Semantic Shifts in Mental Health-Related Concepts

Naomi Baes[ϕ] **Nick Haslam**[ϕ] **Ekaterina Vylomova**[λ]

[ϕ] Melbourne School of Psychological Sciences

[λ]School of Computing and Information Systems

The University of Melbourne

{n.baes, nhaslam, vylomovae}@unimelb.edu.au

Abstract

The present study evaluates semantic shifts in mental health-related concepts in two diachronic corpora spanning 1970–2016, one academic and one general. It evaluates whether their meanings have broadened to encompass less severe phenomena and whether they have become more pathology related. It applies a recently proposed methodology (Baes et al., 2023) to examine whether words collocating with a sample of mental health concepts have become less emotionally intense and develops a new way to examine whether the concepts increasingly co-occur with pathology-related terms. In support of the first hypothesis, mental health-related concepts became associated with less emotionally intense language in the psychology corpus (*addiction*, *anger*, *stress*, *worry*) and in the general corpus (*addiction*, *grief*, *stress*, *worry*). In support of the second hypothesis, mental health-related concepts came to be more associated with pathology-related language in psychology (*addiction*, *grief*, *stress*, *worry*) and in the general corpus (*grief*, *stress*). Findings demonstrate that some mental health concepts have become normalized and/or pathologized, a conclusion with important social and cultural implications.

1 Introduction

Mental health has become more culturally salient in recent years. Concurrently, concepts of mental illness have expanded their meanings to include new and milder phenomena (Haslam, 2016, 'concept creep'). Critics have argued that psychiatry has transformed everyday sadness into major depression (Horwitz and Wakefield, 2007) and adaptive worries and inhibitions into anxiety disorders (Horwitz and Wakefield, 2012). Others argue that this pattern extends to colloquial language, where people use 'depression' to refer to ordinary sadness or low mood (Bröer and Besseling, 2017). Brinkmann (2016) explains this trend

as an ongoing cultural process of 'pathologization', where traits and behaviors that were once considered normal human problems (e.g., inattentiveness) are now conceptualized as mental disorders (e.g., attention-deficit/hyperactivity disorder) to be diagnosed and treated. Critics argue that pathologization leads to increasing vulnerability, which can partly be explained by the adoption of illness identities (Furedi, 2004), and to false positive diagnoses, resulting in the misallocation of treatment resources (Wakefield, 2010). Despite these arguments, whether people are indeed normalizing mental illness and pathologizing everyday life remains a largely untested empirical question.

The present study aims to clarify the nature of these semantic shifts in a new sample of mental health-related concepts: *addiction*, *anger*, *distress*, *grief*, *stress*, and *worry*. It first investigates whether they have undergone vertical concept creep (come to include less severe phenomena) and then tests whether they have become pathologized using a new index based on a dictionary of pathology-related terms. It hypothesizes that words collocating with mental health-related concepts have (1) become less emotionally severe (vertical concept creep) and (2) come to co-occur with pathology-related terms (pathologization).

2 Concept Creep Theory

According to concept creep theory (Haslam, 2016), harm-related concepts are susceptible to two kinds of semantic expansion, broadening to encompass qualitatively new phenomena (horizontal creep) and quantitatively less severe phenomena (vertical creep). Linguistically, horizontal creep resembles semantic widening, including via metaphorical extension, while vertical creep resembles hyperbole, where words shift from a stronger to a weaker meaning (Vylomova and Haslam, 2021). Both forms of creep can occur simultaneously. Theorized causes of concept creep include cultural

shifts towards greater sensitivity to harm, post-materialist values and diminished exposure to adversity (Haslam et al., 2020). As with other harm concepts, the consequences of inflated concepts of mental illness are mixed. Negative consequences include excessive self-diagnosis, prescription of inappropriate health services and treatments (Xiao et al., 2023) and heightened emotional vulnerability (Jones and McNally, 2022). Positive consequences include recognizing and addressing previously neglected forms of suffering (Tse and Haslam, 2021; Foulkes and Andrews, 2023).

3 Related Work

Advances in computational linguistics have facilitated the detection and quantification of diachronic lexical semantic shifts (Tahmasebi and Dubossarsky, 2023), as outlined in pioneering survey papers (Kutuzov et al., 2018; Tahmasebi et al., 2021). New techniques to digitize, process, store, and quantify written language worldwide have enabled non-computational disciplines to use text corpora to explore questions with a social dimension. For instance, linguist Price (2022) used methods from computational linguistics to track the construction of mental illness in the UK press. In psychological science, text mining approaches are gaining traction as researchers begin to harness the advances in modern computational technologies and digital data sources by using natural language processing (NLP) as a tool to understand people and culture at an unprecedented scale. For reviews explaining this paradigm shift in psychology, see Berger and Packard (2022); Jackson et al. (2022); Pennebaker (2022); Demszky et al. (2023). Nevertheless, the field is only beginning to reap the benefits of using NLP to examine social and cultural change (Charlesworth et al., 2023; Leach et al., 2023).

