

Comunicación durante el curso

Para dudas, compartir contenido interesante, y catársis sobre interfaces mal diseñadas

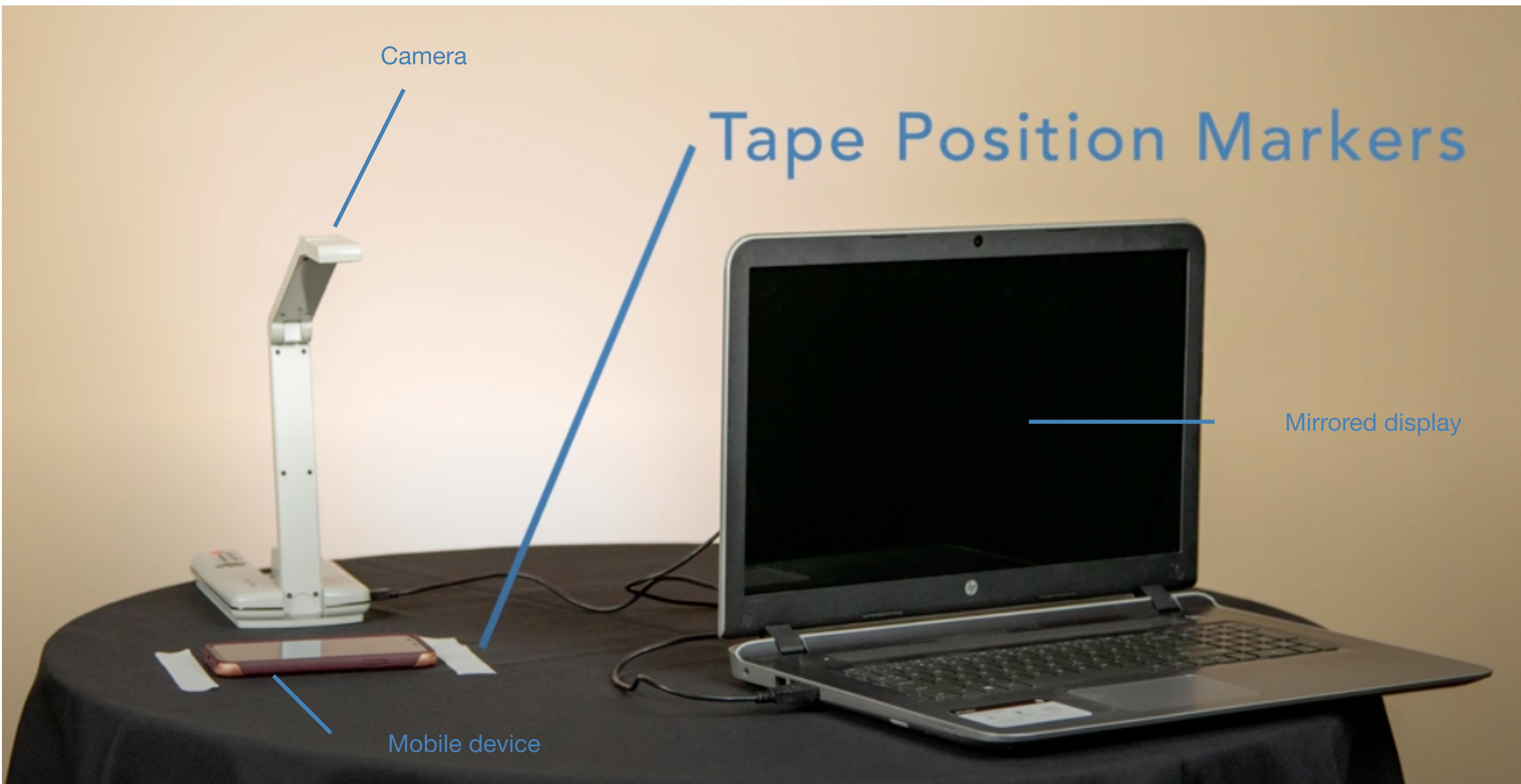


Grupo de Slack de SIGCHI LAIHC
Latin American HCI Community.
Únanse al channel **#eci-uba-2024**
para mantenernos comunicados.

[usability test + think aloud]

[cómo operacionalizar? Enfocarse en
tiempo de completitud + numero de
errores]

Tests de usabilidad



<https://www.youtube.com/watch?v=NdqTbpyvalg>

Think Aloud

Pensar en voz alta

Método de evaluación simple que nos permite entender mejor el modelo mental del usuario.

- 1) Se define una lista de tareas a evaluar en el sistema
- 2) Se le pide al participante pensar en voz alta mientras realiza cada tarea.
- 3) Se anotan todos los problemas de usabilidad que surjan durante la tarea, o cualquier comentario que resulte interesante o sorprendente.

Roles:

- Un facilitador: le da instrucciones al participante, y se enfoca en que siga hablando todo el tiempo. Si el participante deja de hablar, se le puede decir “En qué estás pensando?”.
- Una o más personas toman notas, en silencio.

NASA TLX

Un cuestionario estandarizado para medir el nivel de esfuerzo en una tarea. Se puede usar para cuantificar distintos aspectos de qué tan difícil de usar es un sistema, en tests de usabilidad y también experimentos, pero también para ayudarle a los participantes a reflexionar sobre su experiencia y complementar con una entrevista post-tarea.

Demandada Mental

¿Cuánto esfuerzo mental requirió la tarea? ¿Fue fácil o exigente, simple o compleja?

Demandada Física

¿Cuánto esfuerzo físico requirió la tarea? ¿Fue fácil o exigente, relajada o agotadora?

Demandada Temporal

¿Cuánta presión de tiempo sentiste debido al ritmo al que ocurrieron las tareas o elementos de la tarea? ¿Fue lento o rápido?

Rendimiento Propio

¿Qué tan bien realizaste la tarea? ¿Qué tan satisfecho estás con tu rendimiento?

Esfuerzo

¿Cuánto tuviste que trabajar (mental y físicamente) para lograr tu nivel de rendimiento?

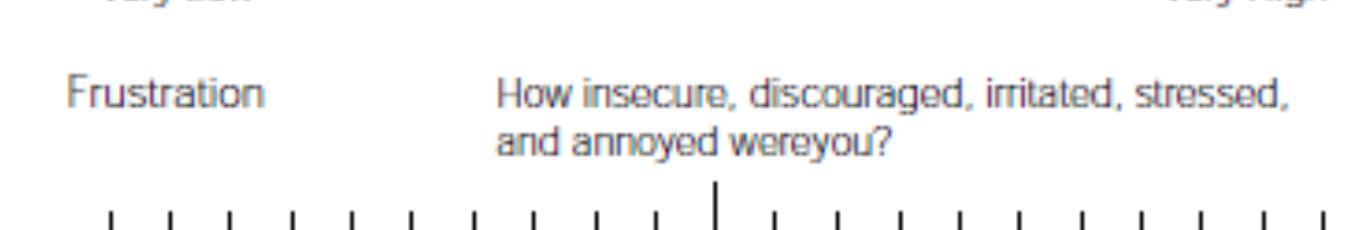
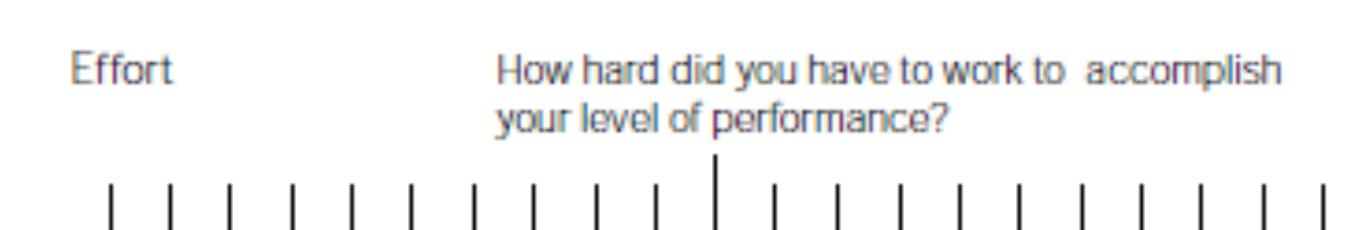
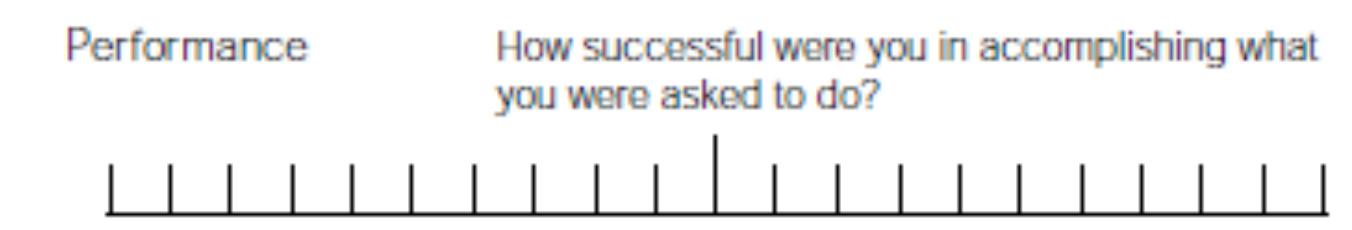
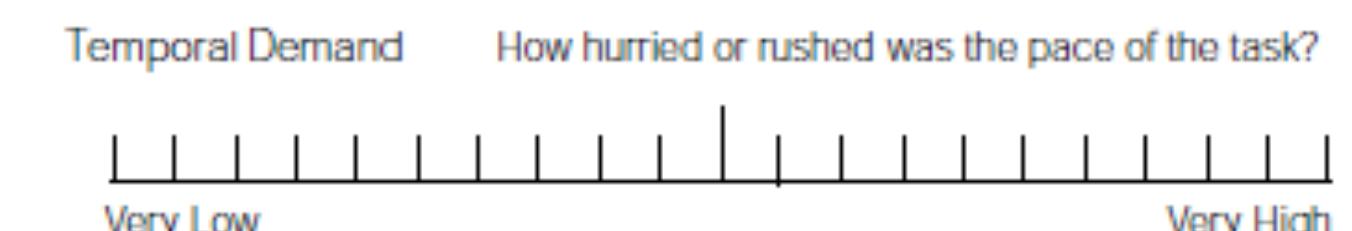
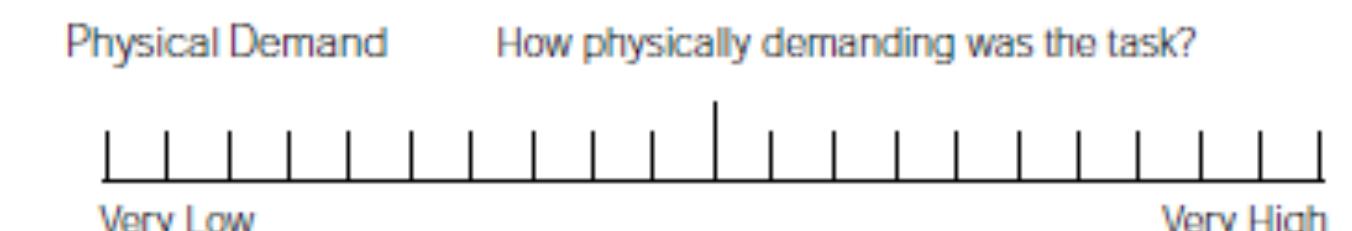
Nivel de Frustración

¿Cómo te sentiste durante la tarea? Qué tanto te hizo dudar, sentir desmotivación o estrés?

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
------	------	------



Otros cuestionarios estandarizados

Ojo: para que estos cuestionarios sirvan, tienen que estar alineados con el objetivo de la evaluación.

