this article.

References

1. World Health Organization. *Preventing Suicide: A Global Imperative*; World Health Organization: Geneva, Switzerland, 2014.
2. Oquendo, M.A.; Galfalvy, H.C.; Choo, T.H.; Kandlur, R.; Burke, A.K.; Sublette, M.E.; Miller, J.M.; Mann, J.J.; Stanley, B.H. Highly variable suicidal ideation: A phenotypic marker for stress induced suicide risk. *Mol. Psychiatry* **2021**, *26*, 5079–5086. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Rizk, M.M.; Choo, T.H.; Galfalvy, H.; Biggs, E.; Brodsky, B.S.; Oquendo, M.A.; Mann, J.J.; Stanley, B. Variability in suicidal ideation is associated with affective instability in suicide attempters with borderline personality disorder. *Psychiatry* **2019**, *82*, 173–178. [\[CrossRef\]](#) [\[PubMed\]](#)
4. McAdams, D.P. The psychology of life stories. *Rev. Gen. Psychol.* **2001**, *5*, 100–122. [\[CrossRef\]](#)
5. Proulx, T.; Heine, S.J. Death and black diamonds: Meaning, mortality, and the meaning maintenance model. *Psychol. Inq.* **2006**, *17*, 309–318. [\[CrossRef\]](#)
6. Schneidman, E.S. Suicide notes and tragic lives. *Suicide Life Threat. Behav.* **1981**, *11*, 286–299. [\[CrossRef\]](#)
7. Pestian, J.P.; Matykiewicz, P.; Linn-Gust, M.; South, B.; Uzuner, O.; Wiebe, J.; Cohen, K.B.; Hurdle, J.; Brew, C. Sentiment analysis of suicide notes: A shared task. *Biomed. Inform. Insights* **2012**, *5*, BII-S9042. [\[CrossRef\]](#)
8. Handelsman, L.D.; Lester, D. The content of suicide notes from attempters and completers. *Crisis* **2007**, *28*, 102–104. [\[CrossRef\]](#)
9. 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.* **2018**, *6*, 529–542. [\[CrossRef\]](#)
10. Pestian, J.; Nasrallah, H.; Matykiewicz, P.; Bennett, A.; Leenaars, A. Suicide note classification using natural language processing: A content analysis. *Biomed. Inform. Insights* **2010**, *3*, BII-S4706. [\[CrossRef\]](#)