Concept creep research provides an innovative engagement between social psychology and NLP. As the only theory in social psychology with a focus on lexical semantic change and its non-linguistic (societal, politically motivated, and cultural) causes and social consequences, researchers have employed NLP techniques to characterize and to track the theorized causes of concept creep (Haslam et al., 2020). Studies have revealed increases in the relative frequency of words reflecting harm-based morality since 1900 in the Google Books English corpus (Wheeler et al., 2019), and

of prejudice-denoting terms in popular U.S. newspapers (Rozado et al., 2023). These trends align with the claims of concept creep theory regarding an increase in harm-based morality (Graham et al., 2013) and a rising cultural sensitivity to harm. Prior empirical work characterizing concept creep has focused on evaluating its horizontal expansion as increases in the semantic breadth of target concepts, evaluated as the average cosine (dis)similarity of a concept's semantic vectors. It has demonstrated that *addiction*, *bullying*, *empathy*, *harassment*, *prejudice*, *racism*, and *trauma* have broadened in recent decades (Vylomova et al., 2019; Haslam et al., 2020; Vylomova and Haslam, 2021).

Baes et al. (2023) recently developed a non-computationally intensive method to evaluate vertical creep by capturing whether a concept has come to be used in less severe contexts. It tests whether words collocating with a centre term have become less intensely negative in their connotation. Research using this new method has yielded mixed findings to date. *Trauma* came to be used in less emotionally severe semantic contexts from 1970–2019 in a corpus of psychology article abstracts (Baes et al., 2023, >133 million words). However, *anxiety* and *depression* showed the opposite trend in the abstracts corpus and in a corpus of everyday American English (Xiao et al., 2023, >500 million words). Subsequent analyses suggested a rising tendency to view these terms through a pathological (i.e., disease-related) lens in academic psychology and society at large.

4 Materials

4.1 Corpora

Two corpora enabled an analysis of shifts in the meaning of mental health-related terms in academic psychology and in the wider society. The psychology corpus contained 871,340 abstracts from 875 psychology journals, ranging from 1930 to 2019, sourced from E-Research and PubMed databases (Vylomova et al., 2019). The journal set was distributed across all subdisciplines of psychology. The final corpus of psychology abstracts was limited to 1970-2016 data due to the relatively small number of abstracts outside this period (Vylomova et al., 2019).

The general corpus combines two corpora: the Corpus of Historical American English (Davies, 2008, CoHA) and the Corpus of Contemporary American English (Davies, 2008, CoCA). CoHA

contains approx. 400 million words from 1810–2009, from 115,000 texts distributed across everyday publications (fiction, magazines, newspapers, and non-fiction books). CoCA contains approx. 560 million words from 1990–2019 from 500,000 texts (extracted from spoken language, TV shows, academic journals, fiction, magazines, newspapers, and blogs). Some CoCA texts were removed, before merging CoCA with CoHA, to prevent overlap with the psychology corpus (removing academic journal texts) and due to missing year data (blogs). The combined general corpus has previously been demonstrated to be reliable (Haslam et al., 2021b; Xiao et al., 2023). It contains magazines (36%), newspapers (31%), spoken language (16%), fiction books (10%), TV shows (7%), and non-fiction books (<1%). Pre-processing both corpora involved tokenization (lowercasing, removing punctuation and stop words) and lemmatization.

4.2 Mental Health Terms

Six mental health-related terms were chosen for analysis: *addiction*, *anger*, *distress*, *grief*, *stress*, *worry*. None of the terms are mental illnesses but all refer to common emotional or behavioral states that can be construed as abnormal or pathological. Critics have argued that concepts such as these are increasingly understood in this manner (e.g., Ridner, 2004; Wakefield and First, 2013). Figure 1 shows that all the terms increased in their relative frequency in the psychology corpus and, for some, the general corpus, making them good candidates to use as centre terms. Analyses expected to indicate they had undergone semantic inflation (vertical creep) and/or pathologization (rising association with illness-related terminology).

