Scientometric Analysis of Psychology and Related Fields

Himadri Mandal (BS2327)

B.Stat II Year, Indian Statistical Institute, Kolkata

Under the guidance of Dr. Garga Chatterjee and Dr. Kuntal Ghosh

May 2025

Outline

Abstract

- Scalable, data-driven bibliometric analysis of psychology (2000–2025).
- Previous work: [1]
- Focus on author-assigned keywords in top 20 h5-index SCOPUS journals
- Metrics: keyword frequency (popularity), citation counts (impact), impact-density
- Highlight divergences between popularity and impact rankings to reveal emerging areas

Introduction

- Interdisciplinary evolution shaped by technology and computational methods
- Challenges: manual review infeasible at scale
- Research questions:
  1. Most frequent keywords over time
  2. Most impactful keywords over time
  3. Highest impact density topics
  4. Divergences between popularity and impact rankings

Data Sources

- Explored: Semantic Scholar, OpenAlex, Web of Science
- Chose SCOPUS: access available, exports limited to 20,000 entries
- Focused on top 20 journals by h5-index (March 2025)


```
================================================
🅼  Journals 🅼
================================================

▶ Frontiers in Psychology
▶ Trends in Cognitive Sciences
▶ Personality and Individual Differences
▶ Current Opinion in Psychology
▶ Perspectives on Psychological Science
▶ Clinical Psychology Review
▶ Journal of Applied Psychology
▶ American Psychologist
▶ Current Psychology
▶ Journal of Vocational Behavior
▶ Journal of Personality and Social Psychology
▶ Annual Review of Psychology
▶ Psychological Bulletin
▶ Child Development
▶ Journal of Child Psychology and Psychiatry
▶ Psychological Science
▶ Educational Psychology Review
▶ Addictive Behaviors
▶ Psychological Trauma: Theory, Research, Practice, and Policy
▶ European Child & Adolescent Psychiatry
```

Figure: Selected Journals

Data Extraction and Cleanup

- Extracted author keywords via Pandas
- Normalization: remove parentheticals, punctuation, collapse whitespace, title case
- Timeline: 2000–2025

Dataset Size

Size before dropping missing keywords: 124,512
Size after dropping missing keywords: 87,352

Keyword Normalization

```python
import re

def normalize_keyword(kw: str) -> str:
    kw = kw.lower().strip()
    # remove parentheticals, e.g. "(CSA)" → ""
    kw = re.sub(r"\(.*?\)", "", kw)
    # remove punctuation
    kw = re.sub(r"[^\w\s]", "", kw)
    # collapse whitespace
    kw = re.sub(r"\s+", " ", kw)
    return kw
```
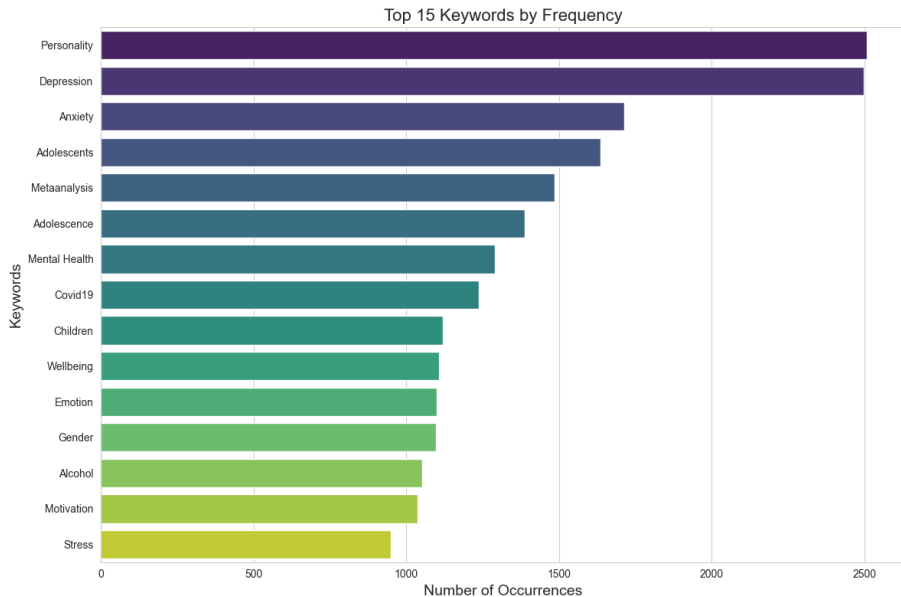
Figure: Keyword Normalization Pipeline