- Copy AttrakDiff: measuring the attractiveness [Hassenzahl et al. 2003]. [Attrakdiff3](#) is used here.
- Copy GEQ - Game Experience Questionnaire: [Poels et al. 2007]. Note there are different versions of GEQ exist and the factor structure of the GEQ might not be stable [Law et al. 2018]--please be cautious when applying it. We used the original version with 42 items here.
- Link FSS - Flow Short State: measuring the flow-inducing aspects of the gameplay. The longer form (Flow State Scale) can be found in this work [Jackson and Marsh, 1996]
- Copy IPQ - Igroup Presence Questionnaire: measuring the sense of presence experienced in a virtual environment [Schubert et al. 2001]. The questionnaire follows [this](#) webpage.
- Copy PQ - Presence Questionnaire: measuring the degree to which individuals experience presence in a virtual environment [Witmer 1998].
- Copy ITQ - Immersive Tendencies Questionnaire: to measure differences in the tendencies of individuals to experience presence [Witmer 1998].
- Copy SUS (Usability) - System Usability Scale: measuring system usability [Brooke 1996].
- Copy UEQ - User Experience Questionnaire: measuring the user experience of interactive products [Laugwitz 2008].

Qué pasa afuera del
laboratorio?

Qué pasa afuera del laboratorio?



Lab Vida real

Controlamos el contexto

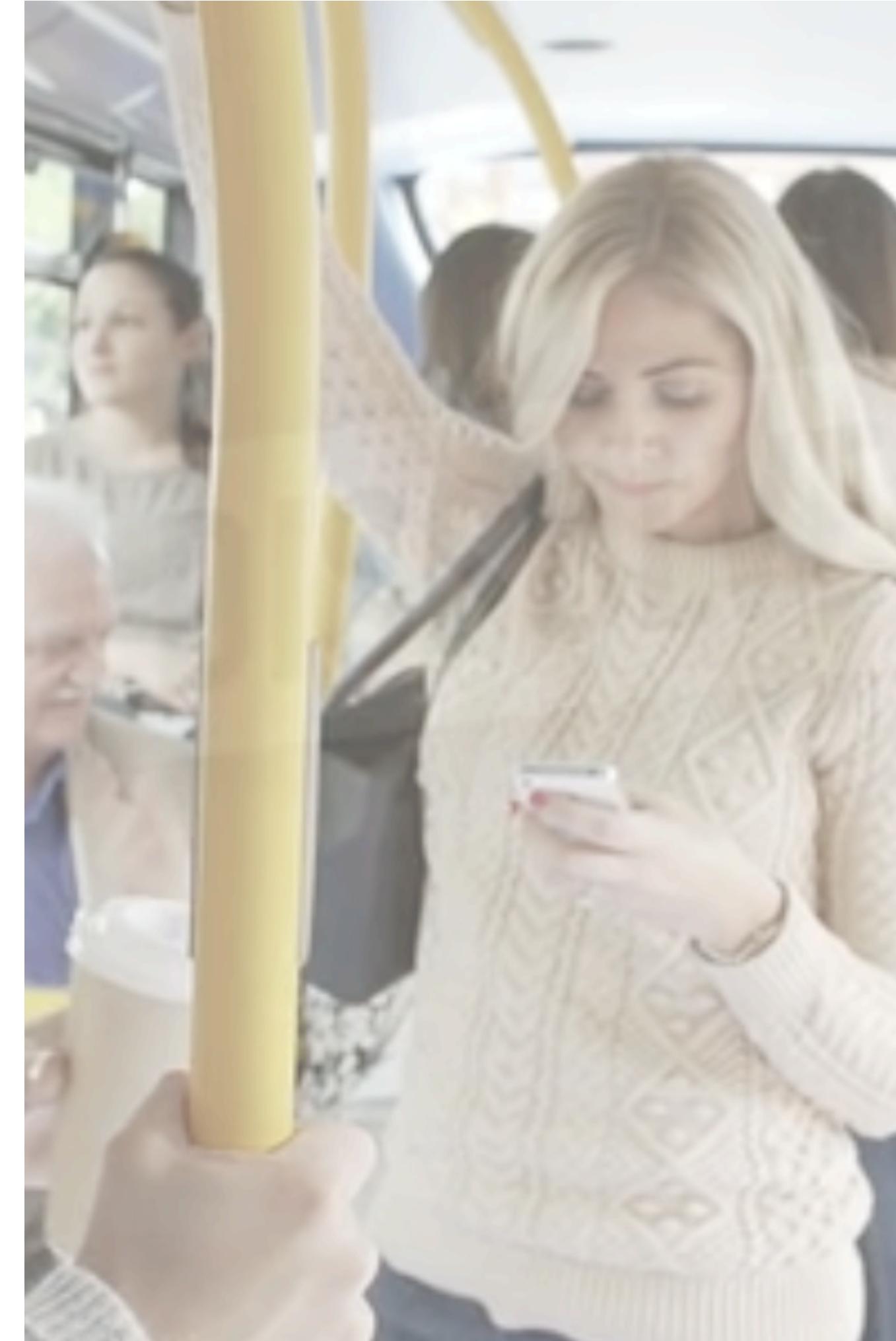
El sistema se usa
en aislamiento

Tareas pre-definidas

El contexto puede
variar inesperadamente

El sistema se usa en
una ecología de
artefactos

Acción situada (los
planes cambian,
y los objetivos se
resuelven
con lo que hay a mano)



Validez Ecológica (Ecological Validity)

Qué tan probable es que los resultados del estudio en el laboratorio se reflejen en el mundo real?

Estudios de campo (Field Studies)

Se realizan en “condiciones reales”, en el mundo real

Más control



Menos control

Los participantes usan dispositivos provistos por los investigadores

Las tareas del estudio están pre-definidas y tienen instrucciones detalladas

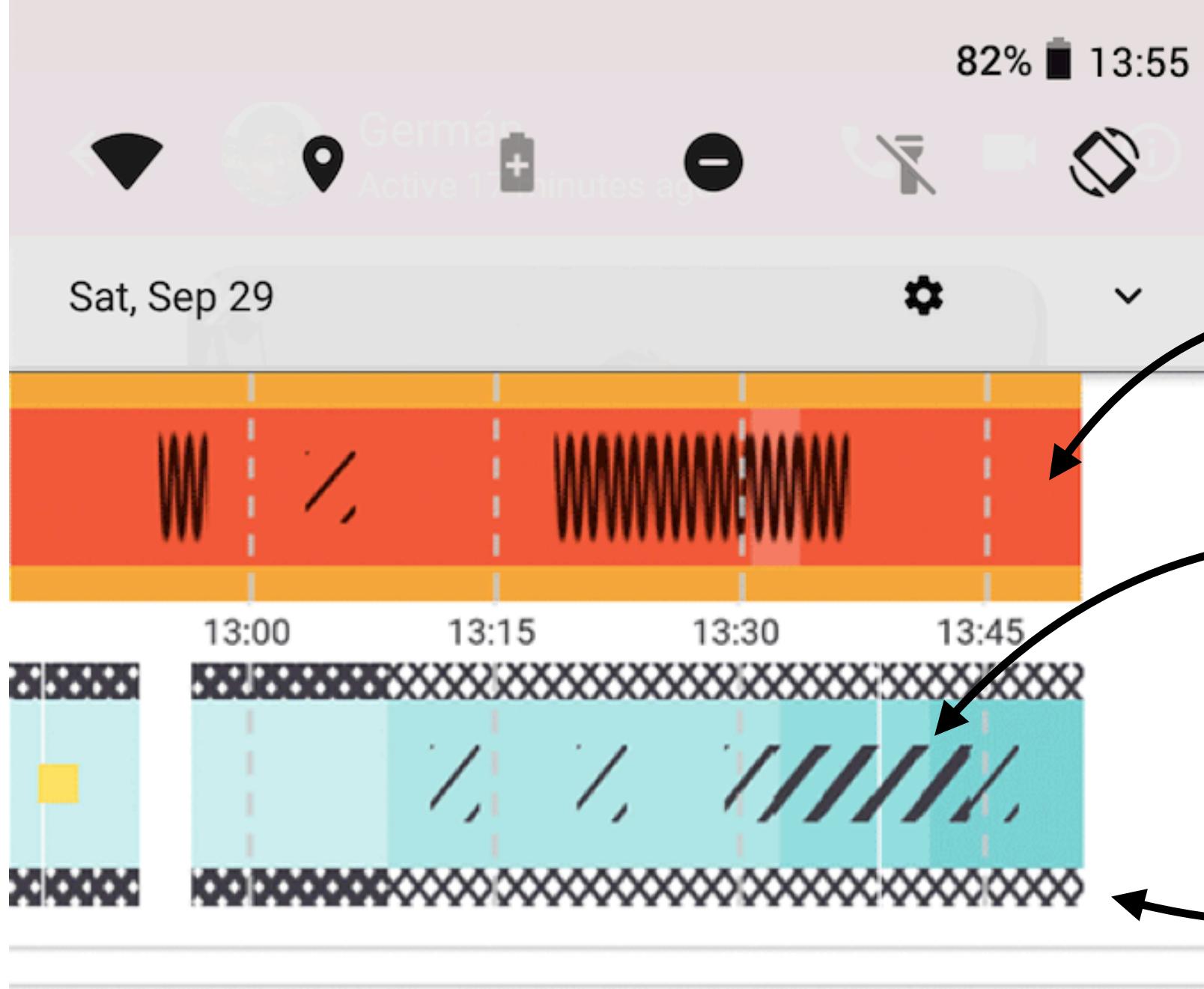
Las tareas se realizan en un tiempo determinado, o en ciertos momentos del día

Los participantes usan sus propios dispositivos

Los participantes usan el sistema como quieren

Los participantes usan el sistema cuando quieren / necesitan

Más diferencias?



Distancia a casa

Pasos

Nivel de batería

Objetivo del estudio: entender qué utilidad le dan las parejas a Lifelines (cómo lo usan en su vida real)

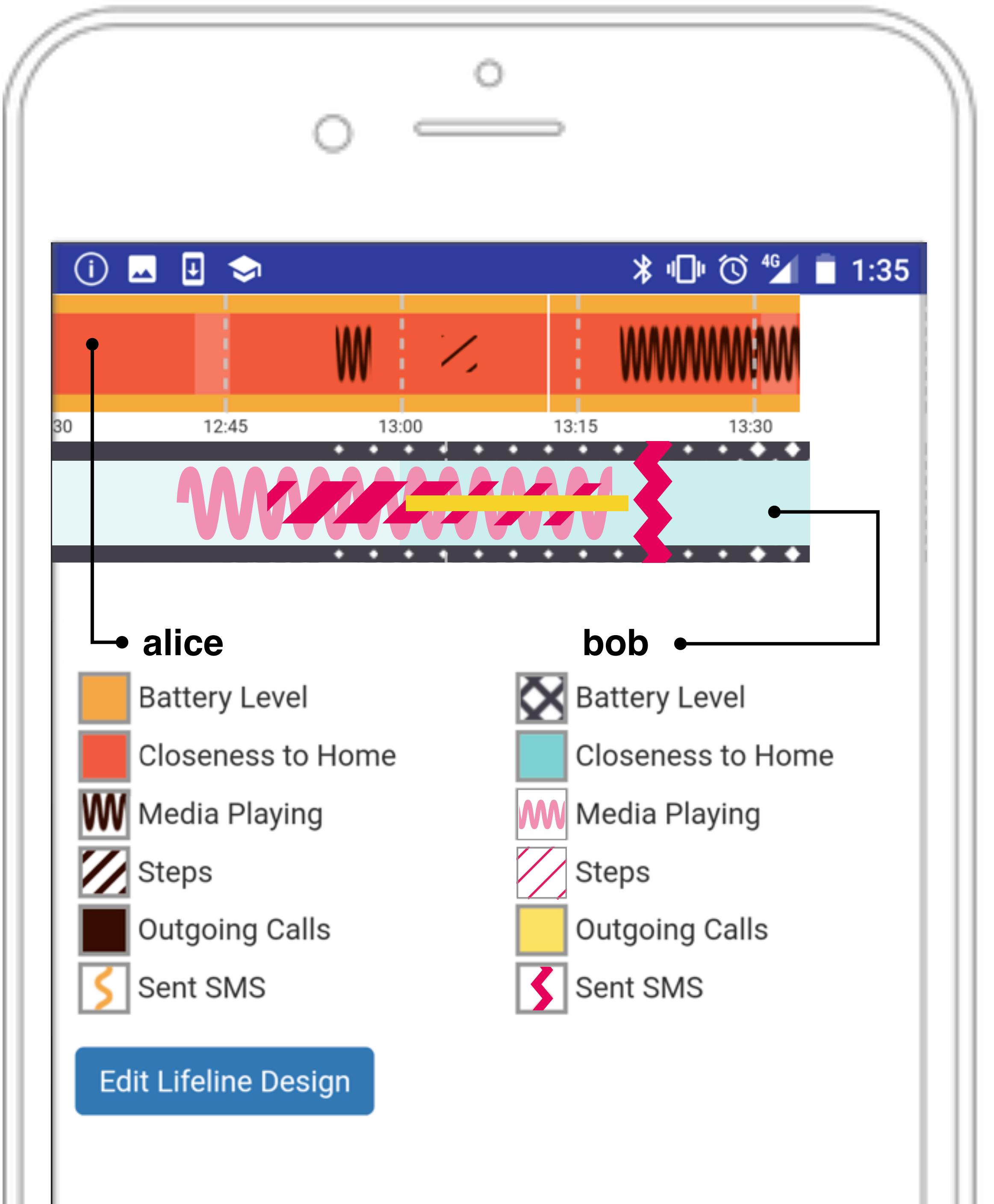
Campo de estudio con 9 parejas usando Lifelines en sus propios celulares