4.3 Warriner Norms Data

Affective meaning norms from a dataset published by Warriner et al. (2013) were used to evaluate the emotional severity of the contexts in which target mental health-related words (i.e., 'centre terms') appeared. It contains norms for valence, arousal, and dominance ratings of 13,915 English lemmas provided by 1,827 United States residents. While reading a word, participants rated how they felt on a series of scales ranging from 1 (low) to 9 (high). For the present study, only valence and arousal ratings were used. For the valence rating ($n = 723$, $M = 5.1$, $SD = 1.7$), 1 corresponded to feeling extremely "annoyed", "bored", "despaired", "melancholic", "unhappy", or, "unsatisfied", and 9 corre-
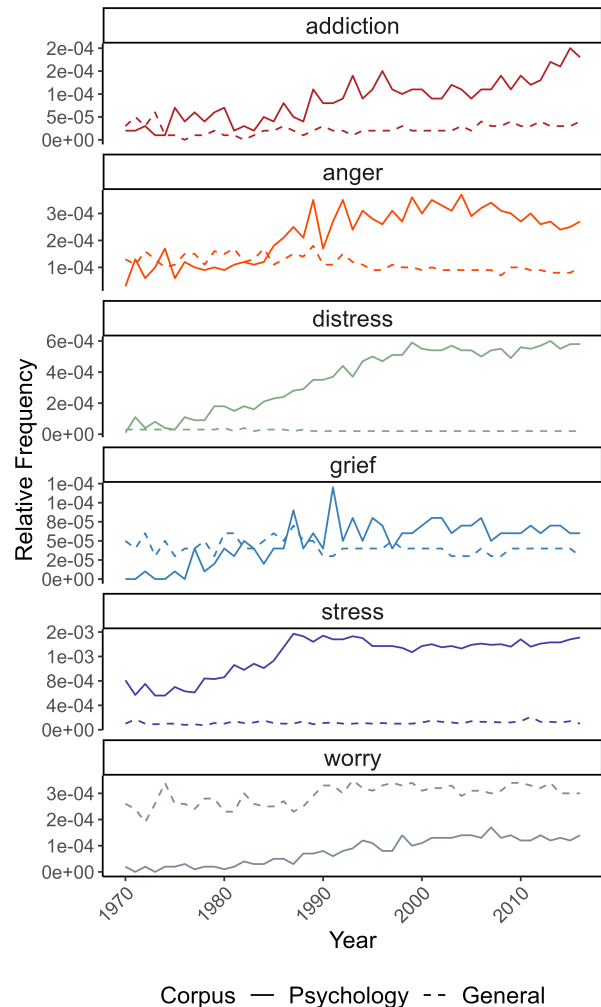

Figure 1: Relative frequencies of mental health-related terms over the study period (1970–2016).

sponded to feeling extremely "contented", "happy", "hopeful", "pleased", or "satisfied". For the arousal rating ($n = 745$, $M = 4.2$, $SD = 2.3$), 1 represented feeling "calm", "dull", "relaxed", "sleepy", "sluggish", or "unaroused", while 9 indicated feeling "agitated", "aroused", "excited", "frenzied", "jittery", "stimulated", or "wide-awake".

5 Method

5.1 Severity Index

A procedure developed by Baes et al. (2023) was used to compute an index for evaluating annual changes in the mean emotional severity of mental health-related terms. For each preprocessed corpus, collocates within a ±5-word context window of the centre terms and their annual count statistics were extracted and linked to their ratings on valence and arousal. Ratings were then summed to generate an index of emotional severity for each collocate

ranging from 2–18. The simplest approach was taken to approximate severity by summing negative valence and arousal ratings measured on the same scale. Valence ratings were reverse scored (1 = happy; 9 = unhappy), arousal ratings were not (1 = calm, 9 = aroused). Words judged as emotionally positive and calm had low scores; words judged as unpleasant and intense had high scores. The index was computed by calculating the weighted average collocate severity for each year (S), weighting the severity rating (x_i) for each collocate (n) by the number of times it appeared in the year (w_i). See the index formula below representing the mean emotional intensity of terms collocating with centre terms. For each centre term, we calculated its annual severity index as follows:

$$S = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} x_i} \quad (1)$$

The Warriner-matched dataset provided at least 80% coverage for each collocate in psychology (addiction: 82%, anger: 83%, distress: 83%, grief: 84%, stress: 80%, worry: 82%) and at least 77% coverage for each collocate in the general corpus (addiction: 81%, anger: 79%, distress: 80%, grief: 81%, stress: 81%, worry: 77%). Furthermore, in most decades, the same terms appeared among the top 10 collocates in each decade when comparing overall collocates to the Warriner-matched ones (with a 0-8% difference across all decades). See the link in the Supplementary materials for the top 10 ranked collocates in each grouping: overall collocates, Warriner-matched collocates, and non-Warriner-matched collocates.

5.2 Pathologization Index

To compute the pathologization index, a list of terms reflecting disease and illness were selected. First, six unambiguously disease-related words with restricted range in meaning (e.g., excluding "condition") were generated: "clinical", "disorder", "symptom", "illness", "pathology", "disease". Next, their 'Small World of Words' associations (De Deyne et al., 2019) were listed and duplicates were removed. See Appendix A for the final list.

Specifically, forward associations (participant responses to a cue word) for each disease-related term were documented using a web user interface. It graphs word association norms from the English Small World of Words project (SWOW-EN)[1]

[1]https://smallworldofwords.org

which contains data collected between 2011 and 2018 for 12,929 cues made by more than 90,000 fluent English speakers – making it the largest existing English-language resource.

Authors filtered the list for terms reflecting pathologization (i.e., to view or characterize as medically or psychologically abnormal), leaving 17 terms: "ailment", "clinical", "clinic", "cure", "diagnosis", "disease", "disorder", "ill", "illness", "medical", "medicine", "pathology", "prognosis", "sick", "sickness", "symptom", "treatment". Next, the collocates for each centre term were searched for the final list of 17 pathologization terms and their appearances were summed across each year before being divided by the total sum of all collocates in that year.