Popularity and Impact Metrics

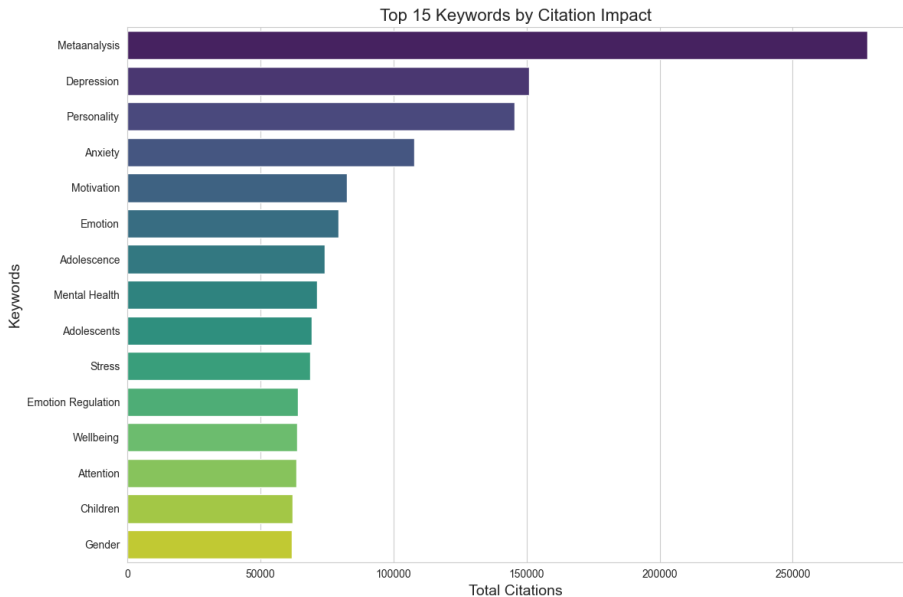**Popularity** ($N(Y,K)$): Count of keyword K in year Y.s
Impact ($c(K,Y)$): Total citations of papers with keyword K in year Y.
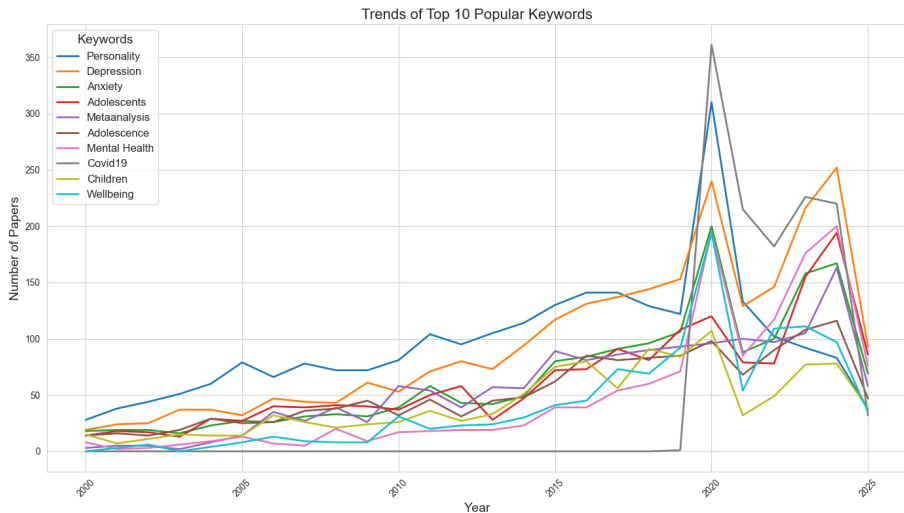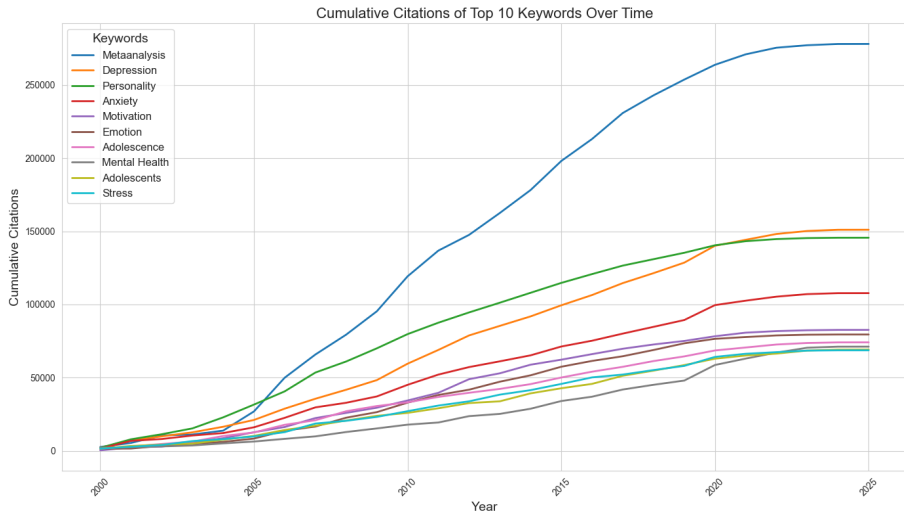**Impact Density**: $\sum_Y c(K,Y)/\sum_Y N(K,Y)$