11. 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, Berlin, Germany, 11 August 2016; pp. 128–133.
12. 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* **2020**, *476*, 20190825. [\[CrossRef\]](#)
13. Teixeira, A.S.; Talaga, S.; Swanson, T.J.; Stella, M. Revealing semantic and emotional structure of suicide notes with cognitive network science. *Sci. Rep.* **2021**, *11*, 1–15. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Stella, M.; Swanson, T.J.; Li, Y.; Hills, T.T.; Teixeira, A.S. Cognitive networks detect structural patterns and emotional complexity in suicide notes. *Front. Psychol.* **2022**, *13*, 917630. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Showers, C. Compartmentalization of positive and negative self-knowledge: Keeping bad apples out of the bunch. *J. Personal. Soc. Psychol.* **1992**, *62*, 1036. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Stella, M. Cognitive network science for understanding online social cognitions: A brief review. *Top. Cogn. Sci.* **2022**, *14*, 143–162. [\[CrossRef\]](#)
17. Lovibond, P.F.; Lovibond, S.H. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav. Res. Ther.* **1995**, *33*, 335–343. [\[CrossRef\]](#)
18. Fatima, A.; Li, Y.; Hills, T.T.; Stella, M. Dasentimental: Detecting depression, anxiety, and stress in texts via emotional recall, cognitive networks, and machine learning. *Big Data Cogn. Comput.* **2021**, *5*, 77. [\[CrossRef\]](#)
19. Schoene, A.M.; Turner, A.; De Mel, G.R.; Dethlefs, N. Hierarchical multiscale recurrent neural networks for detecting suicide notes. *IEEE Trans. Affect. Comput.* **2021**, *14*, 153–164. [\[CrossRef\]](#)
20. Sandin, B.; Chorot, P.; Santed, M.A.; Valiente, R.M.; Joiner Jr, T.E. Negative life events and adolescent suicidal behavior: A critical analysis from the stress process perspective. *J. Adolesc.* **1998**, *21*, 415–426. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Hawton, K.; Comabella, C.C.; Haw, C.; Saunders, K. Risk factors for suicide in individuals with depression: A systematic review. *J. Affect. Disord.* **2013**, *147*, 17–28. [\[CrossRef\]](#)
22. Sareen, J.; Cox, B.J.; Afifi, T.O.; de Graaf, R.; Asmundson, G.J.; Ten Have, M.; Stein, M.B. Anxiety disorders and risk for suicidal ideation and suicide attempts: A population-based longitudinal study of adults. *Arch. Gen. Psychiatry* **2005**, *62*, 1249–1257. [\[CrossRef\]](#)
23. Naragon-Gainey, K.; Watson, D. The anxiety disorders and suicidal ideation: Accounting for co-morbidity via underlying personality traits. *Psychol. Med.* **2011**, *41*, 1437–1447. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Nepon, J.; Belik, S.L.; Bolton, J.; Sareen, J. The relationship between anxiety disorders and suicide attempts: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Depress. Anxiety* **2010**, *27*, 791–798. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Ng, C.W.M.; How, C.H.; Ng, Y.P. Depression in primary care: Assessing suicide risk. *Singap. Med J.* **2017**, *58*, 72. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Li, Y.; Masitah, A.; Hills, T.T. The Emotional Recall Task: Juxtaposing recall and recognition-based affect scales. *J. Exp. Psychol. Learn. Mem. Cogn.* **2020**, *46*, 1782. [\[CrossRef\]](#)
27. Boyd, R.L. Psychological text analysis in the digital humanities. In *Data Analytics in Digital Humanities*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 161–189.
28. Goñi, J.; Arrondo, G.; Sepulcre, J.; Martincorena, I.; De Mendizábal, N.V.; Corominas-Murtra, B.; Bejarano, B.; Ardanza-Trevijano, S.; Peraita, H.; Wall, D.P.; et al. The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cogn. Process.* **2011**, *12*, 183–196. [\[CrossRef\]](#)
29. Mohammad, S. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia, 15–20 July 2018; pp. 174–184.
30. Mohammad, S.M.; Turney, P.D. Crowdsourcing a word–emotion association lexicon. *Comput. Intell.* **2013**, *29*, 436–465. [\[CrossRef\]](#)
31. Fellbaum, C. WordNet. In *Theory and Applications of Ontology: Computer Applications*; Poli, R., Healy, M., Kameas, A., Eds.; Springer: Dordrecht, The Netherlands, 2012.
32. Plutchik, R. Emotion. In *A Psychoevolutionary Synthesis*; Academic Press: Cambridge, MA, USA, 1980.
33. Golino, H.F.; Epskamp, S. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE* **2017**, *12*, e0174035. [\[CrossRef\]](#)
34. Links, P.S.; Eynan, R.; Heisel, M.J.; Barr, A.; Korzekwa, M.; McMain, S.; Ball, J.S. Affective instability and suicidal ideation and behavior in patients with borderline personality disorder. *J. Personal. Disord.* **2007**, *21*, 72–86. [\[CrossRef\]](#)
35. May, A.M.; Klonsky, E.D. What distinguishes suicide attempters from suicide ideators? A meta-analysis of potential factors. *Clin. Psychol. Sci. Pract.* **2016**, *23*, 5.
36. Cannon, W.B. *Bodily Changes in Pain, Hunger, Fear and Rage: An Account of Recent Researches into the Function of Emotional Excitement*; D. Appleton & Co.: New York, NY, USA, 1925.

