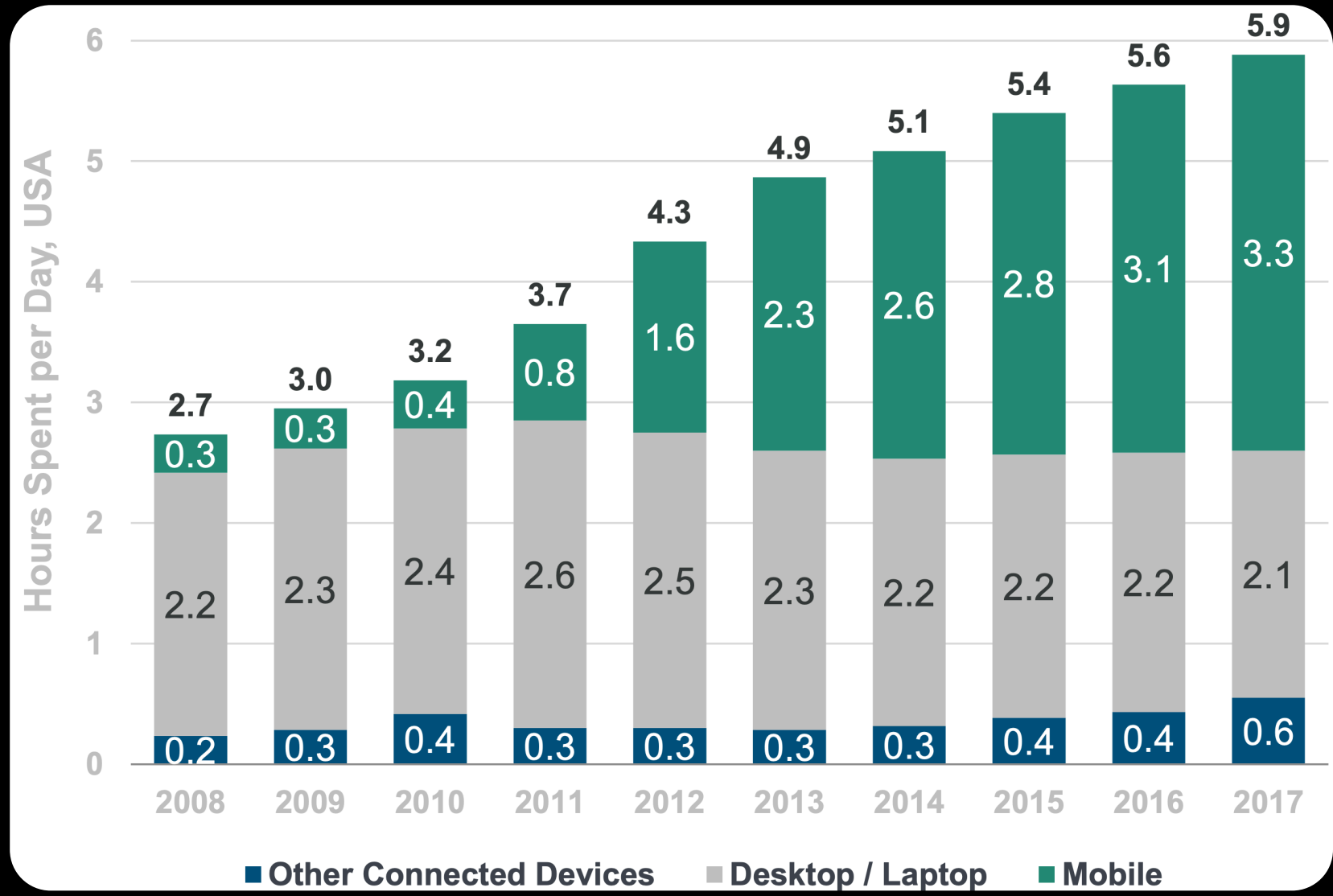
HabitLab: In-the-wild Behavior Change Experiments at Scale

Geza Kovacs

Stanford HCI Group

People spend increasing amounts of time online

Average US adult spends 5.9 hours per day with digital media



Source: Kleiner Perkins 2018 Internet Trends

People struggle to reduce their time online

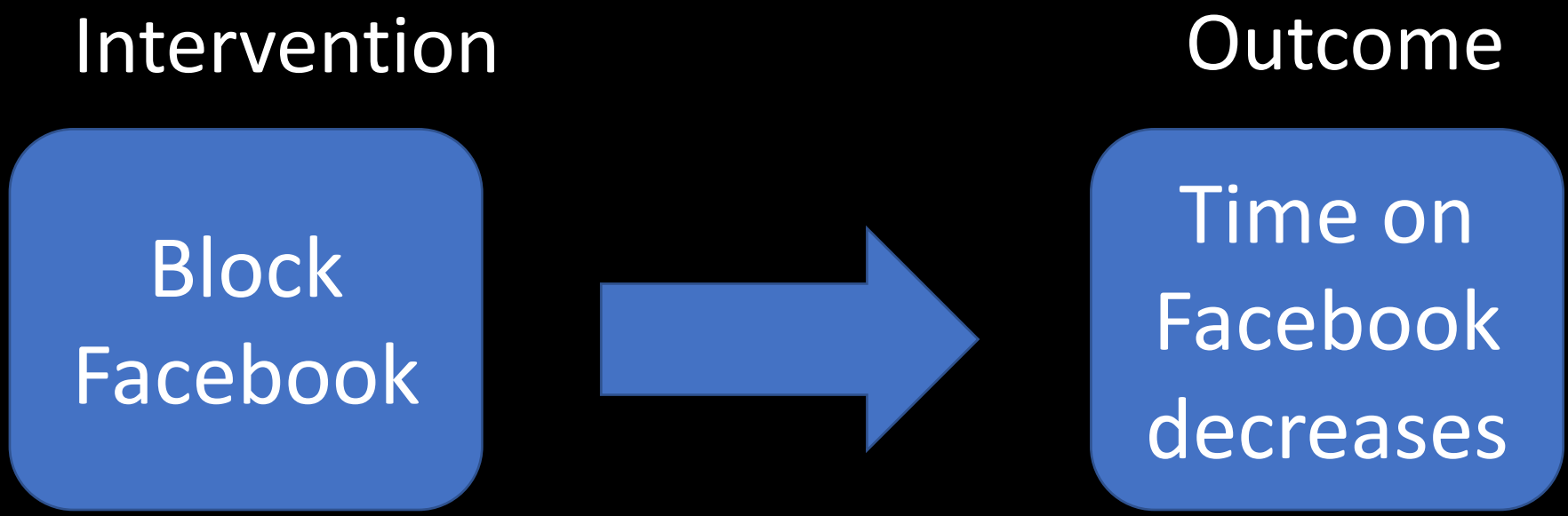
Kim, Young-Ho, et al. "TimeAware: Leveraging framing effects to enhance personal productivity." Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. ACM, 2016.