Figure 2 shows that the pathology-related terms appeared in both corpora, making them good candidates to test the hypothesis that pathology-related terms increasingly accompany the target terms (words representing mental-health related concepts). Furthermore, as might be expected, pathology-related terms had higher total relative frequency and rose more steeply in the psychology corpus compared to the general corpus.
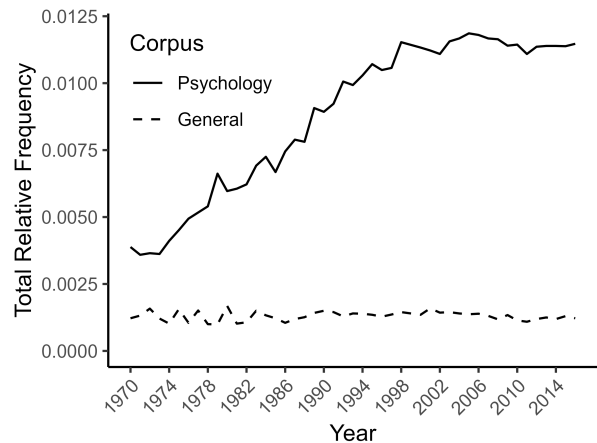

Figure 2: Total relative frequency of pathologization-related terms over the study period (1970-2016).

5.3 Analytic Strategy

Linear regression analyses were performed to test the statistical significance of the predicted trends in the first two hypotheses. An ordinary least squares estimator was used, unless autocorrelation was present (Durbin Watson test: $p < .05$), in which case the outcome variable was fit with a generalized least squares estimator to account for autocorrelated residuals.

6 Results

The linear regression models testing the hypothesized declining trend for the severity index showed some support for the concept creep hypothesis (see Figure 3). Irregularities in the data in earlier years are due to low sample size. In the psychology corpus, there was a significant declining trend in the severity of words related to two of the six concepts: *addiction* and *anger*, and a significant increasing trend in the severity of words related to *stress*. In the general corpus, there was a significant declining trend in the severity of words related to four of the six concepts: *addiction*, *grief*, *stress*, *worry*.
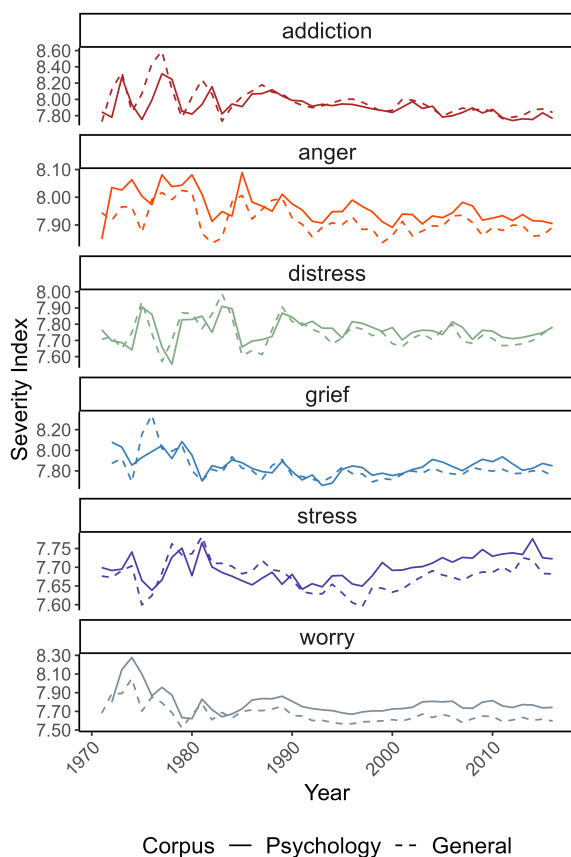

Figure 3: Severity index (3-year rolling mean) of mental health-related terms over the study period (1970-2016).

The linear regression models testing the hypothesized rising trend for the pathologization index showed some support for the hypothesis that mental health-related terms have become pathologized, as Figure 4 illustrates. Irregularities in the data in earlier years are due to low instances of pathology terms in respective context windows. In the psychology corpus, there were significant increases in the pathologization index for four of the six concepts (*addiction*, *grief*, *stress*, *worry*). For the general corpus, there was a significant increasing pathologization trend for two of the six concepts (*grief*, *stress*).
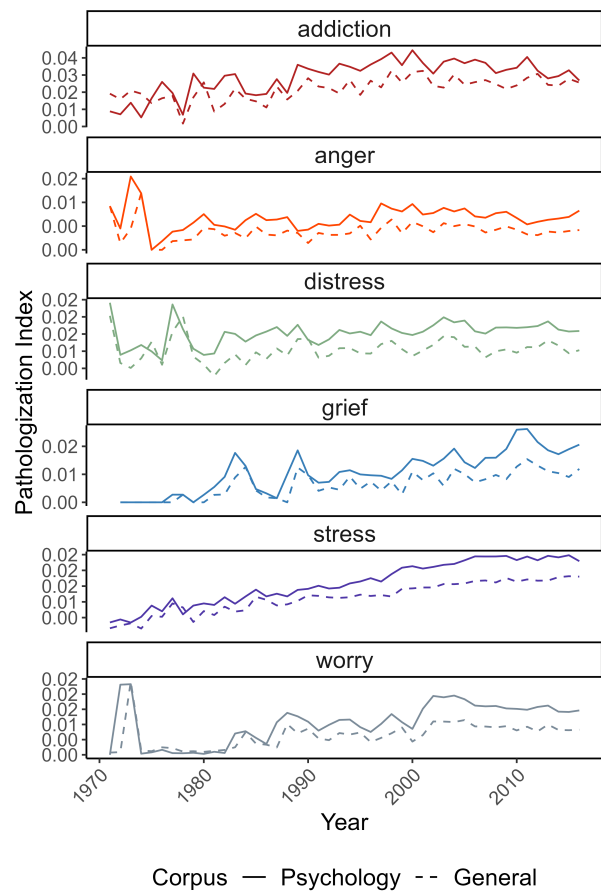