Lifelines

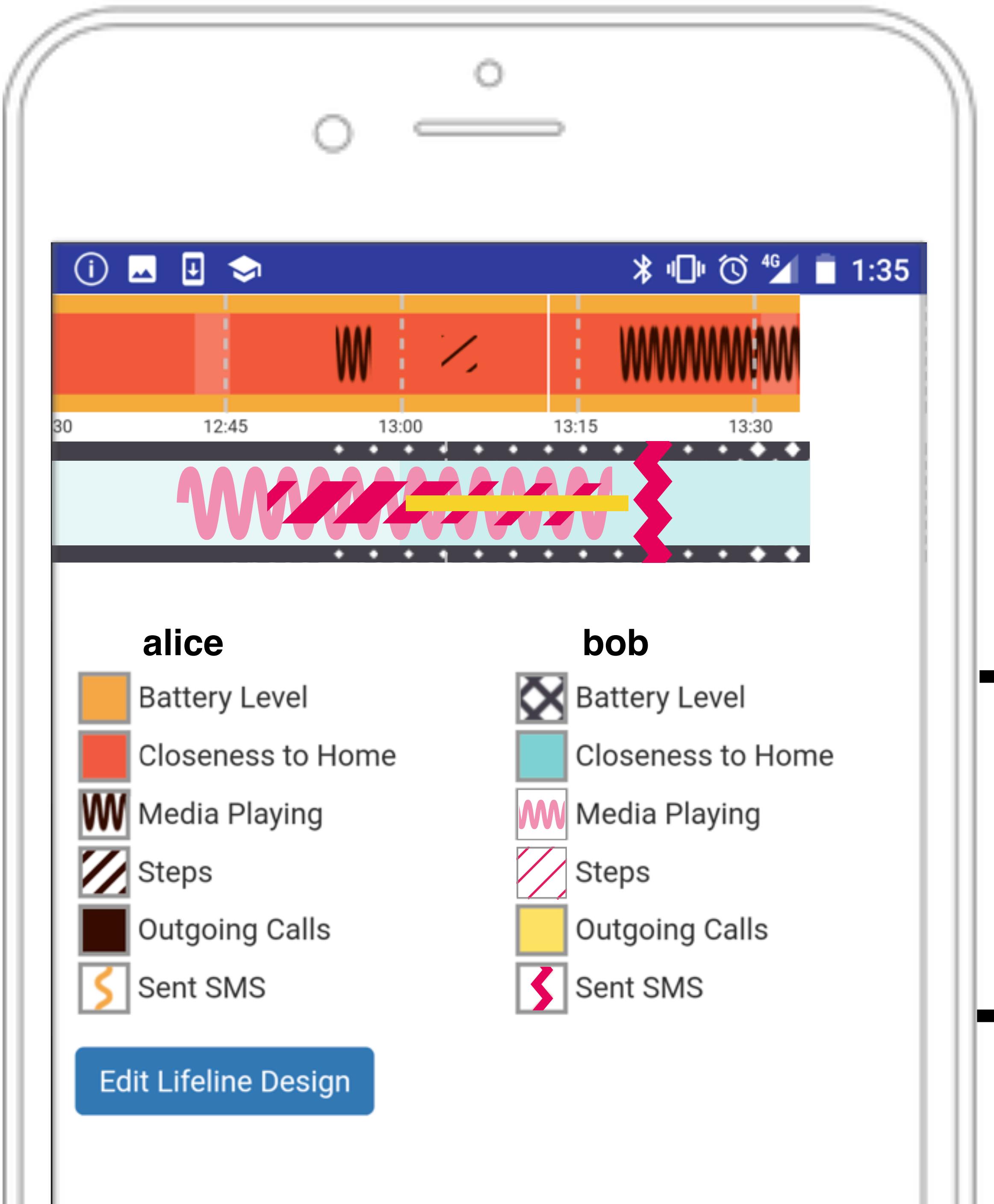
Griggio, Nouwens, McGrenere and Mackay. CHI 2019.
<https://doi.org/10.1145/3290605.3300853>