Andreassen, Cecilie Schou, et al. "Development of a Facebook addiction scale." Psychological reports 110.2 (2012): 501-517.

Users use productivity tools to help them reduce time online

Interventions and resulting outcomes

If your goal is to spend less time on Facebook



Online behavior change is a domain well-suited for studying interventions and outcomes

Outcomes

- Can measure outcomes precisely (time spent on each site each visit)

Interventions

- Many interventions possible
- Can adapt interventions quickly (new one each visit)

Current productivity tools make assumptions about interventions and outcomes

Assumptions about **outcomes**

- Effectiveness persists over time
- There are no negative externalities

Assumptions about **interventions**

- A single intervention can meet most needs
- Users are good at predicting what interventions will work for them

We developed HabitLab to study whether these assumptions are actually true

Questions about **outcomes**

- Does effectiveness remain constant over time?
- Do externalities exist?

Questions about **interventions**

- Does a single intervention meet most needs?
- Are users good at predicting what interventions will work well for them?

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

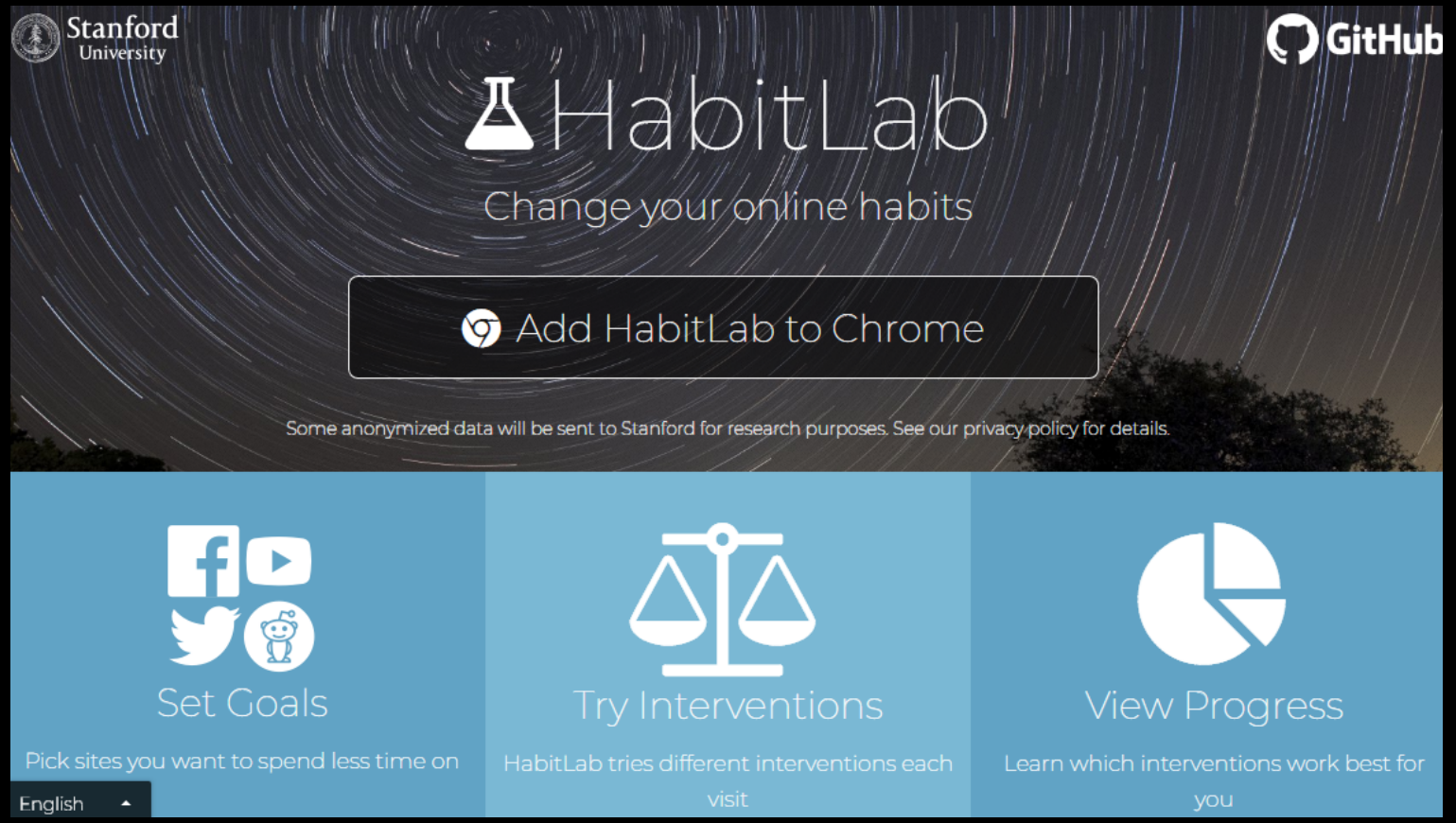
Interventions

How do users' preferences change over time?