Figure 4: Pathologization index (3-year rolling mean) of mental health-related terms over the study period (1970-2016).

Control analyses were then run, holding the pathologization index and then the severity index constant (see Table 1 for predictive effects). These analyses were conducted because the two indices were partially confounded: terms in the pathologization dictionary tended to have above-average severity index scores (8.92) compared to the average severity index score of collocates for mental health concepts.[2] In the psychology corpus, holding the pathologization index constant showed significant decreases in the severity of four of the six concepts: *addiction*, *anger*, *stress*, *worry* (revealing the significant normalizing effect for worry and reversing the direction of the trend for *stress*). The

---

[2] *addiction* (psychology: 7.87; general: 8.02), *anger* (psychology: 8.00; general: 7.87), *distress* (psychology: 7.78; general: 7.73), *grief* (psychology: 7.90; general: 7.78), *stress* (psychology: 7.72; general: 7.66), *worry* (psychology: 7.91; general: 7.54)

general corpus showed the same significant effects when controlling for pathologization (with severity trends for four concepts: *addiction*, *grief*, *stress*, *worry*). Furthermore, when holding the severity index constant in the psychology corpus, the same four concepts showed significant increases in pathologization: *addiction*, *grief*, *stress*, *worry*. Similarly, in the general corpus, the same two concepts became associated with pathology-related language: *grief* and *stress*. Appendix B documents tables with summary statistics for all analyses.

Term	β(sev)	β(sev)+	β(path)	β(path)+
addict	-0.004*	-0.005*	0.0007*	0.0008*
	-0.008*	-0.008*	0.00004	-0.000002
anger	-0.002*	-0.002*	0.00002	0.000007
	-0.0009	-0.0009	0.00002	0.00002
distress	0.0007	0.004	0.0001	0.0001
	-0.002	-0.002	-0.000007	-0.000007
grief	-0.002	-0.004	0.0007*	0.0007*
	-0.005*	-0.006*	0.00006*	0.00007*
stress	0.002*	-0.001*	0.0006*	0.0006*
	-0.002*	-0.003*	0.0002*	0.0002*
worry	-0.004	-0.006*	0.0005*	0.0005*
	-0.005*	-0.005*	1.1139	0.00001

Table 1: Standardized Regression Coefficients for Year Predicting Severity Index (sev) and Pathologization Index (path) (row 1 = psychology; row 2 = COHCA). + = control analysis. * = $p < .05$. *addict* = addiction.

Post hoc correlation analyses indicated no relationship between rising pathologization and decreasing severity, except for in the psychology corpus for stress-related terminology, where there was a significant positive association (see Appendix C), likely influenced by the presence of high severity pathology terms (e.g., "disorder", "symptom") among top ranked collocates in later decades. Results indicate that the two indices rise and fall independently, apart from when (high severity) pathologization terms make up part of the severity index.

7 Discussion

The present findings support the concept creep hypothesis, which predicted that mental health concepts are increasingly used in the context of less emotionally intense language. *Addiction*, *stress*, and *worry* were normalized in this way within academic psychology and in the general corpus, whereas *anger* was only normalized in the former, and *grief* only in the latter. Only *distress* did not support the hypothesis in either corpus.

Findings also support the pathologization hypothesis, which predicted that mental health-related concepts are increasingly associated with pathology-related terminology. *Grief* and *stress* became pathologized in both academic psychology and the general corpus, whereas *addiction* and *worry* only became pathologized in the psychology corpus. *Anger* and *distress* showed no signs of pathologization.

Our findings indicate that the meanings of some mental health-related concepts have broadened to be used in less emotionally intense contexts and in more pathology-related contexts. This pattern can be observed in academic and general language. The semantic expansion of some mental health concepts in psychology (*addiction*, *anger*, *worry*) may have contributed to similar trends for these concepts and others (*grief*, *stress*) in non-academic language use. Previous work on the cultural dynamics of concept creep (Haslam et al., 2021b) demonstrates that semantic shifts of harm related concepts in psychology can influence those observed in society at large.