The Lifelines technology probe

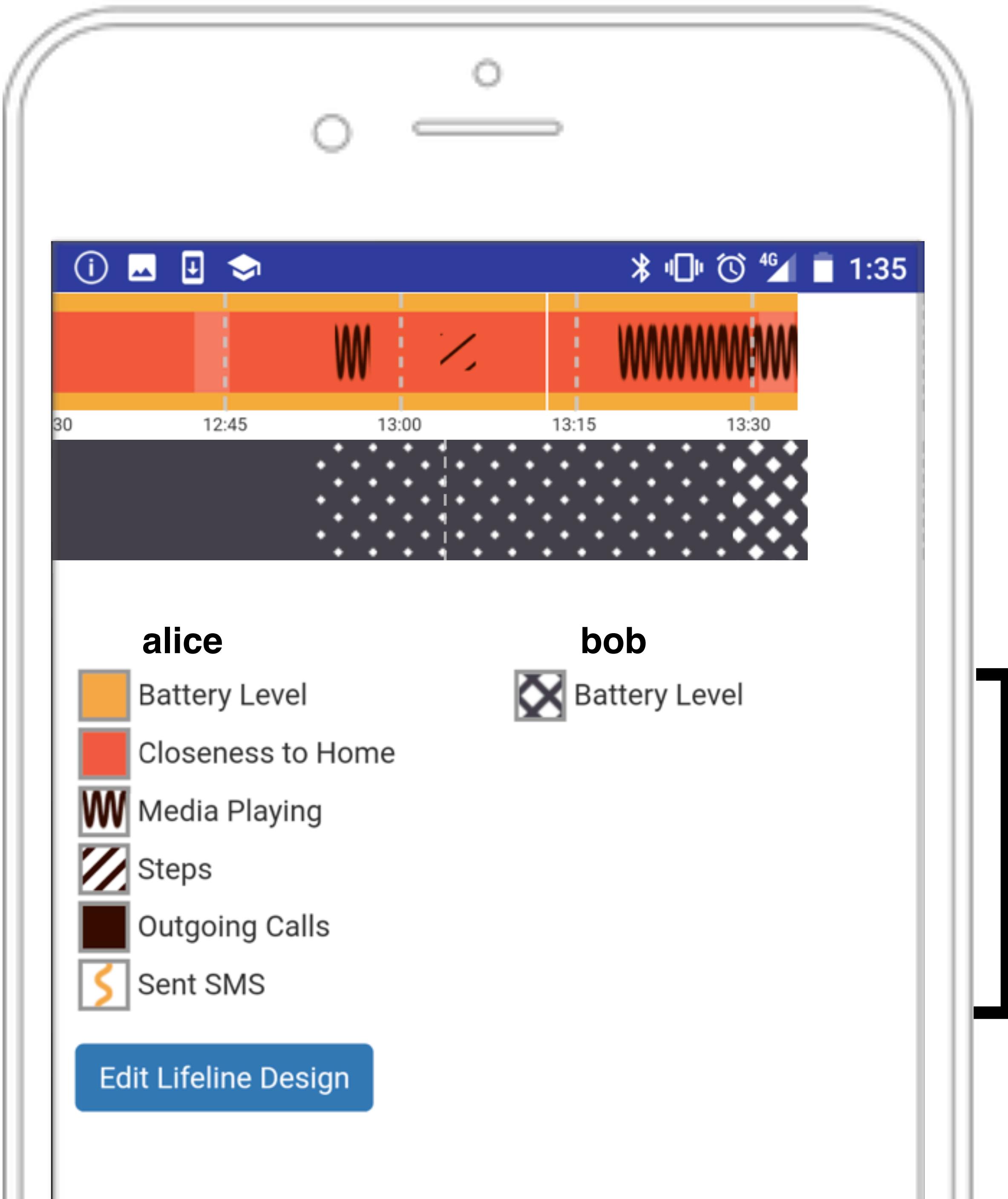




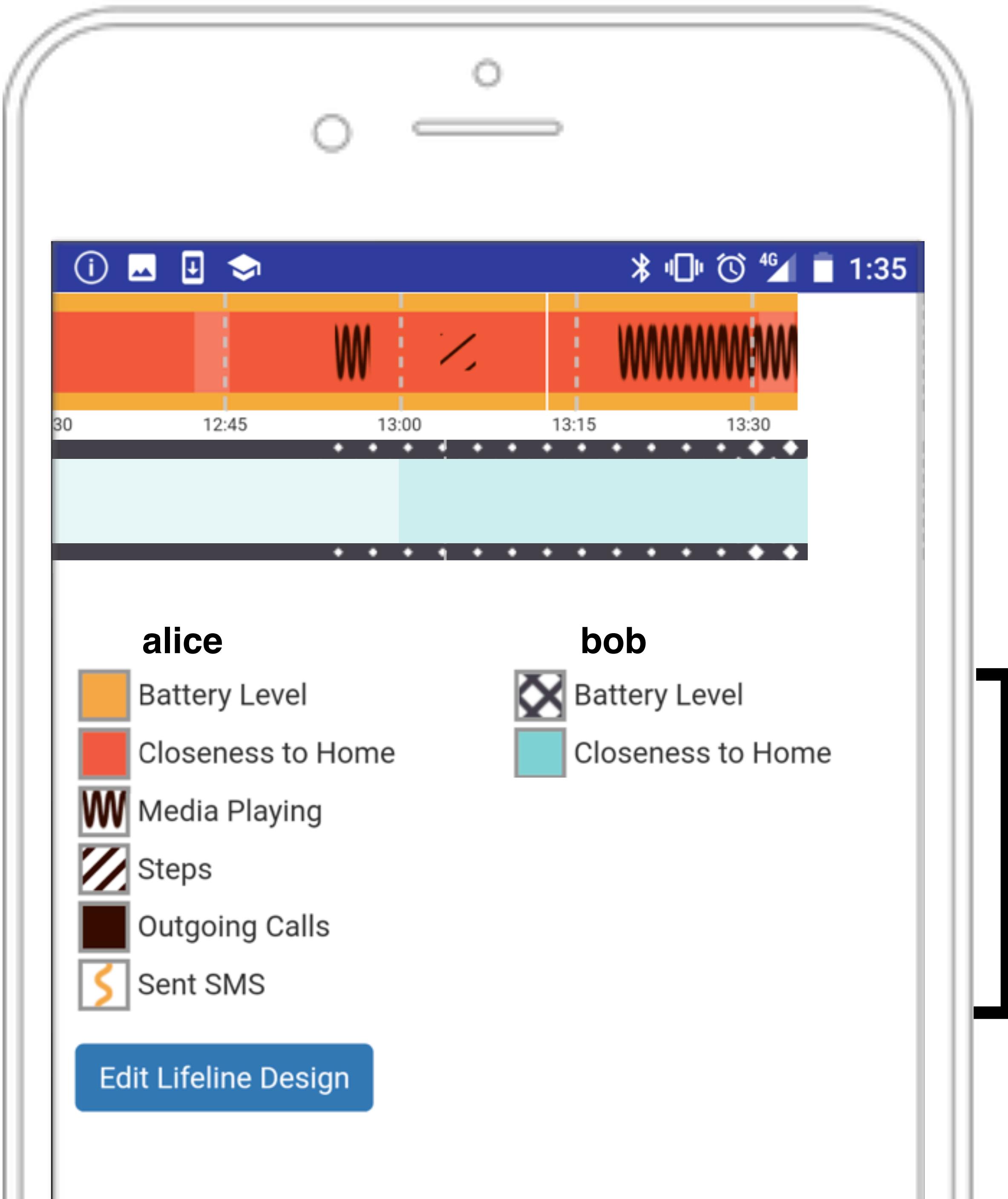
The Linebuilder



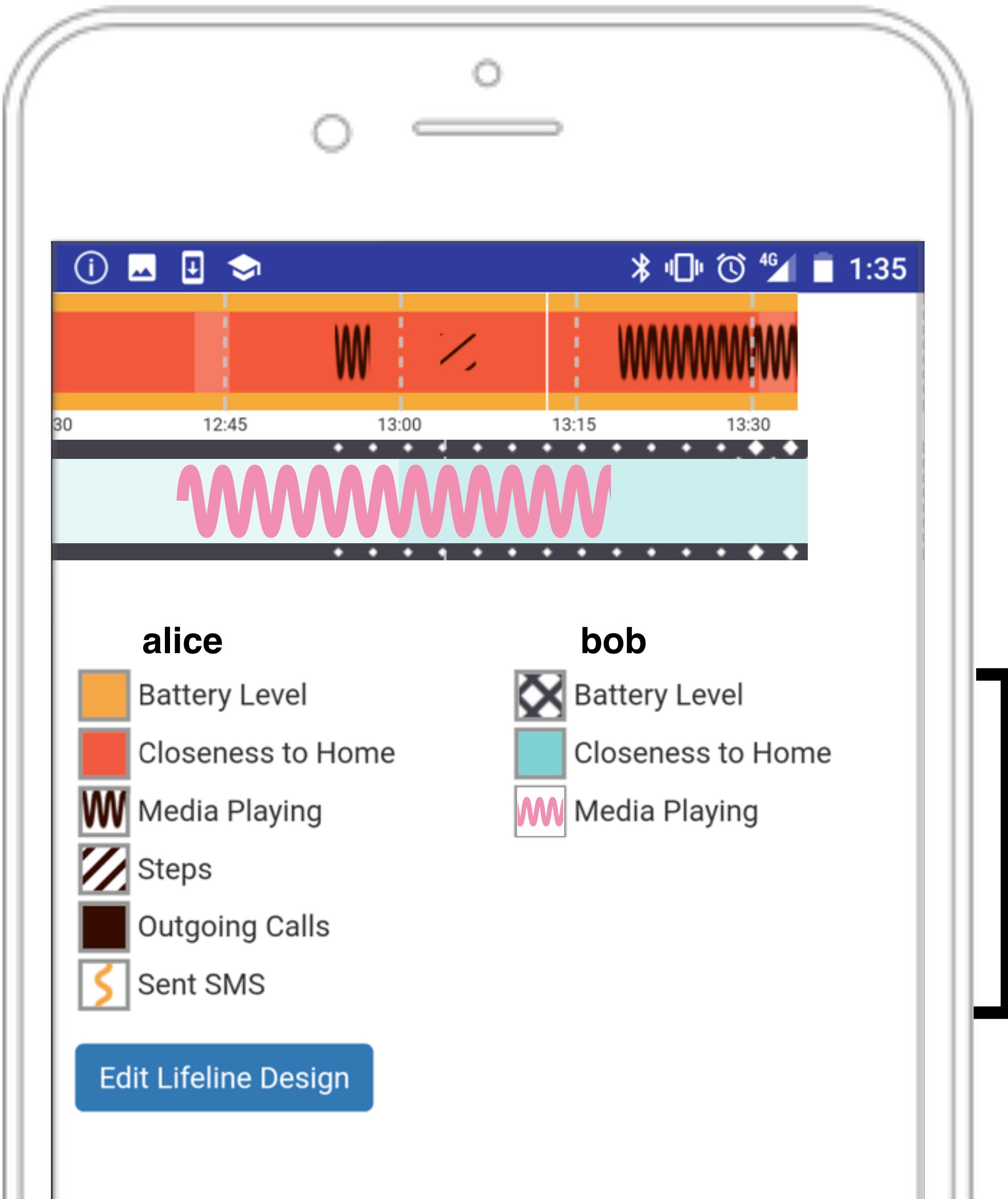
6 streams of
contextual
information



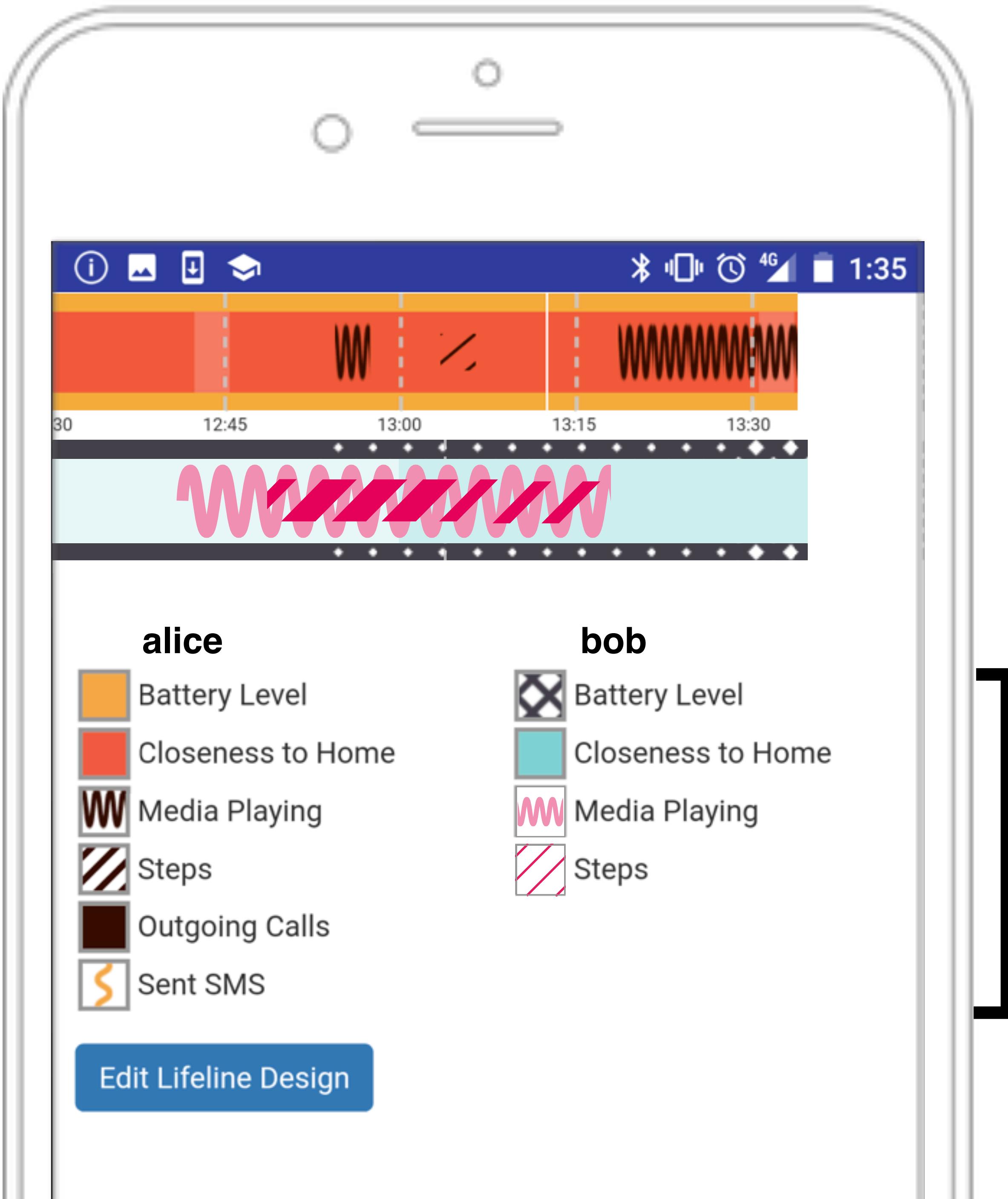
6 streams of
contextual
information



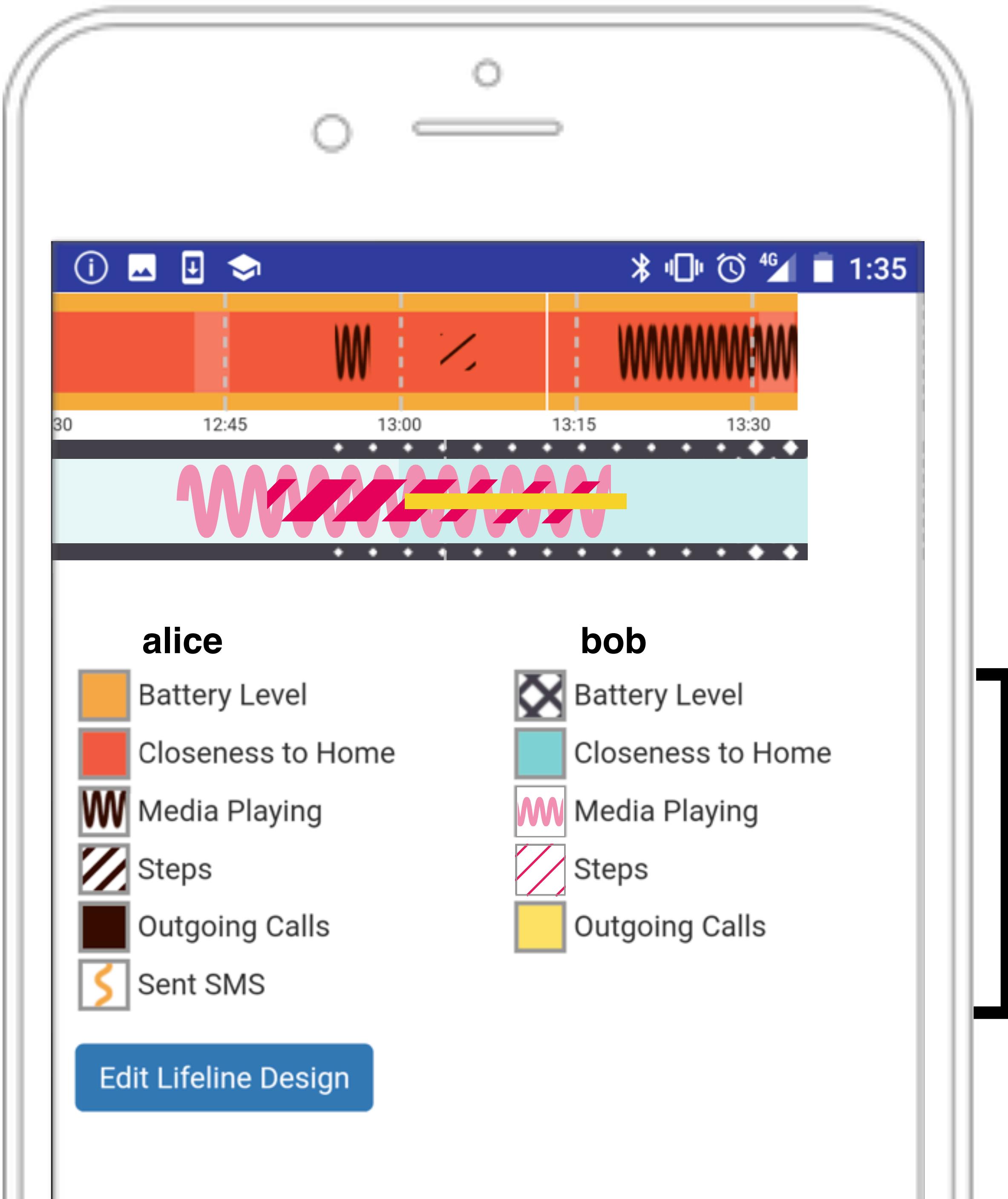
6 streams of
contextual
information



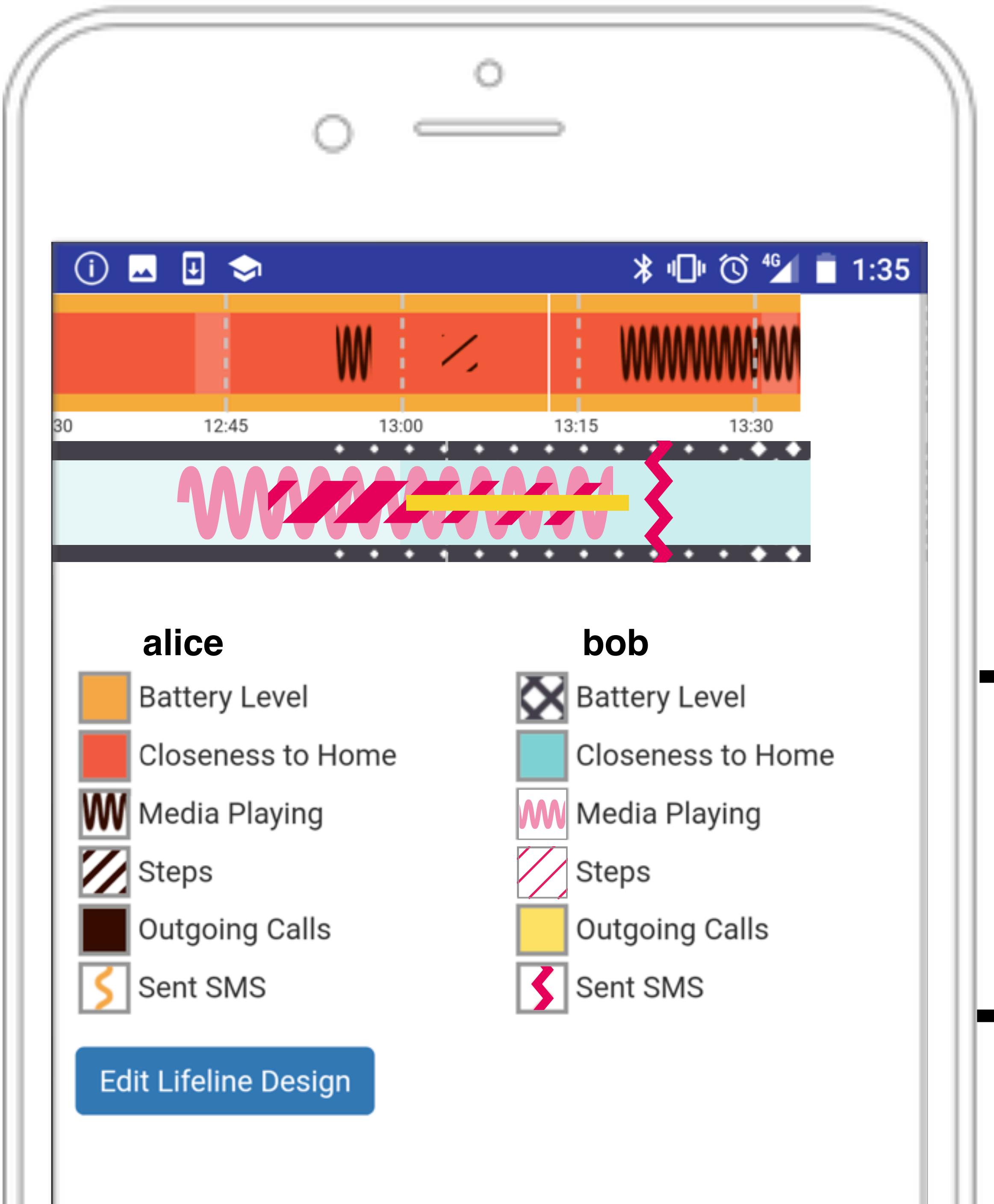
6 streams of
contextual
information



6 streams of
contextual
information



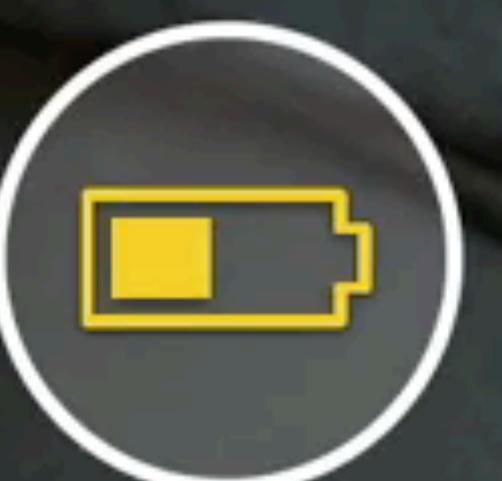
6 streams of
contextual
information



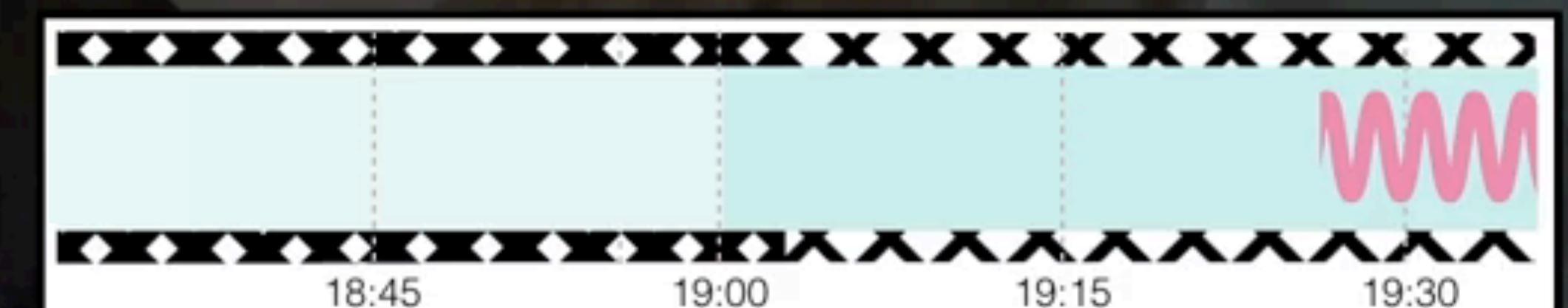
6 streams of
contextual
information



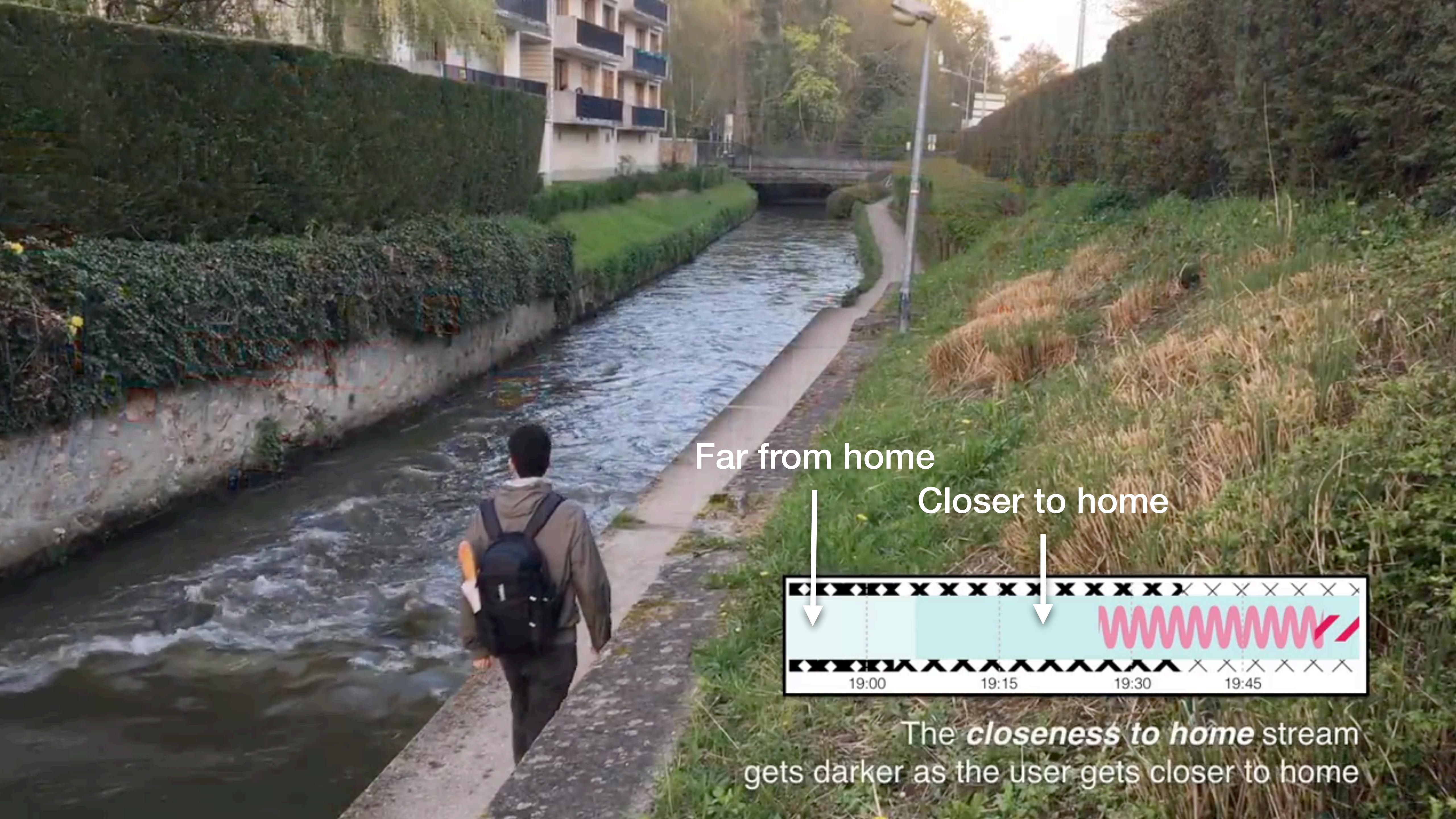
1 hour ago



now