HabitLab

Our behavior change platform

12,000+ in-the-wild active users (Browser+Android)

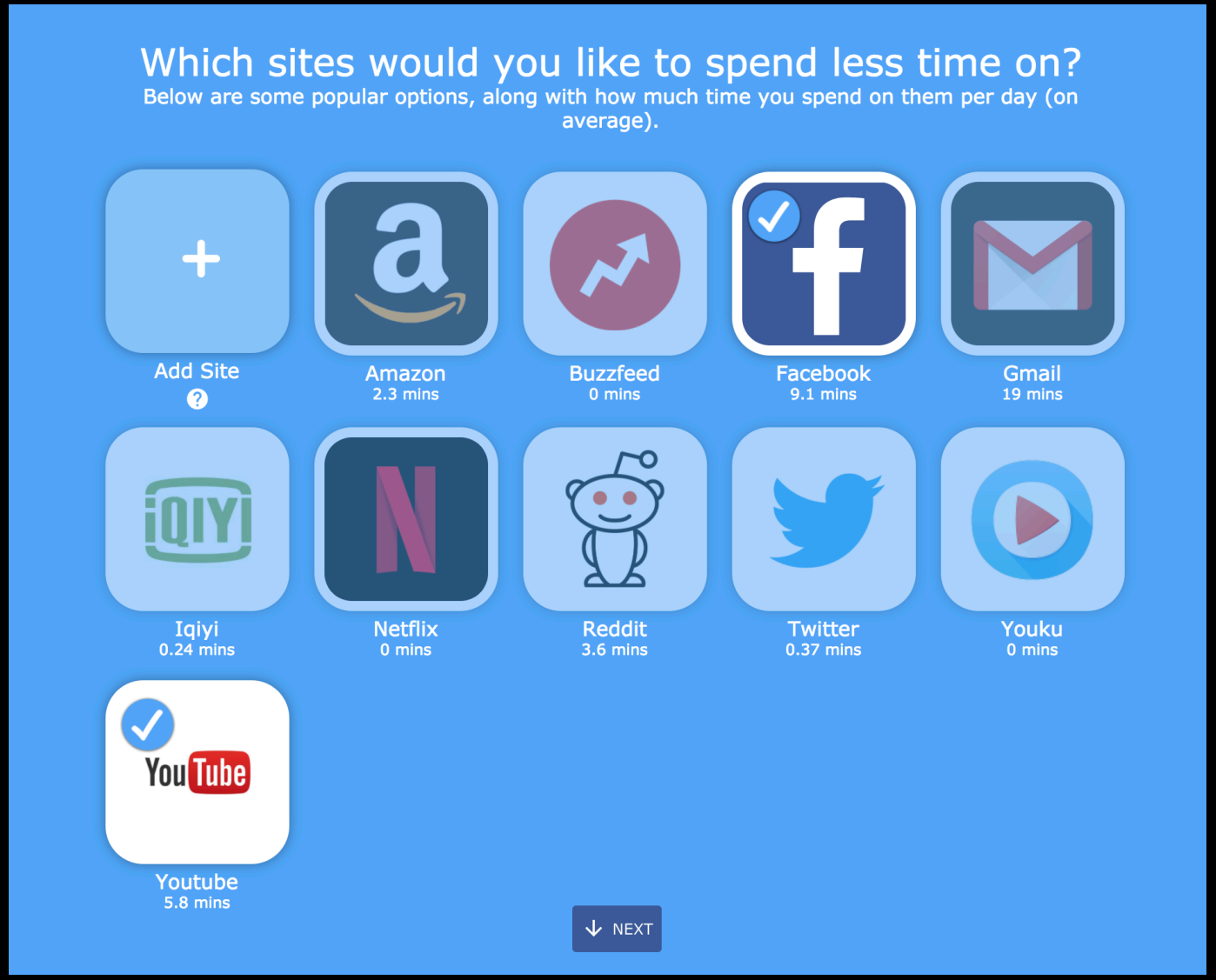


HabitLab

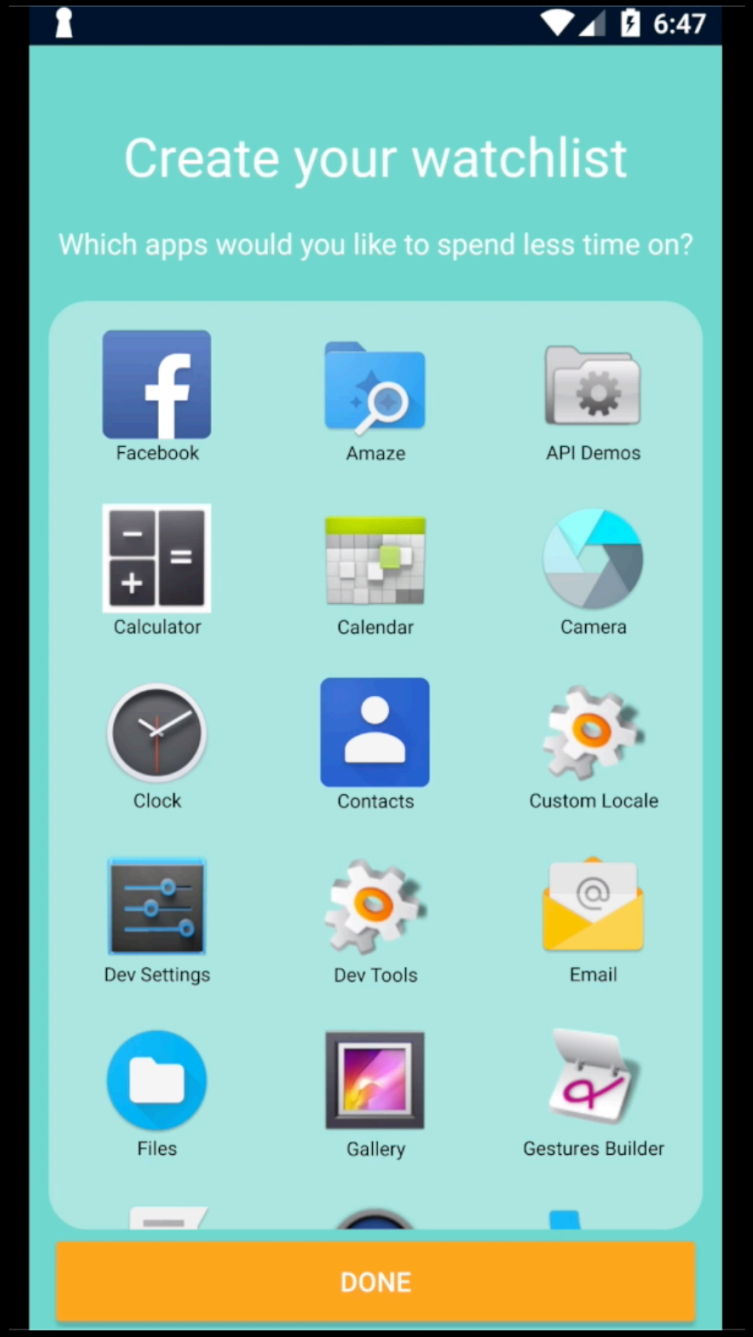
Our behavior change platform

