Universität Marburg

Introducing openESM: A database of openly available experience sampling datasets

Björn Siepe[1]

October 18th, 2025 – Talk at the Emotional Disorder Genesis and Evolution (EDGE) Lab Meeting

[1]Psychological Methods Lab, Department of Psychology, Philipps-Universität Marburg

Licensed under CC-BY 4.0

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

Your research question:

- Does this generalize?
- What about heterogeneity?
- How robust is this finding?

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

Your research question:

- Does this generalize?
- What about heterogeneity?
- How robust is this finding?

Your options:

1. Run 10 more studies (years)

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

Your research question:

- Does this generalize?
- What about heterogeneity?
- How robust is this finding?

Your options:

1. Run 10 more studies (years)
2. Email 50 researchers for data (crickets)

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

Your research question:

- Does this generalize?
- What about heterogeneity?
- How robust is this finding?

Your options:

1. Run 10 more studies (years)
2. Email 50 researchers for data (crickets)
3. Harmonize datasets manually (months)

ESM Research: A Scale Problem

Typical ESM Study:

- 100-200 participants
- 1,000-10,000 observations
- multiple months to collect and clean

Your research question:

- Does this generalize?
- What about heterogeneity?
- How robust is this finding?

Your options:

1. Run 10 more studies (years)
2. Email 50 researchers for data (crickets)
3. Harmonize datasets manually (months)
4. Use openESM (minutes)

Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

and packages for easy access

[1] R logo: https://www.r-project.org/logo/

[2] Python logo: https://www.python.org/community/logos/

Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

and packages for easy access

Substantive Research

Test theories across contexts

---

[1] R logo: https://www.r-project.org/logo/

[2] Python logo: https://www.python.org/community/logos/

Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

and packages for easy access

Substantive Research

Test theories across contexts

Design Research

Optimize measurement choices

[1] R logo: https://www.r-project.org/logo/

[2] Python logo: https://www.python.org/community/logos/

Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

and packages for easy access

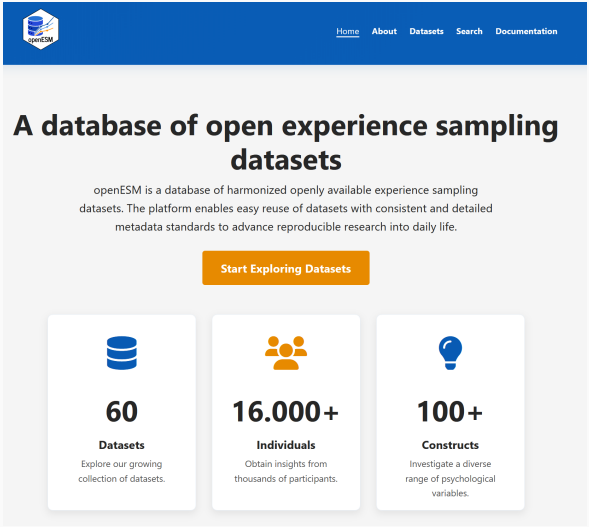
Substantive Research
Test theories across contexts

Design Research
Optimize measurement choices

Statistical Methods
Benchmark and develop methods

[1] R logo: https://www.r-project.org/logo/
[2] Python logo: https://www.python.org/community/logos/

Live Demo


openESM
Home
About
Datasets
Search
Documentation
A database of open experience sampling datasets
openESM is a database of harmonized openly available experience sampling datasets. The platform enables easy reuse of datasets with consistent and detailed metadata standards to advance reproducible research into daily life.
Start Exploring Datasets
60
Datasets
Explore our growing collection of datasets.
16.000+
Individuals
Obtain insights from thousands of participants.
100+
Constructs
Investigate a diverse range of psychological variables.

