

ExpiWell

ESM WORKSHOP SERIES · PART ONE

Setting Up the *Study*

Deciding whether ESM fits your question, then setting the parameters that shape everything downstream.

By Louis Tay and Marilyn Uy

expiwell.com

Setting Up the Study

1

Decide whether ESM is what you need

Is the phenomenon episodic? Can your participants report in the moment? How much control do you need over the situation?

2

Set study parameters

Protocol type, sampling periods, and survey length — the three decisions that define the burden you place on participants.

STEP ONE

Decide whether ESM is what you need

- For episodic (experiential) as opposed to semantic (conceptual).
- Retrospective reporting may be more useful for predicting intentions than aggregated online assessments.
- Target participants who are willing and able to provide momentary reports.

CONTROL OVER THE SITUATION

● More control needed

Lab

● Natural environment needed

ESM

Set study parameters

1

Protocol type

How reports are triggered — by interval, signal, event, or context.

2

Sampling periods

How many observations you need for a stable within-person estimate.

3

Survey length

How much you can ask before compliance starts to fall away.

Four ways to trigger a report

● Interval contingent

e.g., every 10am and 6pm; at the end of every day

Least burdensome. In ExpiWell: a time-based Calendar or Rolling schedule.

● Signal contingent

e.g., when signaled by a notification — scheduling can be pure random or stratified within a block

Most burdensome. In ExpiWell: random signals inside a time slot you define.

● Event contingent

e.g., after every interaction with your direct supervisor lasting more than 10 minutes

Define what counts as an event. In ExpiWell: participants open the survey themselves.

● Context contingent

e.g., based on location or a specific event

ExpiWell can deliver a survey when a participant is at, enters or exits a defined area (GEMA license).

The decision rests on three things

1

Prevalence of the target behavior or focal event

Rare events push you toward event- or context-contingent designs.

2

Susceptibility of the phenomenon to retrospective bias

The more memory distorts it, the closer to the moment you need to be.

3

Participant response burden

Every gain in fidelity is paid for in participant effort.

Signal and time contingent: framing the window

Consider the temporal framing of assessments and the time frame of the questions themselves.



Immediate / online

“What are you doing right now?”

- Concern about recall biases
- May not capture specific behaviors when the signal occurs



Short-term retrospective

“What were you doing in the last two hours?”

- Concern about capture of behaviors when the signal occurs



Prospective

“What do you intend to do in the next two hours?”

- Less concern about recall bias

How many observations do you need?

The number of observations needed for a stable estimate of a phenomenon within each person.

- Run a multilevel power analysis (Maas & Hox, 2005; Snijders & Bosker, 1999); depends on level of analysis (within- vs between-) and effect size.
- Statistical power for cross-level interaction: hermanaguinis.com/crosslevel.html
- If you want to increase power, it is better to add people than observations (Bolger, Stadler, & Laurenceau, 2012).

Around 50 level-2 units is typical of most ESM-based organizational studies.

GUIDELINE FOR CROSS-LEVEL INTERACTION (LEVEL 2 MODERATOR)

● 30	people	with 30 observations
● 60	people	with 25 observations
● 150	people	with 5 observations

Kreft & DeLeeuw, 1998

Additional references

1

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research.

Journal of Social and Personal Relationships, 35, 7–31.

2

Lane, S. P., & Hennes, E. P. (2019). Conducting sensitivity analyses to identify and buffer power vulnerabilities in studies examining substance use over time.

Addictive Behaviors, 94, 117–123.

3

Companion materials and code.

osf.io/ez2rm

Keep it brief — and expect to pay for length



< 2

minutes

Ideal length of a single survey.
Shortened scales are typically used.



15–20

min / day

Daily response burden, or less
where possible.



50–95

% response

Typical range. Can run higher in
some fields, such as clinical
psychology.

Participant burden and anticipated compliance should be estimated before launch, not discovered during it.

Cadence and duration

- Naturalistic incidence of events and states should guide the schedule.
- Enough to capture fluctuations — ideally no more than 6 times per day, and no more than 3 weeks (Delespaul, 1992).
- “Breaks” can be given between intensive longitudinal assessments.
- There are studies that go on for months.

● 6×

per day, at most

● 3

weeks, at most

Delespaul, 1992

Three schedule types carry these parameters

Whichever you pick, surveys are scheduled in each participant's own time zone, app notifications are sent automatically, and a participant's surveys are linked across the study without an ID question.



Calendar

Fixed dates, same for everyone

- For studies tied to a real-world date
- Time-, signal- and event-contingent variants



Rolling

The clock starts when a participant joins

- For when start dates cannot be aligned
- Time-, signal- and event-contingent variants



Static

No dates at all

- Surveys stay open across a period
- One-off and event-contingent

ExpiWell

NEXT IN THE SERIES

Part Two: Implement the Study

Pilot, evaluate participant experience, build onboarding materials, then run at full scale.

By Louis Tay and Marilyn Uy

expiwell.com