12,000+ in-the-wild active users (Browser+Android)

Users select sites or apps to reduce time on (goals)

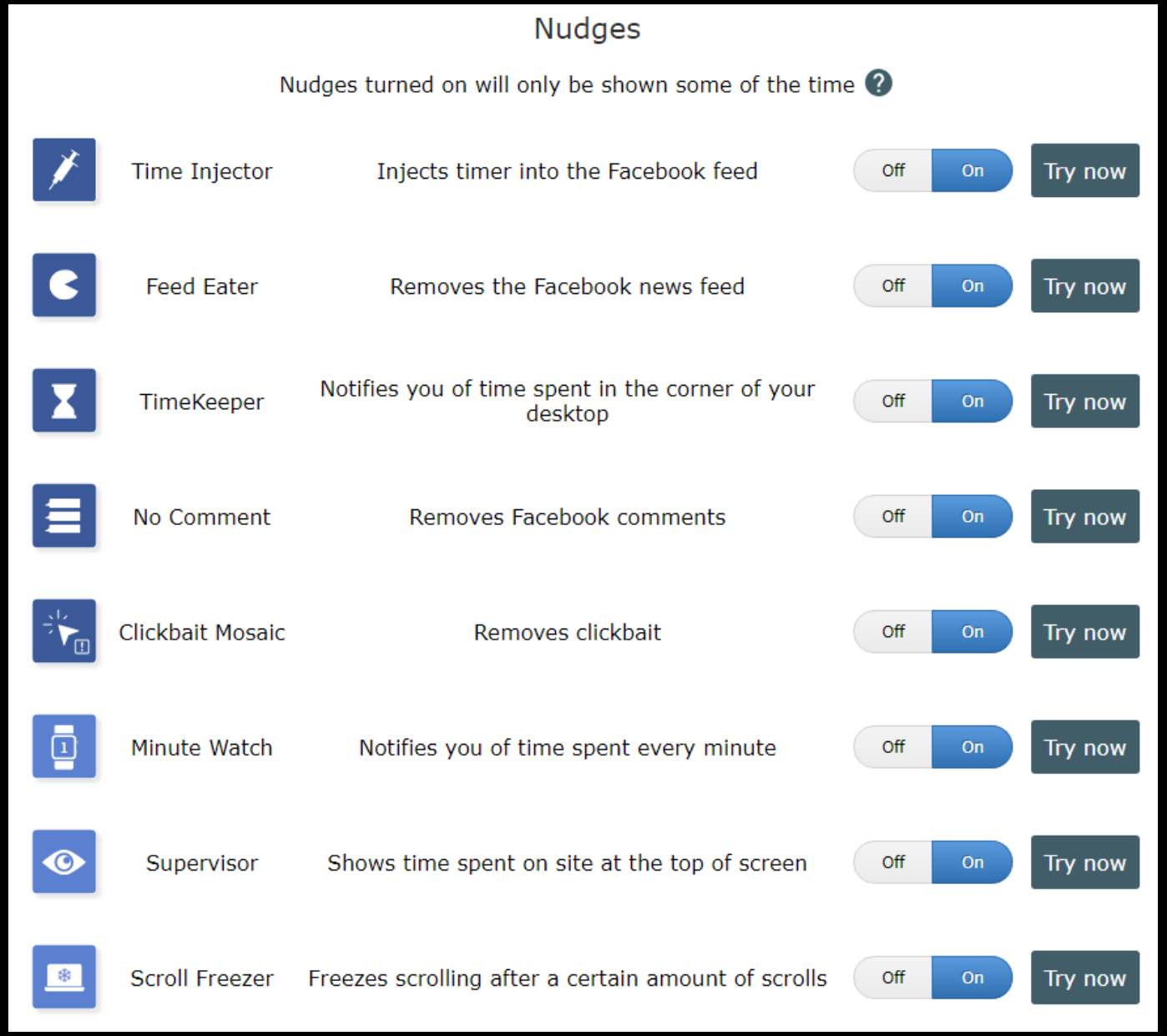


Users select sites or apps to reduce time on (goals)