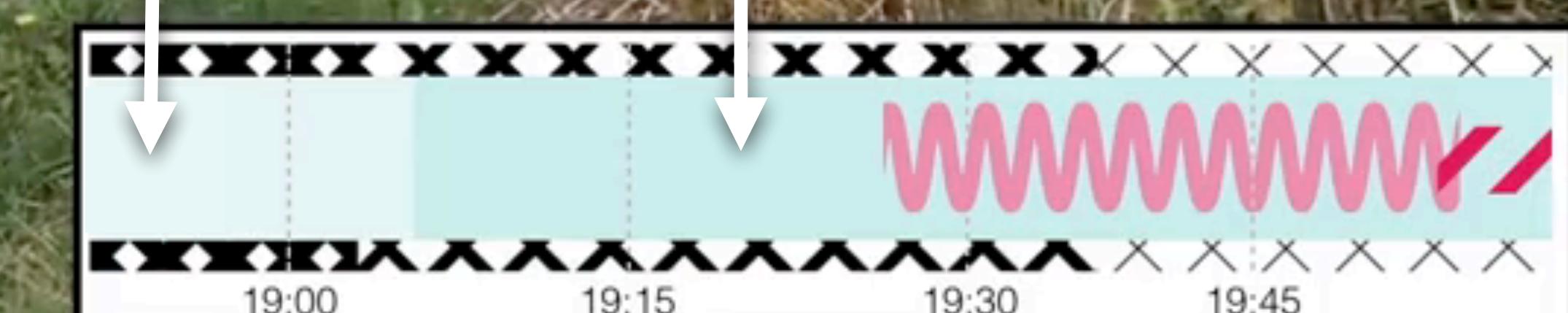


His **battery level** stream
gets thinner as the phone's battery drops



Far from home

Closer to home



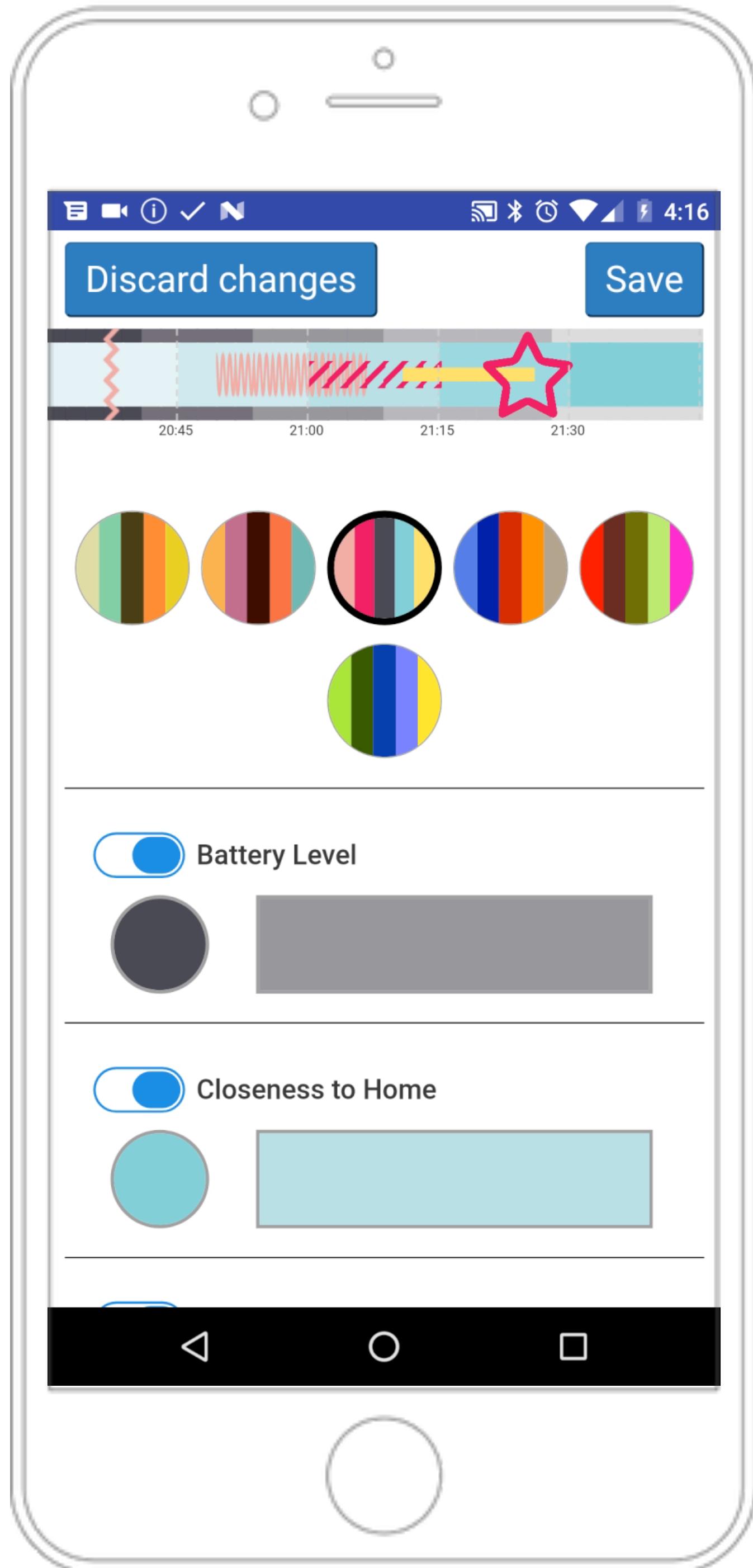
The **closeness to home** stream gets darker as the user gets closer to home



Very close
to home



The **closeness to home** stream
gets darker as the user gets closer to home



**Partners individually choose
the streams they want to share with the other**

A considerar en el diseño de estudios de campo

- 1) Cómo hacemos para obtener datos sobre qué está pasando en el día a día, si no estamos ahí para observar?
- 2) Cómo compensamos el “novelty effect”? Por cuánto tiempo tenemos que seguir recolectando datos?