37. Weishaar, M.E.; Beck, A.T. Hopelessness and suicide. *Int. Rev. Psychiatry* **1992**, *4*, 177–184. [\[CrossRef\]](#)
38. Joiner, T.E. *Why People Die by Suicide*; Harvard University Press: Cambridge, MA, USA, 2005.
39. Wiebenga, J.X.; Eikelenboom, M.; Heering, H.D.; van Oppen, P.; Penninx, B.W. Suicide ideation versus suicide attempt: Examining overlapping and differential determinants in a large cohort of patients with depression and/or anxiety. *Aust. N. Z. J. Psychiatry* **2021**, *55*, 167–179. [\[CrossRef\]](#)
40. Ribeiro, J.D.; Yen, S.; Joiner, T.; Siegler, I.C. Capability for suicide interacts with states of heightened arousal to predict death by suicide beyond the effects of depression and hopelessness. *J. Affect. Disord.* **2015**, *188*, 53–59. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Stanley, I.H.; Boffa, J.W.; Rogers, M.L.; Hom, M.A.; Albanese, B.J.; Chu, C.; Capron, D.W.; Schmidt, N.B.; Joiner, T.E. Anxiety sensitivity and suicidal ideation/suicide risk: A meta-analysis. *J. Consult. Clin. Psychol.* **2018**, *86*, 946. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Ng, W. Clarifying the relation between neuroticism and positive emotions. *Personal. Individ. Differ.* **2009**, *47*, 69–72. [\[CrossRef\]](#)
43. Van Orden, K.A.; Witte, T.K.; Cukrowicz, K.C.; Braithwaite, S.R.; Selby, E.A.; Joiner Jr, T.E. The interpersonal theory of suicide. *Psychol. Rev.* **2010**, *117*, 575. [\[CrossRef\]](#)
44. Klonsky, E.D.; May, A.M. The three-step theory (3ST): A new theory of suicide rooted in the “ideation-to-action” framework. *Int. J. Cogn. Ther.* **2015**, *8*, 114–129. [\[CrossRef\]](#)
45. O'Connor, R.C. Towards an integrated motivational–volitional model of suicidal behaviour. *Int. Handb. Suicide Prev. Res. Policy Pract.* **2011**, *1*, 181–198.
46. Bayliss, L.T.; Christensen, S.; Lamont-Mills, A.; du Plessis, C. Suicide capability within the ideation-to-action framework: A systematic scoping review. *PLoS ONE* **2022**, *17*, e0276070. [\[CrossRef\]](#)
47. Burke, T.A.; Ammerman, B.A.; Knorr, A.C.; Alloy, L.B.; McCloskey, M.S. Measuring acquired capability for suicide within an ideation-to-action framework. *Psychol. Violence* **2018**, *8*, 277. [\[CrossRef\]](#)
48. Smith, P.N.; Cukrowicz, K.C.; Poindexter, E.K.; Hobson, V.; Cohen, L.M. The acquired capability for suicide: A comparison of suicide attempters, suicide ideators, and non-suicidal controls. *Depress. Anxiety* **2010**, *27*, 871–877. [\[CrossRef\]](#)
49. Stephens, C. Narrative analysis in health psychology research: Personal, dialogical and social stories of health. *Health Psychol. Rev.* **2011**, *5*, 62–78. [\[CrossRef\]](#)
50. Pennebaker, J.W. Writing about emotional experiences as a therapeutic process. *Psychol. Sci.* **1997**, *8*, 162–166. [\[CrossRef\]](#)
51. Huff, C.; Tingley, D. “Who are these people?” Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Res. Politics* **2015**, *2*, 2053168015604648. [\[CrossRef\]](#)
52. Pestian, J.P.; Matykievicz, P.; Linn-Gust, M. What’s in a note: Construction of a suicide note corpus. *Biomed. Inform. Insights* **2012**, *5*, BII-S10213. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [\[CrossRef\]](#) [\[PubMed\]](#)
54. O'Driscoll, C.; Buckman, J.E.; Fried, E.I.; Saunders, R.; Cohen, Z.D.; Ambler, G.; DeRubeis, R.J.; Gilbody, S.; Hollon, S.D.; Kendrick, T.; et al. The importance of transdiagnostic symptom level assessment to understanding prognosis for depressed adults: Analysis of data from six randomised control trials. *BMC Med.* **2021**, *19*, 109. [\[CrossRef\]](#)
55. Chesney, E.; Goodwin, G.M.; Fazel, S. Risks of all-cause and suicide mortality in mental disorders: A meta-review. *World Psychiatry* **2014**, *13*, 153–160. [\[CrossRef\]](#)
56. Fu, X.L.; Qian, Y.; Jin, X.H.; Yu, H.R.; Wu, H.; Du, L.; Chen, H.L.; Shi, Y.Q. Suicide rates among people with serious mental illness: A systematic review and meta-analysis. *Psychol. Med.* **2023**, *53*, 351–361. [\[CrossRef\]](#)
57. Xue, C.; Ge, Y.; Tang, B.; Liu, Y.; Kang, P.; Wang, M.; Zhang, L. A meta-analysis of risk factors for combat-related PTSD among military personnel and veterans. *PLoS ONE* **2015**, *10*, e0120270. [\[CrossRef\]](#)
58. Ramchand, R.; Rudavsky, R.; Grant, S.; Tanielian, T.; Jaycox, L. Prevalence of, risk factors for, and consequences of posttraumatic stress disorder and other mental health problems in military populations deployed to Iraq and Afghanistan. *Curr. Psychiatry Rep.* **2015**, *17*, 37. [\[CrossRef\]](#)
59. Hills, T.T.; Proto, E.; Sgroi, D.; Seresinhe, C.I. Historical analysis of national subjective wellbeing using millions of digitized books. *Nat. Hum. Behav.* **2019**, *3*, 1271–1275. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Armed Forces Health Surveillance Center. Deaths by suicide while on active duty, active and reserve components, US Armed Forces, 1998–2011. *Msmr* **2012**, *19*, 7–10.
61. Martin, J.; Ghahramanlou-Holloway, M.; Lou, K.; Tucciarone, P. A comparative review of US military and civilian suicide behavior: Implications for OEF/OIF suicide prevention efforts. *J. Ment. Health Couns.* **2009**, *31*, 101–118. [\[CrossRef\]](#)
62. Stella, M. Text-mining forma mentis networks reconstruct public perception of the STEM gender gap in social media. *PeerJ Comput. Sci.* **2020**, *6*, e295. [\[CrossRef\]](#) [\[PubMed\]](#)

63. Newman, M. *Networks*; Oxford University Press: Oxford, UK, 2018.
64. Mohammad, S.; Turney, P. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, CA, USA, 5 June 2010; pp. 26–34.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.