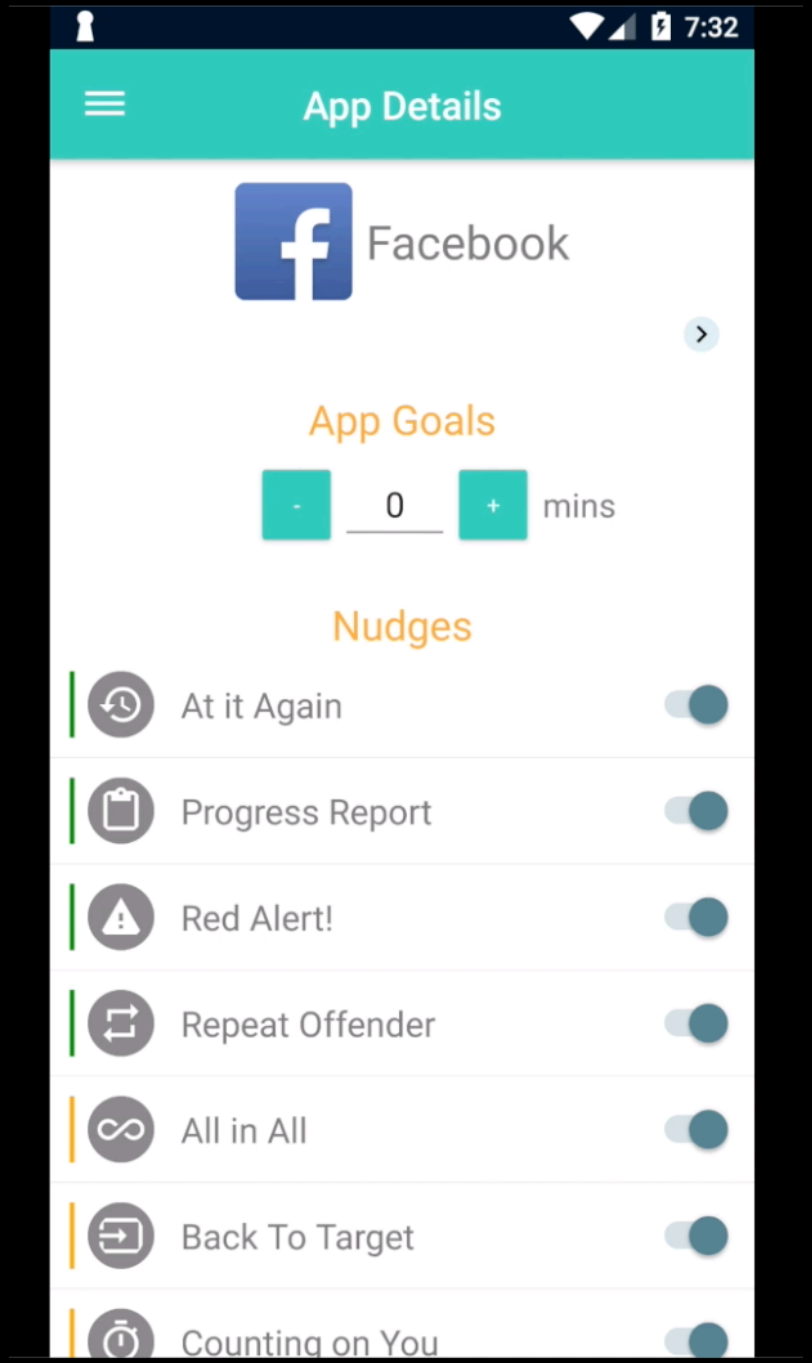


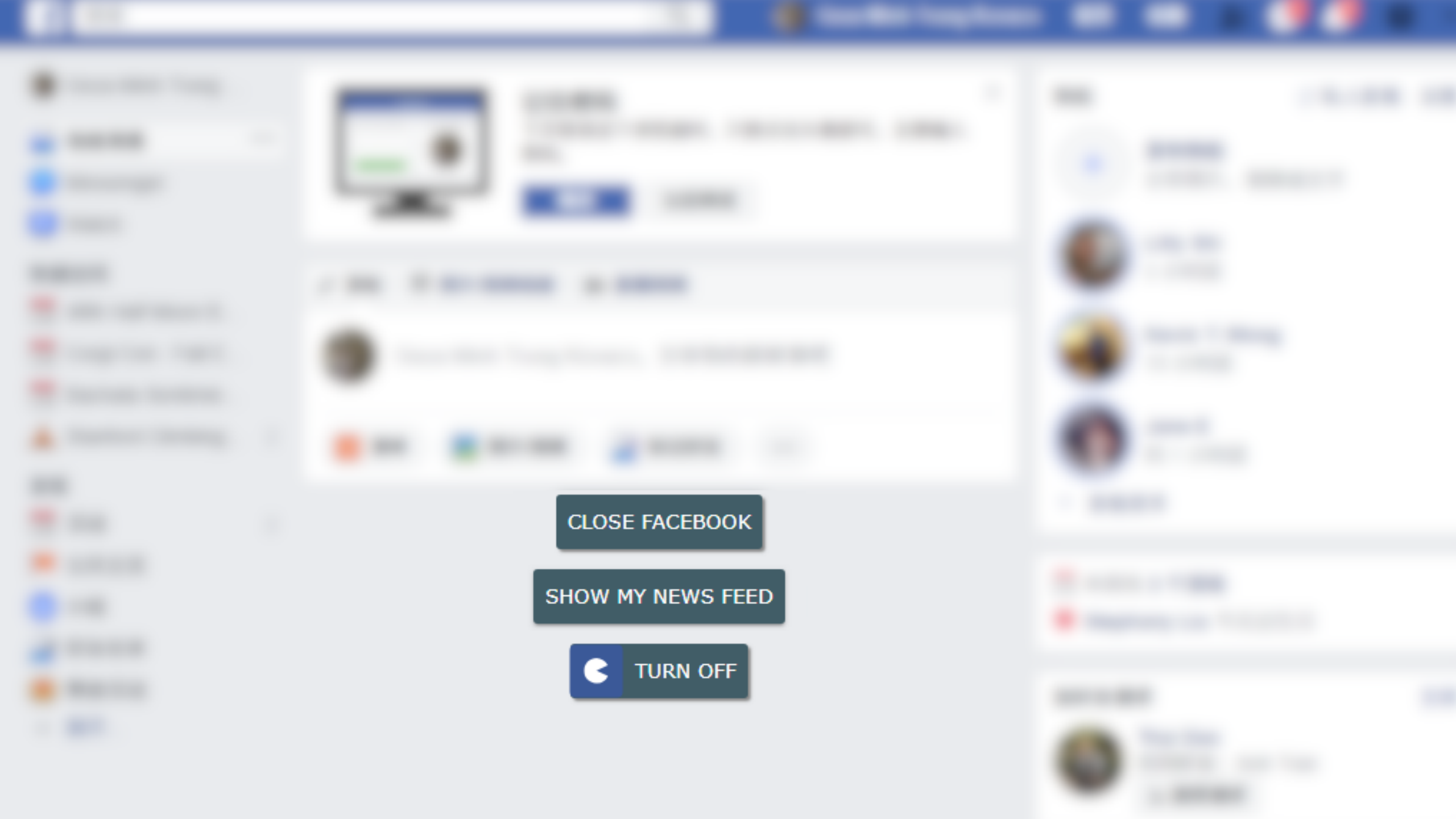
Interventions help reduce time on goal sites and apps

30+ interventions available



30+ interventions available




CLOSE FACEBOOK
SHOW MY NEWS FEED
TURN OFF

Today you have been here for:

2 min 38 sec

CLOSE FACEBOOK

TURN OFF

CLOSE TAB

Time remaining:

0 min 42 sec

ADD TIME