- 3) Cómo recolectar datos sobre el uso del sistema en la vida real sin invadir la privacidad de los participantes?
- 4) Qué queremos controlar, qué NO queremos controlar, y qué NO PODEMOS controlar?

Cómo diseñarían este estudio?

Objetivo del estudio: entender qué utilidad le dan las parejas a Lifelines (cómo lo usan en su vida real)

METHOD

One-month technology-probe study

9 couples (8 living together)

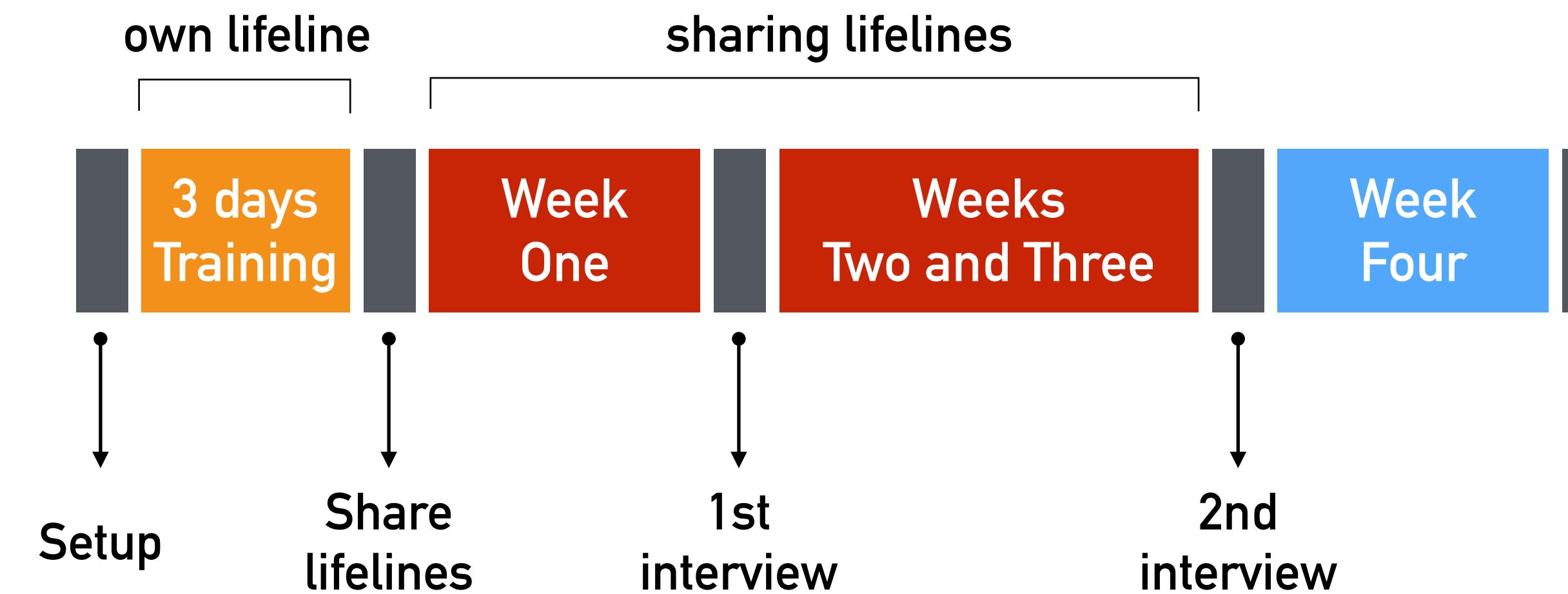
Argentina, Brazil, UK, Switzerland, New Zealand & France

Setup and Interviews via Skype

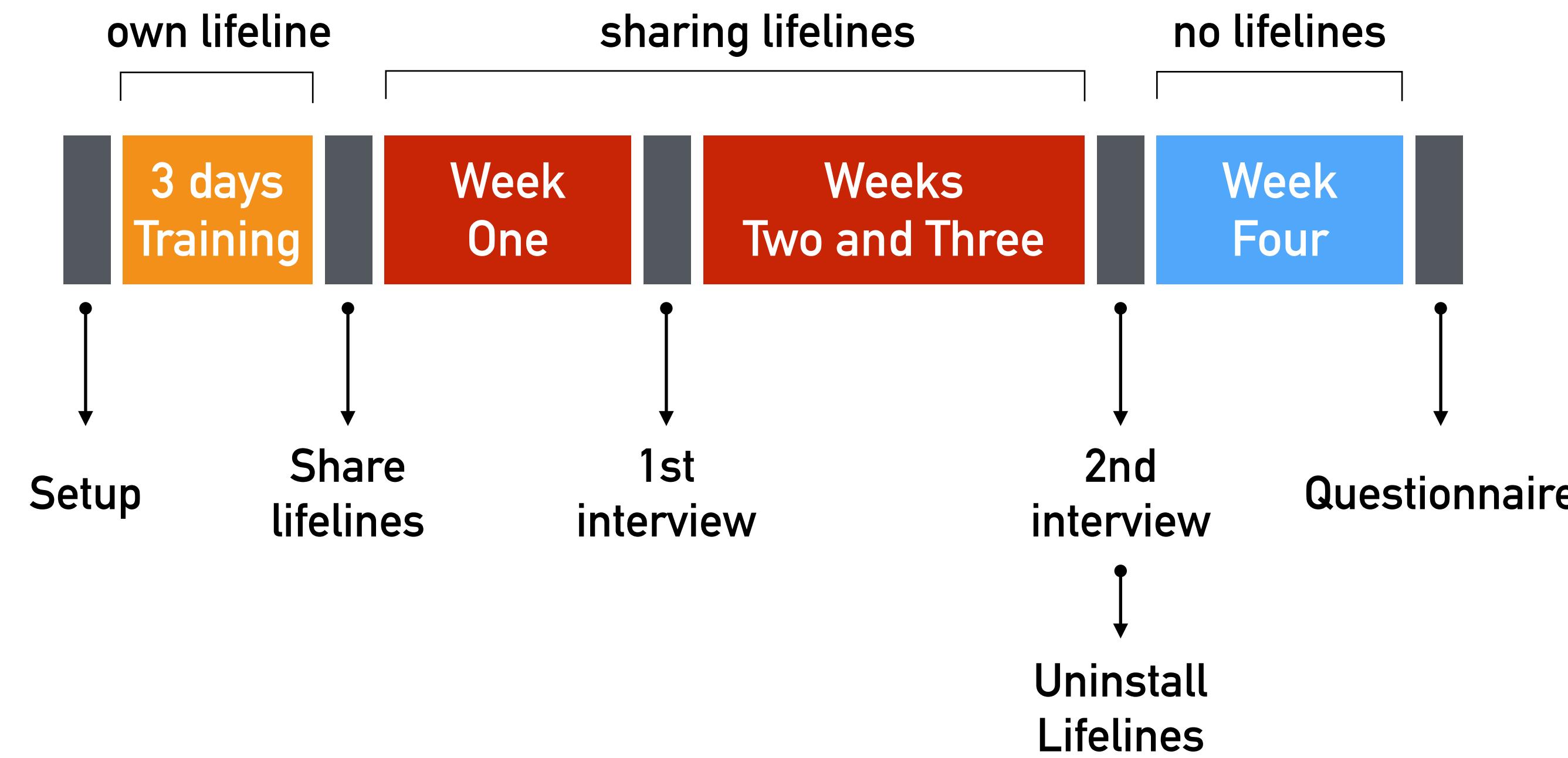
During this “practice” period, participants can get familiar with how their context is visualised, and make informed decisions about what streams they want to share with their partners