Trends like these have social and cultural implications. Concept creep and a rising tendency to view unpleasant emotional states through the lens of pathology may lead people to self-diagnose inappropriately and to seek unnecessary or even harmful treatments. Some evidence indicates that rates of mental illness have risen, alongside increased mental health service utilization, over-diagnosis, over-treatment, and over-prescription (Paris, 2020). Critics argue that the 'psychiatrization' of society (Haslam et al., 2021a) leads people to view ordinary problems in living as medical illnesses (Beeker et al., 2021). Growing awareness of mental ill health may be causing mental health concepts to broaden in ways that may have some benefits (e.g., reductions in stigma) but it may have significant costs as well (Foulkes and Andrews, 2023).

Conclusion

In conclusion, the findings from the present study lend support to the vertical concept creep hypothesis (Haslam, 2016) for mental health-related concepts and concerns that everyday life has become pathologized (Horwitz and Wakefield, 2007, 2012) in academic psychology and society at large. The severity index proved reliable for capturing the emotional intensity of mental health concepts, and future work will ideally apply it to pathological

concepts to explore whether they have also become normalized. The newly developed pathologization index offers a way to quantify a cultural trend, which future work can use to explore whether it influences or tracks alongside the semantic shifts of certain concepts. A key task for future research is to disentangle the relationship between vertical and horizontal forms of concept creep and to quantify their social and cultural factors in corpora representing other disciplines and languages.

Limitations

Limitations inspire future directions. First, future work could explore alternative ways of combining the valence and arousal components in the severity index using different functions. For example, it could (i) sum standardized valence and arousal ratings, rather than summing raw mean scores on the two components or (ii) weight valence more heavily than arousal when combining them, under the assumption that valence is more central to severity. Second, while the severity index tracks historical patterns in the emotional intensity of collocates, it cannot reveal which words or classes contribute to these trends. To examine underlying dimensions in the collocate data, future work could reduce the dimensions of the collocate data with a bottom up dimensionality reduction technique (e.g., k-means clustering) or top-down approach (e.g., using WordNet; Fellbaum, 2010) to capture word sense disambiguation). Third, while the severity index robustly evaluates shifts in the emotional intensity of a concept's meaning, other linguistic shifts might be at play. Future work could parse the text into syntactic dependencies and examine whether there has been a rise in the use of intensifiers, a proxy for hyperbole (Bloomfield, 1933), to modify the centre terms and examine how this relates to severity and pathologization. Finally, count-based co-occurrence methods cannot represent word meaning as comprehensively as word embeddings can, given that their bag-of-words approach disregards grammar, word order, and other contextual signals like metaphor. Future work should compare the present method to a word embeddings approach and evaluate the performance of each method.

Acknowledgements

References

Baes, N., Vylomova, E., Zyphur, M., and Haslam, N. (2023). The semantic inflation of "trauma" in psychology. *Psychology of Language and Communication*, 27(1):23–45.

Beeker, T., Mills, C., Bhugra, D., te Meerman, S., Thoma, S., Heinze, M., and von Peter, S. (2021). Psychiatrization of society: A conceptual framework and call for transdisciplinary research. *Frontiers in Psychiatry*, 12:645556.

Berger, J. and Packard, G. (2022). Using natural language processing to understand people and culture. *American Psychologist*, 77(4):525.

Bloomfield, L. (1933). *Language*. Compton Printing Works Ltd., London.

Brinkmann, S. (2016). *Diagnostic cultures: A cultural approach to the pathologization of modern life*. Routledge, New York.

Bröer, C. and Besseling, B. (2017). Sadness or depression: Making sense of low mood and the medicalization of everyday life. *Social Science & Medicine*, 183:28–36.

Charlesworth, T. E., Sanjeev, N., Hatzenbuehler, M. L., and Banaji, M. R. (2023). Identifying and predicting stereotype change in large language corpora: 72 groups, 115 years (1900–2015), and four text sources. *Journal of Personality and Social Psychology*.

Davies, M. (2008). The corpus of contemporary american english (coca). Available online at `https://www.english-corpora.org/coca/`.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., and Storms, G. (2019). The "small world of words" english word association norms for over 12,000 cue words. *Behavior research methods*, 51:987–1006.

Demszky, D., Yang, D., Yeager, D. S., Bryan, C. J., Clapper, M., Chandhok, S., Eichstaedt, J. C., Hecht, C., Jamieson, J., Johnson, M., et al. (2023). Using large language models in psychology. *Nature Reviews Psychology*, pages 1–14.

Fellbaum, C. (2010). Wordnet. In *Theory and applications of ontology: Computer applications*, pages 231–243. Springer.

Foulkes, L. and Andrews, J. L. (2023). Are mental health awareness efforts contributing to the rise in reported mental health problems? a call to test the prevalence inflation hypothesis. *New Ideas in Psychology*, 69:101010.