Design Process

Design Process

1. Cleaning

- Harmonized NAs
- Removed redundant columns
- Standardized ID, beep, day

Design Process

1. Cleaning

- Harmonized NAs
- Removed redundant columns
- Standardized ID, beep, day

2. Harmonization

- Same variable names for same questions
- E.g., “depressed” across datasets

Design Process

1. Cleaning

- Harmonized NAs
- Removed redundant columns
- Standardized ID, beep, day

2. Harmonization

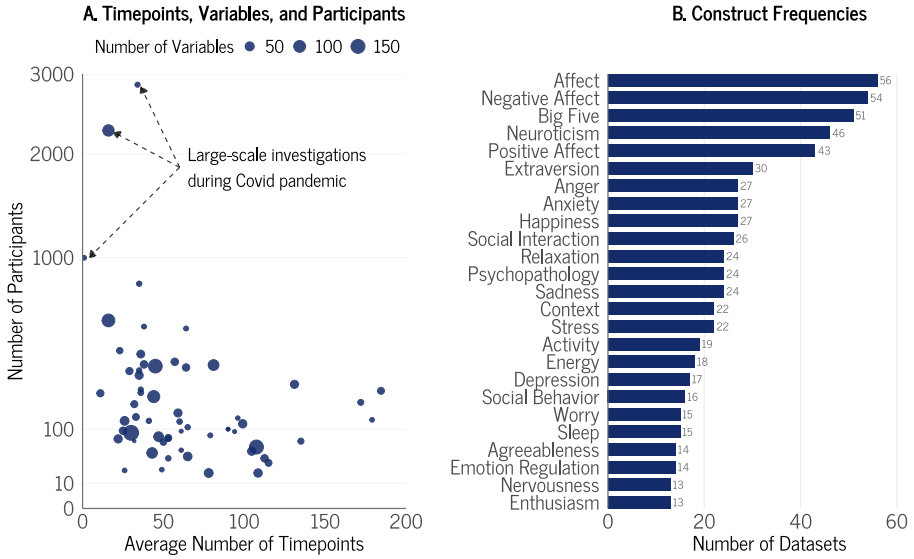
- Same variable names for same questions
- E.g., “depressed” across datasets

3. Annotation

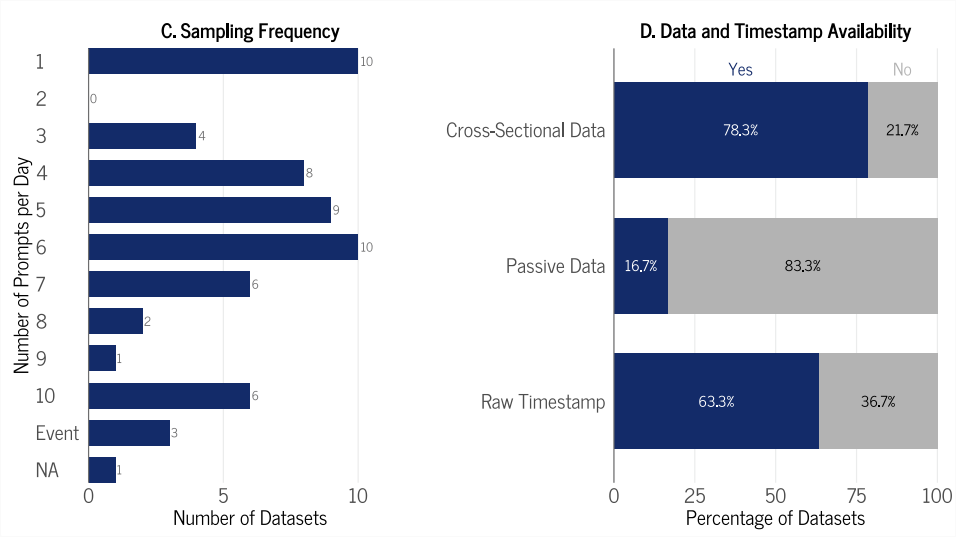
- Hierarchical construct tags
- E.g., “angry” $\rightarrow$ anger $\rightarrow$ negative affect $\rightarrow$ affect