We have a 1st interview after 1 week of sharing the contextual data so we can capture participant's first impressions, taking advantage of novelty effects



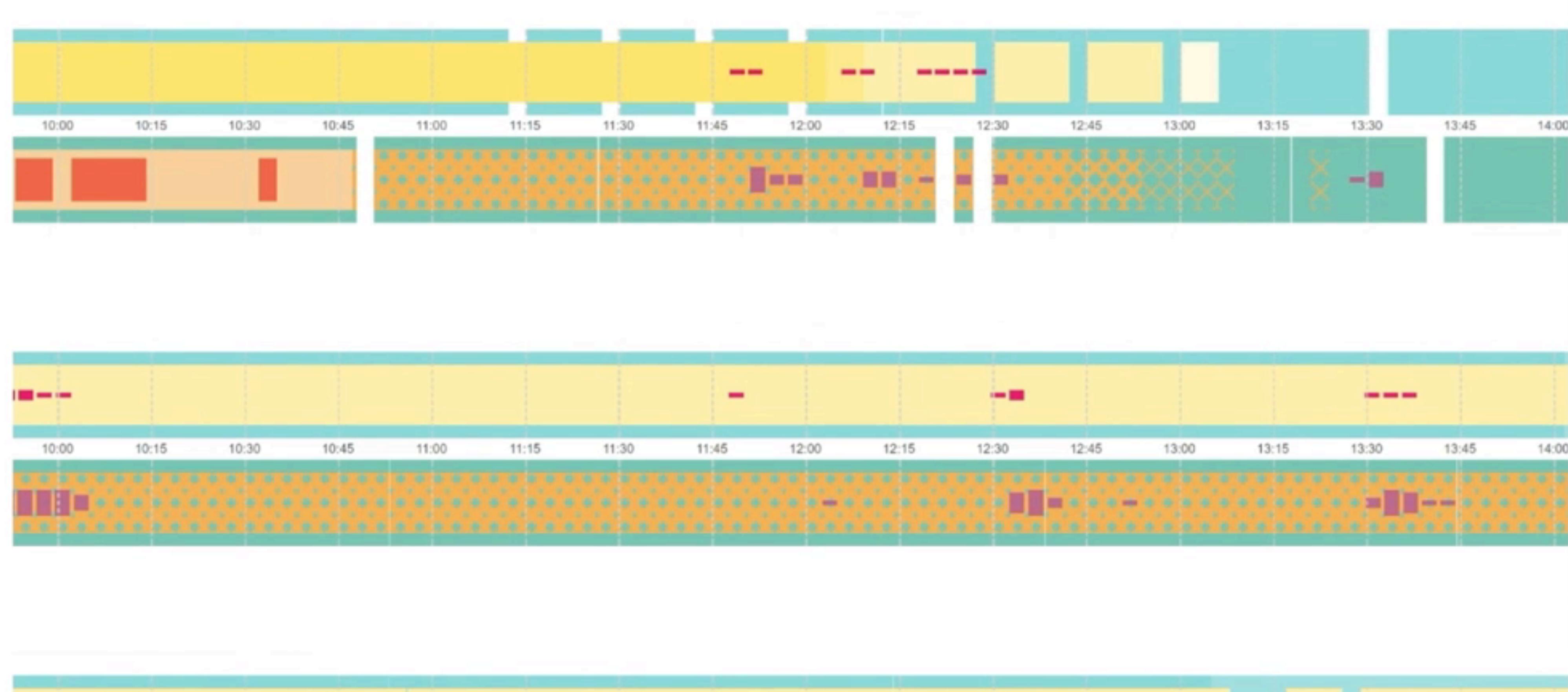
Then we wait 2 more weeks for the 2nd interview, to learn about their experiences past the novelty effects



Last, we ask them to STOP sharing their lifelines and come back to them after a week, so we learn about the contrast of sharing vs. not sharing contextual information

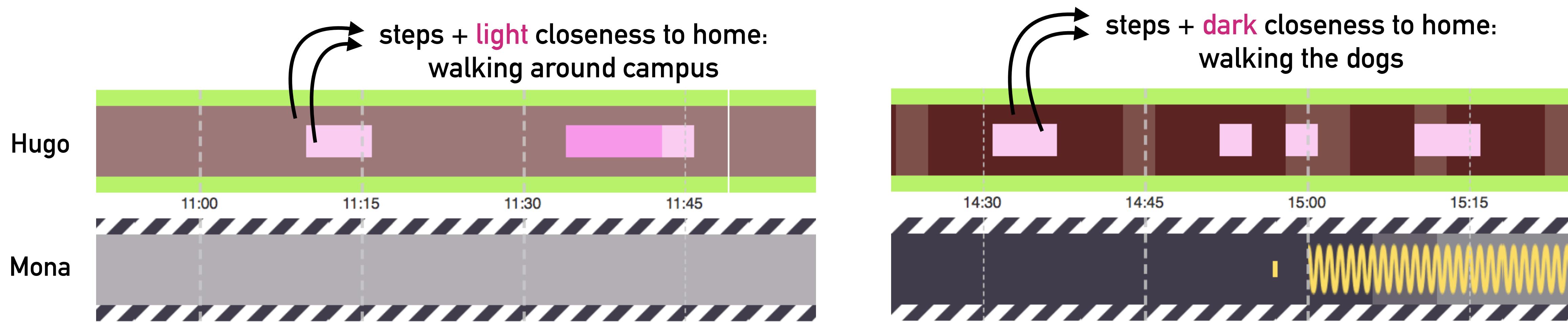
We collected 22 hours of video from 34 Skype interviews

Open coding on videos (Chronoviz) + Thematic Analysis



During interviews, we show them their “old” data to help them remember interesting events and let them explain to us what happened in their own words

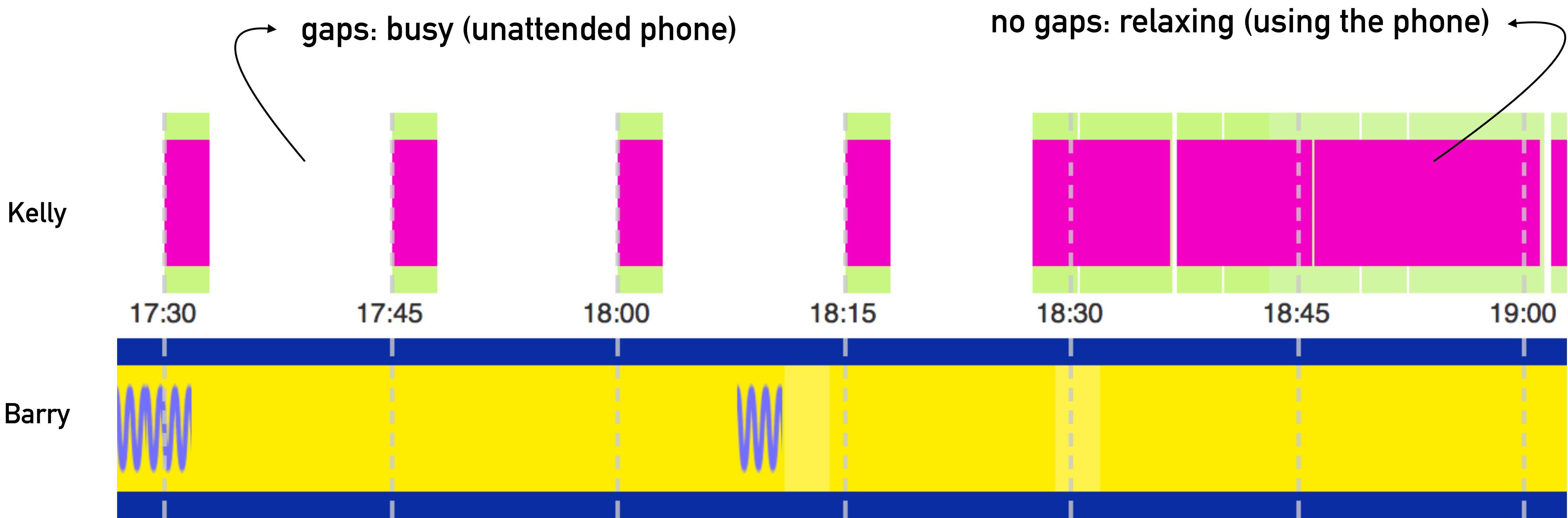
How did multiple streams inform context?



Interpreting **multiple streams together** helps disambiguate the meaning of single streams

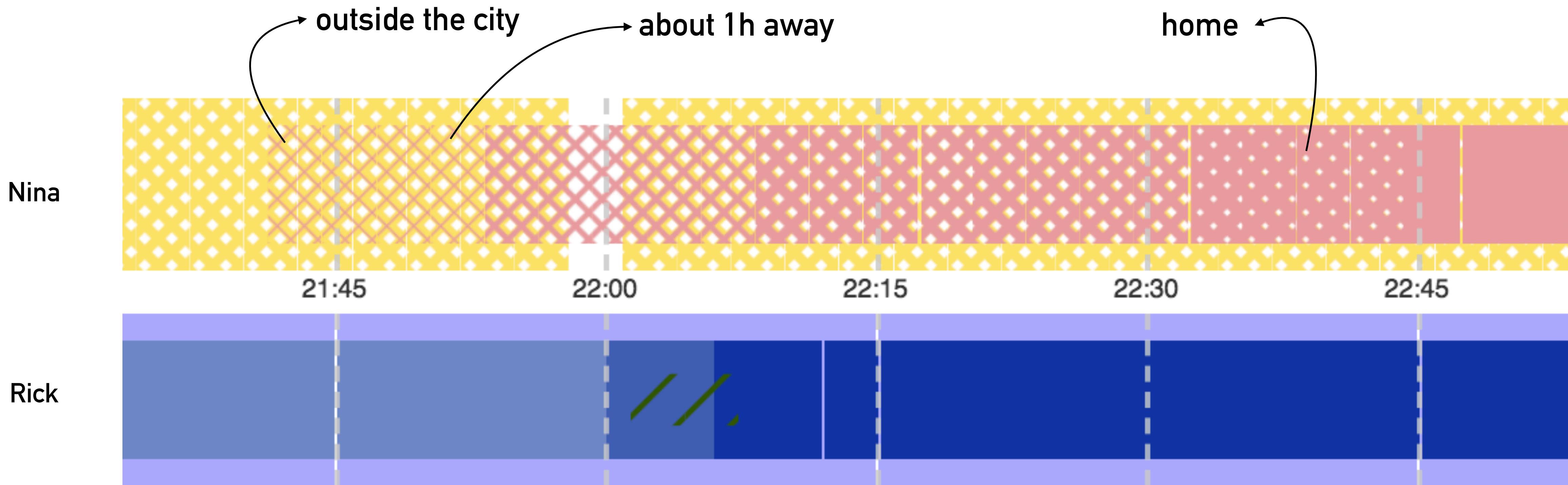
Participants most often looked at **one stream at a time**
according to the data that served their current needs

How did persistent streams inform context?