Time added:

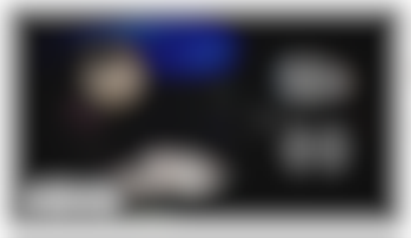
0 min

TURN OFF

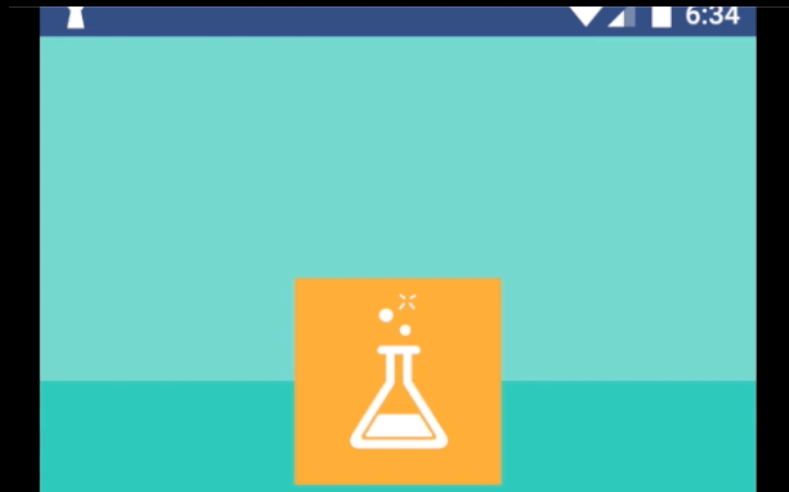
TURN OFF

This video is 3 minutes and 30 seconds long.
Are you sure you want to play it?

CLOSE YOUTUBE WATCH VIDEO


You've opened Facebook 4 times today



Continue to Facebook?

You've been here 39 times today. Want to take a break?

GET ME OUT OF HERE!