Furedi, F. (2004). *Therapy culture: Cultivating vulnerability in an uncertain age*. Routledge Taylor Francis, London.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., and Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in experimental social psychology*, volume 47, pages 55–130. Elsevier.

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry*, 27(1):1–17.

Haslam, N., Dakin, B. C., Fabiano, F., McGrath, M. J., Rhee, J., Vylomova, E., Weaving, M., and Wheeler, M. A. (2020). Harm inflation: Making sense of concept creep. *European Review of Social Psychology*, 31(1):254–286.

Haslam, N., Tse, J. S., and De Deyne, S. (2021a). Concept creep and psychiatrization. *Frontiers in Sociology*, 6:806147.

Haslam, N., Vylomova, E., Zyphur, M., and Kashima, Y. (2021b). The cultural dynamics of concept creep. *American Psychologist*, 76(6):1013.

Horwitz, A. V. and Wakefield, J. C. (2007). *The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder*. Oxford University Press, New York.

Horwitz, A. V. and Wakefield, J. C. (2012). *All we have to fear: Psychiatry's transformation of natural anxieties into mental disorders*. Oxford University Press, New York.

Jackson, J. C., Watts, J., List, J.-M., Puryear, C., Drabble, R., and Lindquist, K. A. (2022). From text to thought: How analyzing language can advance psychological science. *Perspectives on Psychological Science*, 17(3):805–826.

Jones, P. J. and McNally, R. J. (2022). Does broadening one's concept of trauma undermine resilience? *Psychological Trauma: Theory, Research, Practice, and Policy*, 14(S1):S131.

Kutuzov, A., Øvrelid, L., Szymanski, T., and Velldal, E. (2018). Diachronic word embeddings and semantic shifts: A survey. *arXiv preprint arXiv:1806.03537*.

Leach, S., Kitchin, A. P., and Sutton, R. M. (2023). Word embeddings reveal growing moral concern for people, animals and the environment. *British Journal of Social Psychology*.

Paris, J. (2020). *Overdiagnosis in psychiatry: how modern psychiatry lost its way while creating a diagnosis for almost all of life's misfortunes*. Oxford University Press, New York, 2nd edition.

Pennebaker, J. W. (2022). Computer-based language analysis as a paradigm shift. In Dehghani, M. and Boyd, R. L., editors, *Handbook of Language Analysis in Psychology*, pages 576–587. The Guilford Press.

Price, H. (2022). *The language of mental illness: corpus linguistics and the construction of mental illness in the press*. Cambridge University Press, Cambridge.

Ridner, S. H. (2004). Psychological distress: Concept analysis. *Journal of advanced nursing*, 45(5):536–545.

Rozado, D., Al-Gharbi, M., and Halberstadt, J. (2023). Prevalence of prejudice-denoting words in news media discourse: A chronological analysis. *Social Science computer review*, 41(1):99–122.

Tahmasebi, N., Borin, L., and Jatowt, A. (2021). Survey of computational approaches to lexical semantic change detection. In Tahmasebi, N., Borin, L., Jatowt, A., Xu, Y., and Hengchen, S., editors, *Computational approaches to semantic change*, pages 1–91. Language Science Press.

Tahmasebi, N. and Dubossarsky, H. (2023). Computational modeling of semantic change. *arXiv preprint arXiv:2304.06337*.

Tse, J. S. and Haslam, N. (2021). Inclusiveness of the concept of mental disorder and differences in help-seeking between asian and white americans. *Frontiers in Psychology*, 12:699750.

Vylomova, E. and Haslam, N. (2021). Semantic changes in harm-related concepts in english. In Tahmasebi, N., Borin, L., Jatowt, A., Xu, Y., and Hengchen, S., editors, *Computational approaches to semantic change*, pages 93–121. Language Science Press.

Vylomova, E., Murphy, S., and Haslam, N. (2019). Evaluation of semantic change of harm-related concepts in psychology. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*, pages 29–34.

Wakefield, J. C. (2010). Misdiagnosing normality: Psychiatry's failure to address the problem of false positive diagnoses of mental disorder in a changing professional environment. *Journal of Mental Health*, 19(4):337–351.

Wakefield, J. C. and First, M. B. (2013). Clarifying the boundary between normality and disorder: a fundamental conceptual challenge for psychiatry.

Warriner, A. B., Kuperman, V., and Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior Research Methods*, 45(4):1191–1207.

Wheeler, M. A., McGrath, M. J., and Haslam, N. (2019). Twentieth century morality: The rise and fall of moral concepts from 1900 to 2007. *PLOS ONE*, 14(2):e0212267.

Xiao, Y., Baes, N., Vylomova, E., and Haslam, N. (2023). Have the concepts of 'anxiety' and 'depression' been normalized or pathologized? a corpus study of historical semantic change. *PLOS ONE*, 18(6):e0288027.