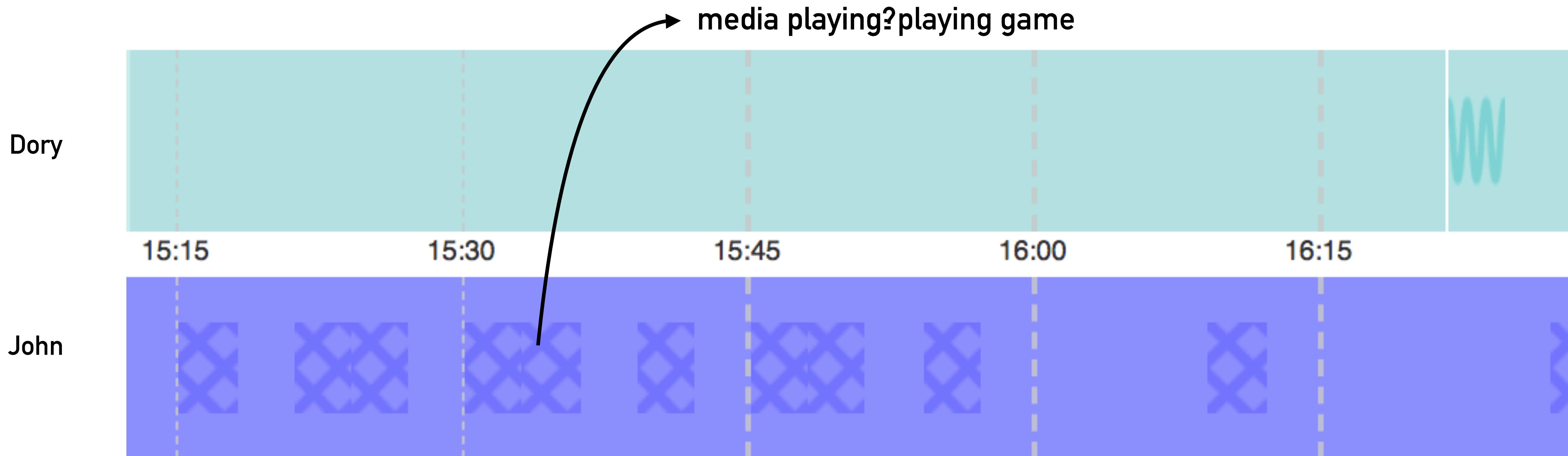
Patterns of missing data provide additional contextual information

How did couples' communication dynamics change?



Contextual information **replaced** direct communication

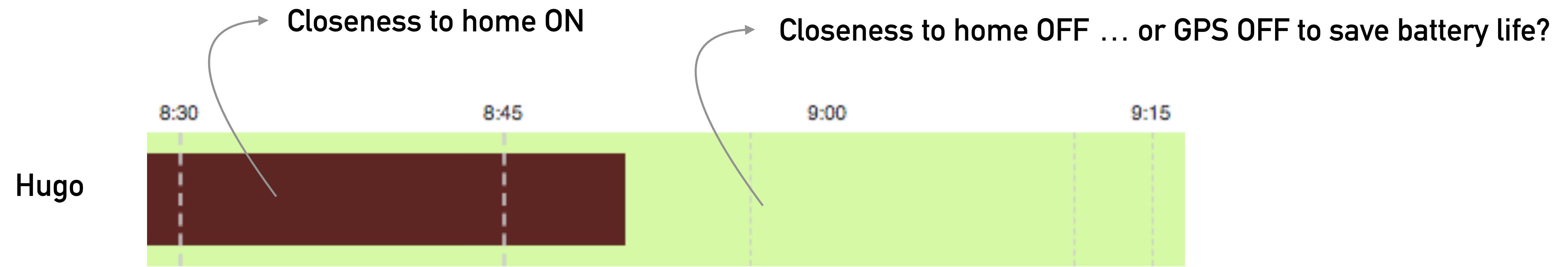
How did couples' communication dynamics change?



Contextual information triggered direct communication

RESULTS

Replacing direct communication
may hurt established communication patterns



Ambiguous visualizations enable plausible deniability

A poor balance between triggering and replacing
direct communication can lead to feeling more distant

Strong individual differences across and within couples

The streams that triggered or replaced communication depended on the routines, needs and intimate knowledge of each couple

Media Playing

For Dory: ask about surprising patterns (trigger)

For Kelly: coordinate pick-ups (replace)

Steps

check that her husband arrived to work safely (replace)

warn her husband about his knee health (trigger)

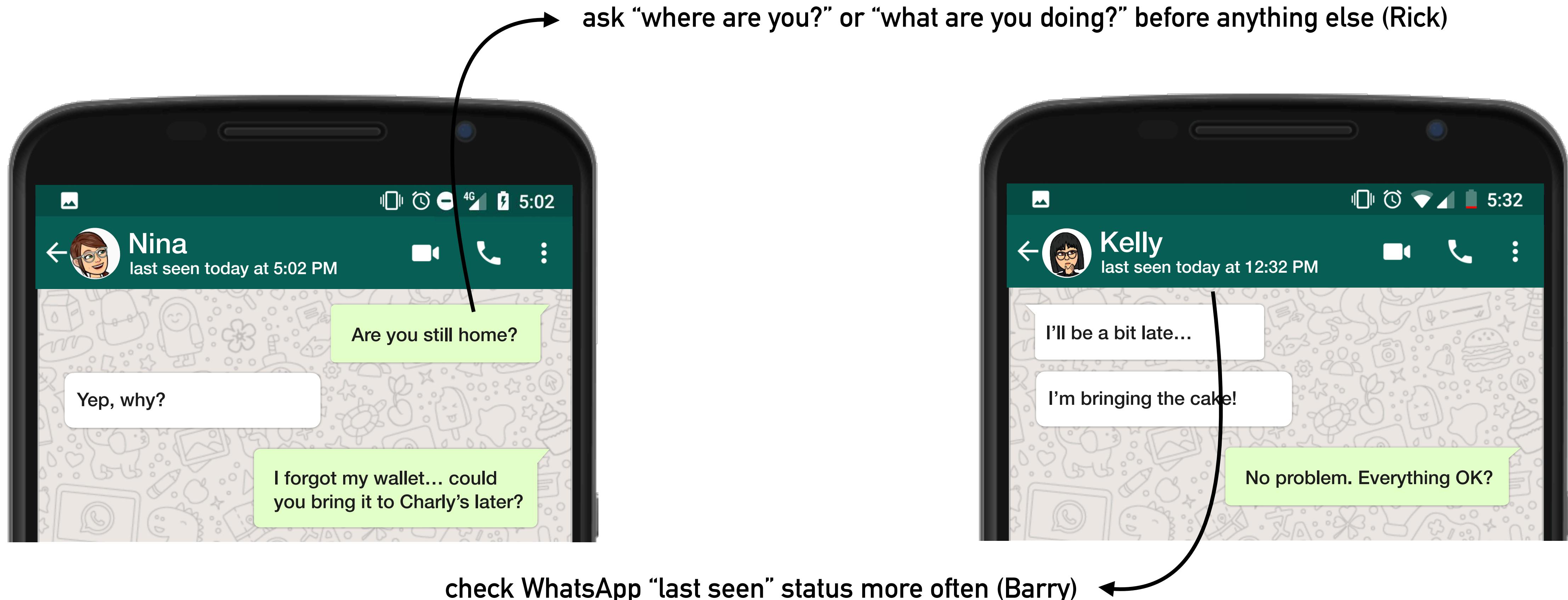
Data that confirms partners' expectations and knowledge of each other replaces direct communication

Data that challenges partners' expectations and knowledge of each other triggers direct communication

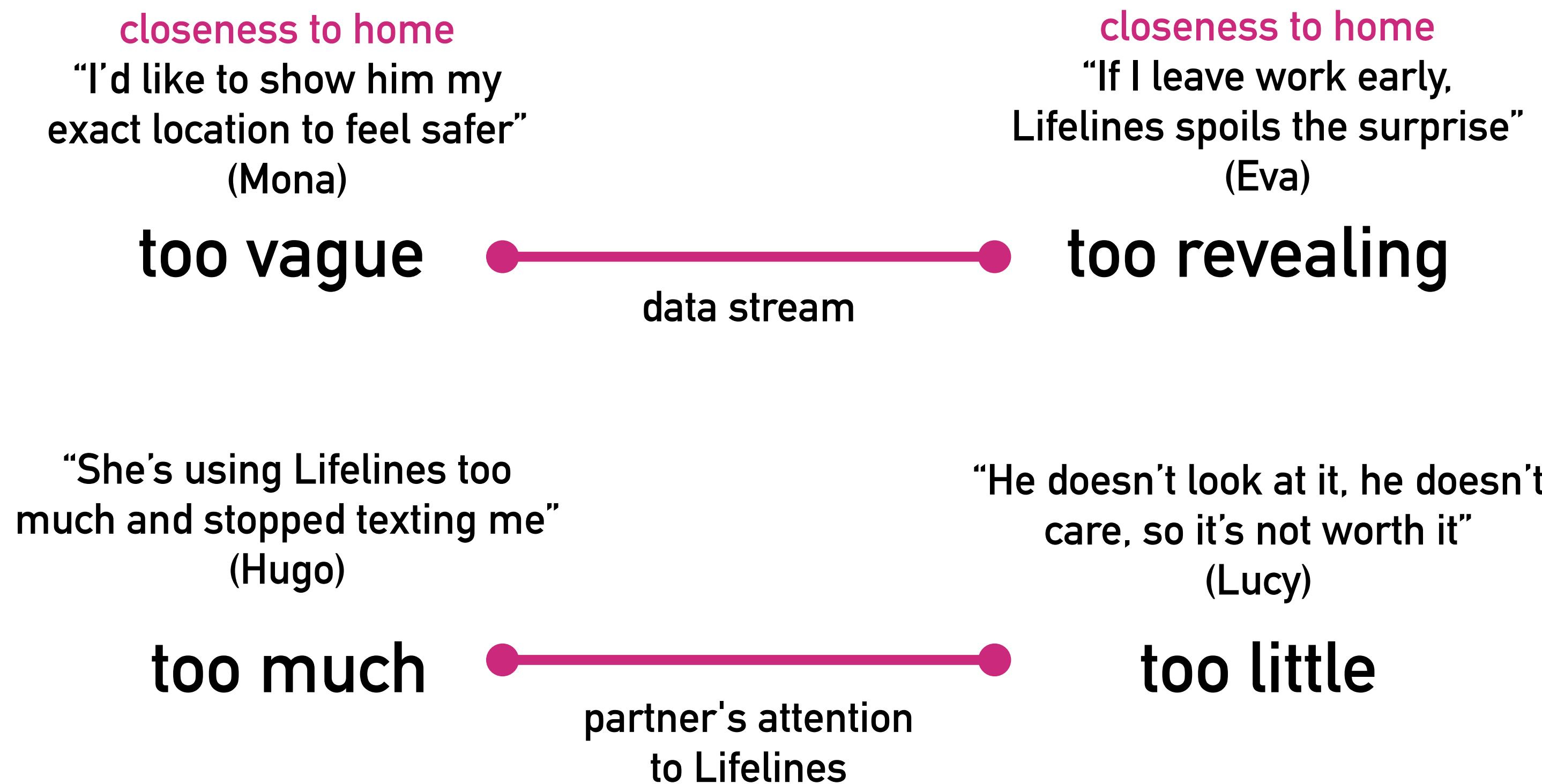
Post-Lifelines

15 out of 18 participants missed sharing at least 2 streams.

Some found ways of compensating the lack of Lifelines:



Individual differences around privacy concerns



Opportunities for design

Based on the strong individual differences in **privacy concerns** across couples: how about **customizable levels of persistence**? Some might prefer live, ephemeral visualisations rather than historical, persistent visualisations.

Based on the strong individual differences in **what data is meaningful** across couples: what if every app had a “context API” for enabling visualisations of more diverse types of contextual data?

Systems sharing contextual information should support **plausible deniability** to protect user’s privacy

Conclusions

Experiments are a great method to compare two or more designs, in a controlled setting, to evaluate how they affect the user experience / performance / other measures

The more we control a user study, the more reliable, but the less “ecologically valid”

Field studies can generate insights with high ecological validity, about the system’s usability but also about why and in what contexts participants find our system the most valuable



LipLearner

Customizable Silent Speech Interactions on Mobile Devices

LipLearner: Customizable Silent Speech Interactions on Mobile Devices.

Zixiong Su, Shitao Fang, and Jun Rekimoto.

In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). <https://doi.org/10.1145/3544548.3581465>