Most Popular Keywords

Top 15 Keywords by Frequency

Most Impactful Keywords

Top 15 Keywords by Citation Impact

Trends of Top 10 Popular Keywords

Impact Trends

Cumulative Citations of Top 10 Keywords Over Time

Popularity vs Impact Rankings

Keyword	Impact Rank	Popularity Rank	Occ.
Metaanalysis	#1	#5	1,486
Depression	#2	#2	2,498
Personality	#3	#1	2,508
Anxiety	#4	#3	1,713
Motivation	#5	#14	1,037
Emotion	#6	#11	1,100
Adolescence	#7	#6	1,387
Mental Health	#8	#7	1,290
Adolescents	#9	#4	1,635
Stress	#10	#15	949

Popularity vs Impact

We wanted to find keywords that have high impact but rank low in popularity rankings. This meant studying

$$\max_{K \in \mathcal{S}} \left(\text{rank}_{\text{impact}}(K) - \text{rank}_{\text{popularity}}(K) \right).$$

We fixed $\mathcal{S}$ to be the top 100 most impactful keywords. Then, we find the top 25 keywords which have the highest deviations.

Deviation Between Rankings

Keyword	Deviation	Impact Rank	Popularity Rank
Review	102	44	146
Job Performance	100	67	167
Power	75	100	175
Positive Psychology	71	90	161
Epidemiology	69	75	144
…			