What’s inside

What's inside

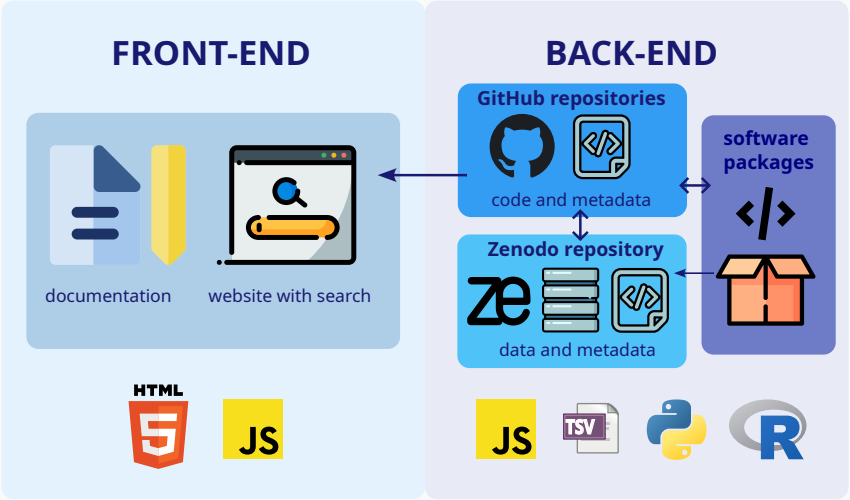


What's inside



Architecture Behind It

Architecture Behind It



Example Analysis

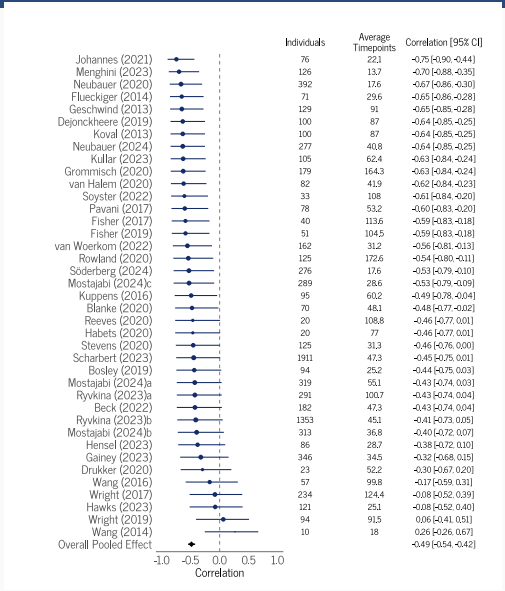
- **Question:** Within-person correlation of positive and negative affect

Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations

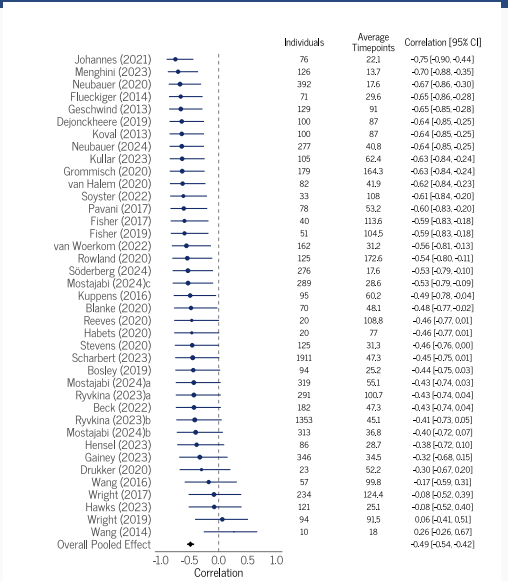
Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled $r = -0.49$ [95% CI: $-0.54, -0.42$]



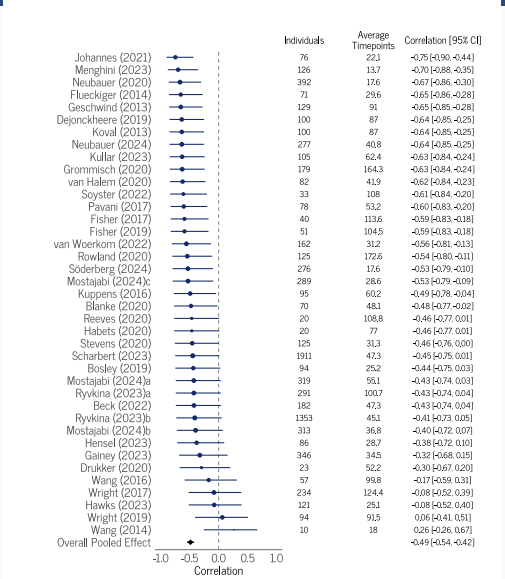
Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled $r = -0.49$ [95% CI: $-0.54, -0.42$]
- **Design insight:** More prompts/day → stronger negative correlation



Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled $r = -0.49$ [95% CI: $-0.54, -0.42$]
- **Design insight:** More prompts/day → stronger negative correlation
- **Robustness:** Stable across 12 preprocessing/modeling choices



Next Steps

- Community Involvement: Get people to help with data, more non-WEIRD experience sampling data
- Extending Database: Data and Metadata
- Extending Software: Benchmarking, more filtering options

Goal: turn openESM into a continuously evolving community resource for cumulative experience sampling research

Get In Touch

- bjoern.siepe@uni-marburg.de
- https://bsiepe.github.io/

Paper & Slides

References i