A Appendix A

'Small World of Words' Associations for selected terms ("clinical", "disorder", "symptom", "illness", "pathology", "disease")": "ailment", "aids", "anxiety", "bad", "bacteria", "bed", "bipolar", "cancer", "cause", "clean", "clinical", "cold", "conduct", "contagious", "cough", "cure", "death", "diagnosis", "disease", "doctor", "epidemic", "fever", "green", "hospital", "ill", "illness", "infection", "list", "malaria", "medical", "medicine", "mental", "pain", "panic", "physical", "plague", "precise", "problem", "psychology", "research", "sad", "sickness", "sick", "surgery", "symptom", "tired", "treatment", "unhealthy", "unwell", "virus", "vomit", "white", "study", "sterile".

B Appendix B

Term	β	p	SE	Accuracy
addict	-0.004	.043	0.002	0.07
	-0.008	.010	-0.008	0.12
anger	-0.002	.034	0.001	0.08
	-0.0009	.329	.0009	-0.0005
distress	0.0007	.510	0.001	-0.01
	-0.002	.130	0.002	0.03
grief	-0.002	.459	0.003	-0.01
	-0.005	.005	0.002	0.15
stress	0.002	<.001	0.0005	-134.48*
	-0.002	.034	0.0008	-89.18*
worry	-0.004	.129	0.002	0.03
	-0.005	<.001	0.0007	0.48

Table 2: Regression Summary Statistics for Year Predicting Severity Index (row 1 = psychology; row 2 = general). Accuracy (model fit) = Adj. R^2. * = AIC. *addict* = addiction.

Term	β	p	SE	Accuracy
addict	0.0007	<.001	0.0001	-236.57*
	0.00004	.704	0.0001	-0.02
anger	0.00002	.757	0.00006	-0.02
	0.00002	.078	0.000009	-491.49*
distress	0.0001	.142	0.00008	0.03
	-0.000007	.905	0.00006	-0.02
grief	0.0007	<.001	0.00008	0.60
	0.00006	.002	0.00002	0.17
stress	0.0006	<.001	0.00003	-386.27*
	0.0002	<.001	0.00003	0.43
worry	0.0005	.0003	0.0001	0.24
	1.11	.919	0.00001	-473.42*

Table 3: Regression Summary Statistics for Year Predicting Pathologization Index (row 1 = psychology; row 2 = general). Accuracy (model fit) = Adj. R^2. * = AIC. *addict* = addiction.

Term	β year	p year	β path	p path	Accuracy
addict	-0.005	.04	1.33	.491	0.06
	-0.008	.012	-5.90	.166	0.14
anger	-0.002	.038	-3.09	.206	0.89
	-0.0009	.347	0.52	.975	-0.02
distress	0.004	.695	2.38	.247	-0.004
	-0.002	.135	0.02	.996	0.007
grief	-0.004	.333	3.20	.514	-0.02
	-0.006	.002	15.41	.244	0.15
stress	-0.001	.035	6.03	.012	-142.49*
	-0.003	.007	6.03	.094	-94.42*
worry	-0.006	.030	5.23	.082	0.08
	-0.005	<.001	9.74	.303	0.48

Table 4: Regression Summary Statistics for Year Predicting Severity Index, controlling for Pathologization (row 1 = psychology; row 2 = general). Accuracy (model fit) = Adj. R^2. * = AIC. *addict* = addiction. *stress* (psychology)** = independent variables have VIF: 11.91.

Term	β year	p year	β path	p path	Accuracy
addict	0.0008	<.001	0.008	.491	-228.00*
	0.00002	.871	-0.007	.166	0.003
anger	0.000007	.921	-1.17	.206	-0.006
	0.00002	.084	0.00005	.975	0.03
distress	0.0001	.175	0.012	.247	0.03
	-0.000007	.909	0.00002	.996	-0.05
grief	0.0007	<.001	0.004	.514	0.577
	0.00007	.001	0.002	.244	0.18
stress	0.0006	<.001	0.02	.012	-383.09*
	0.0002	<.001	0.01	.094	0.45
worry	0.0005	<.001	0.013	.082	0.25
	0.00001	.425	0.002	.303	-462.26*

Table 5: Regression Summary Statistics for Year Predicting Pathologization Index, controlling for Severity (row 1 = psychology; row 2 = general). Accuracy (model fit) = Adj. R^2. * = AIC. *addict* = addiction.

C Appendix C

Term	r	df	p
addiction	-0.09	45	.543
	-0.21	45	.149
anger	-0.19	45	.199
	-0.03	45	.824
distress	0.19	45	.199
	0.005	45	.974
grief	-0.02	42	.881
	-0.03	45	.828
stress	0.66	45	<.001
	-0.03	45	.851
worry	0.11	44	.451
	0.09	45	.505

Table 6: Correlation Statistics for Severity Index and Pathologization Index (row 1 = psychology; row 2 = general).

D Supplementary Material

The data and code are available at the following repository link: https://osf.io/hbzmc/?view_only=f6e3d36f89204eae9c6ecaa501f0015e

Access the spreadsheet with tables of the highest-ranked collocates for concepts related to mental health at the following link: https://osf.io/hbzmc/files/osfstorage/652ddf3728274506adb867cc