Table: Top Deviations

Impact Density Rankings

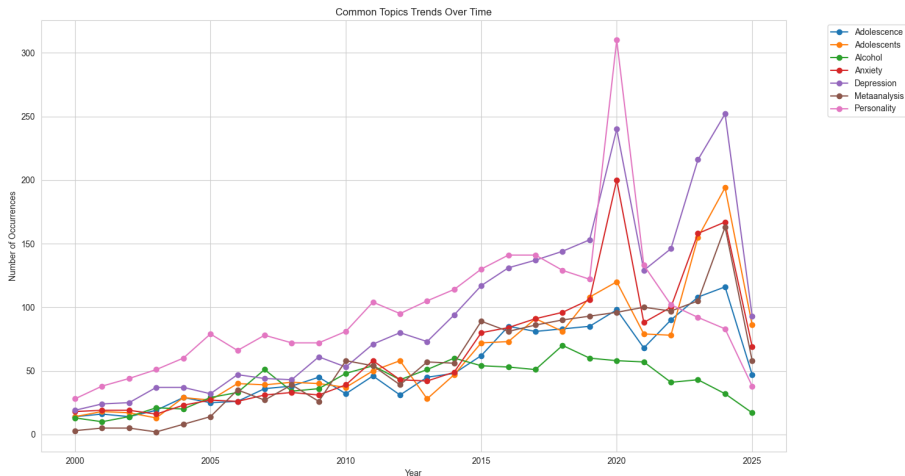Top 15 Keywords by Citations per Paper
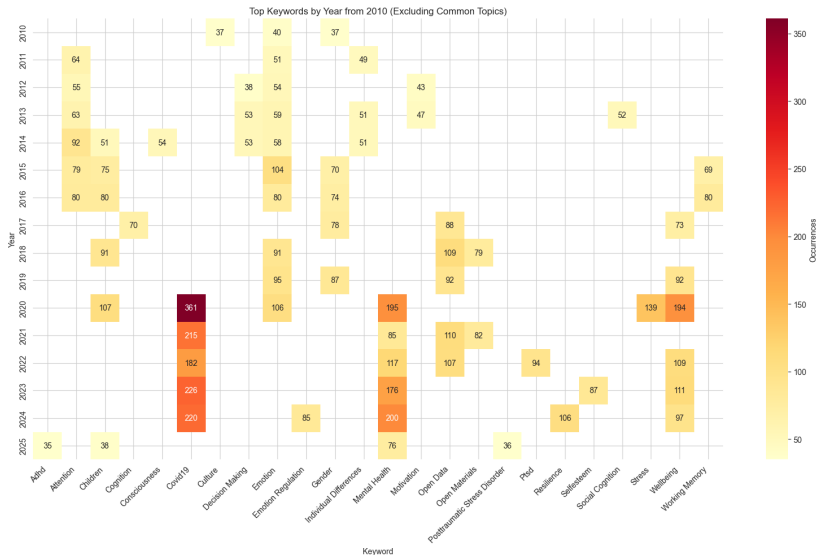(Minimum 25 papers)
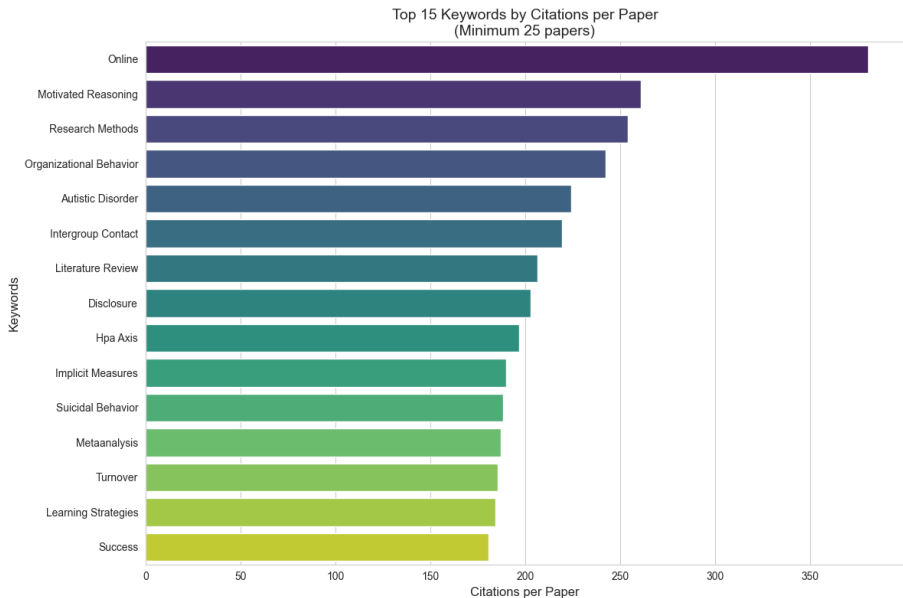
Common Keywords

Filtering keywords that appear in 75% of years **AND** are in top 10 for at least 50% of years

- Adolescence
- Adolescents
- Alcohol
- Anxiety
- Depression
- Metaanalysis
- Personality

Common Keywords Trends

Common Topics Trends Over Time

Top Keywords Excluding Common

Top Keywords by Year from 2010 (Excluding Common Topics)

Conclusion

- Enduring prominence: Personality, Depression, Anxiety
- High impact vs frequency: Meta-analysis, Positive Psychology
- Divergences reveal emergent areas: Burnout, Intervention, Well-being
- Methodology replicable and scalable

Acknowledgements

I sincerely thank **Dr. Garga Chatterjee** and **Dr. Kuntal Ghosh** for their invaluable guidance throughout the project, and for curating a very enjoyable course.

References

[1] Oliver Wieczorek et al. "Mapping the field of psychology: Trends in research topics 1995–2015". en. In: *Scientometrics* 126 (June 2021), pp. 9699–9731. ISSN: 0138-9130. DOI: 10.1007/s11192-021-04069-9. URL: https://doi.org/10.1007/s11192-021-04069-9.