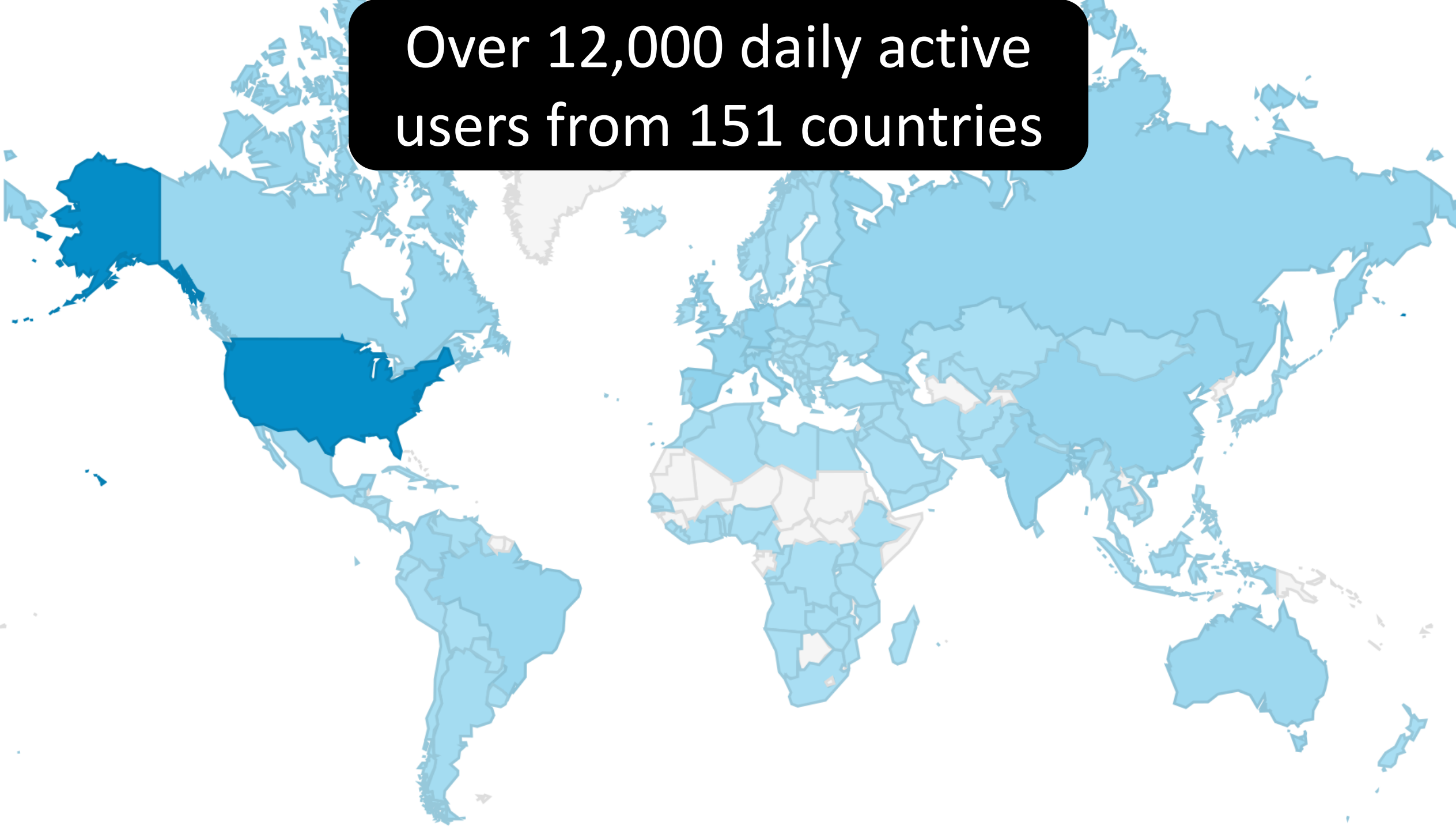
Continue to Facebook

Design process for interventions

Existing interventions on the Chrome Store

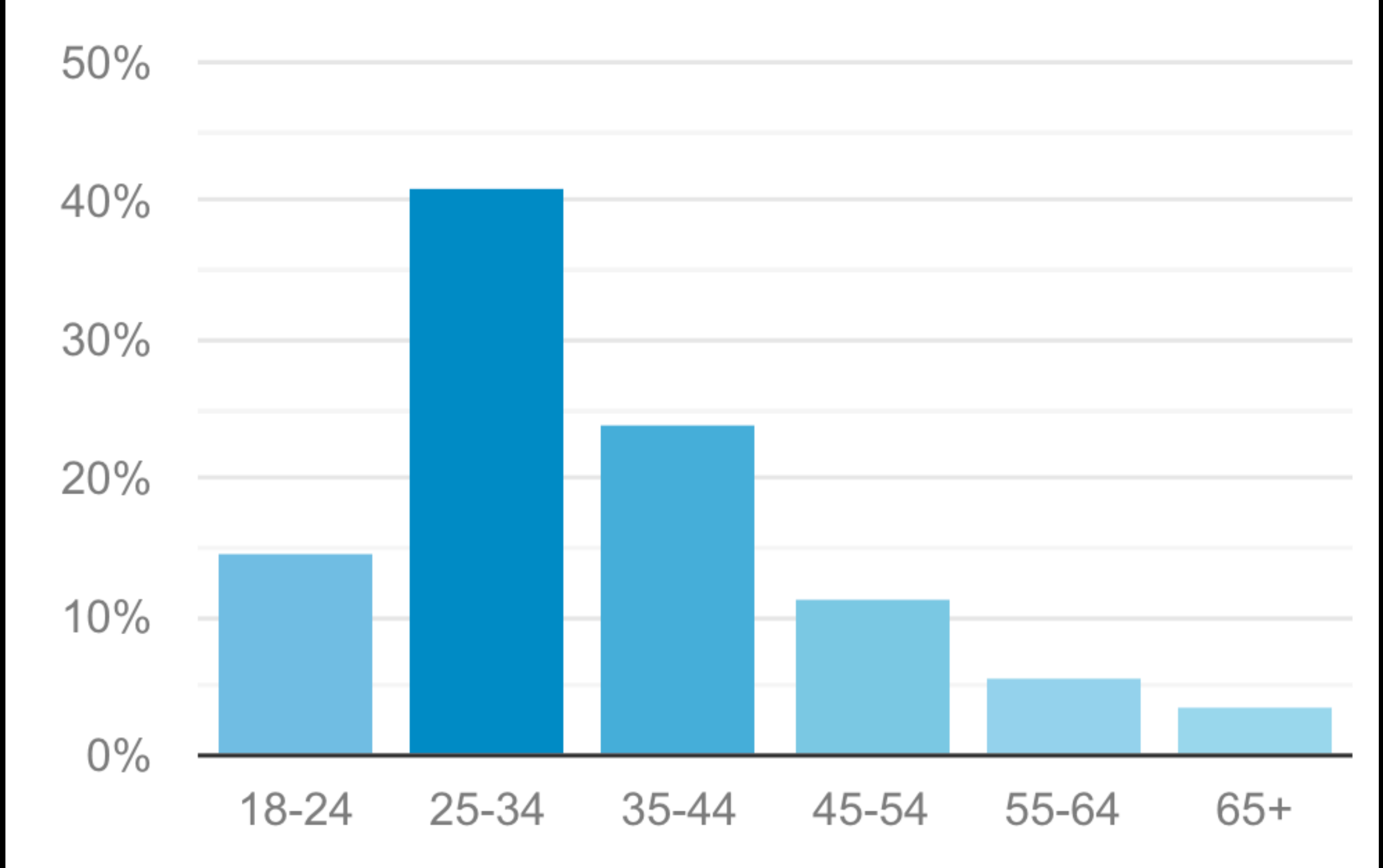
Ideas proposed by experts and users

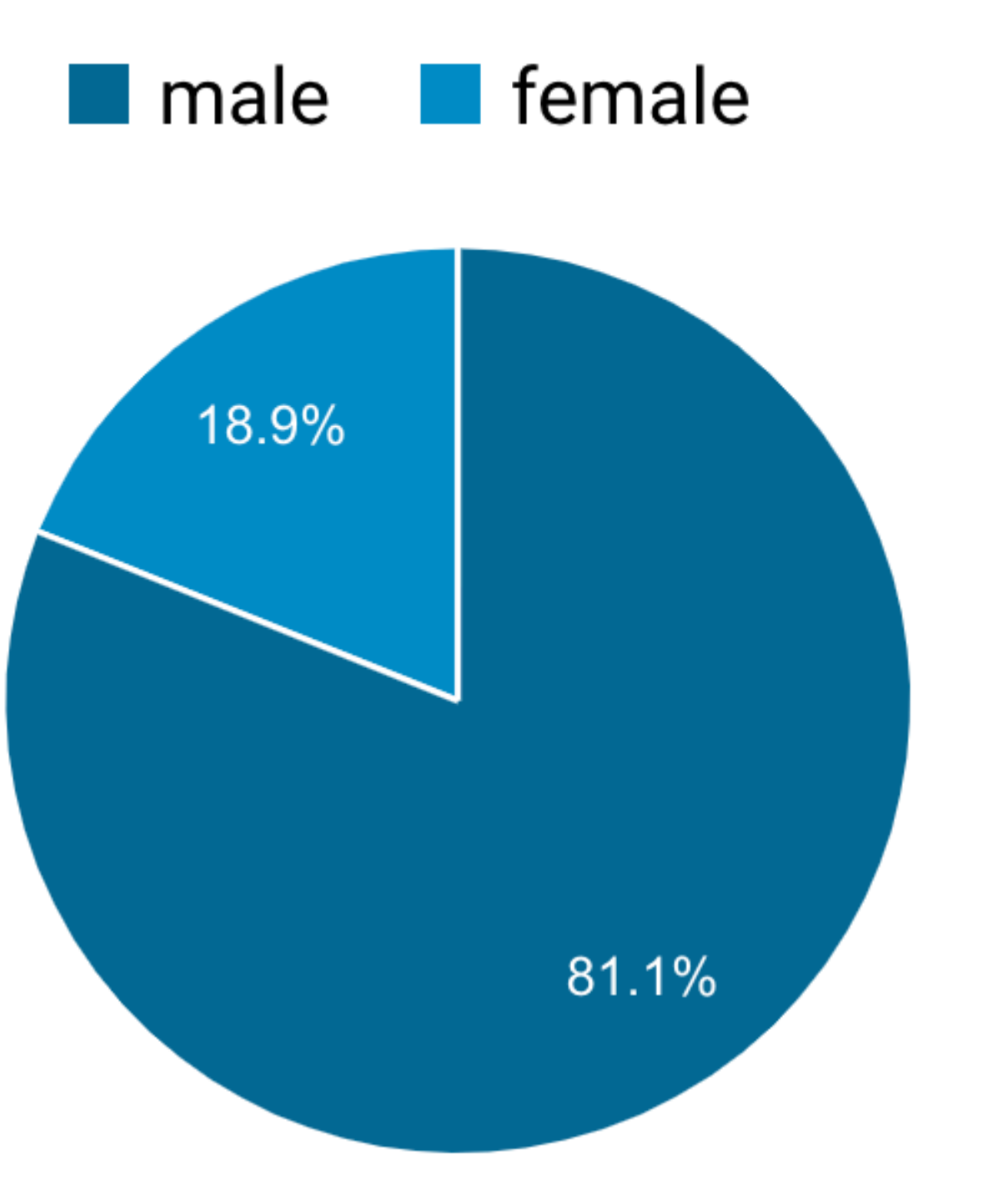
Adaptations of techniques from the literature


Over 12,000 daily active
users from 151 countries

1.	United States	3,710	30.16%
2.	Spain	605	4.92%
3.	Germany	540	4.39%
4.	Russia	461	3.75%
5.	China	450	3.66%
6.	India	448	3.64%
7.	United Kingdom	439	3.57%
8.	France	418	3.40%
9.	Italy	402	3.27%
10.	Canada	368	2.99%

1.	English (US)	6,143	50.00%
2.	Spanish	801	6.52%
3.	English (UK)	725	5.90%
4.	Chinese (Simplified)	462	3.76%
5.	Russian	443	3.61%
6.	Italian	340	2.77%
7.	Portuguese (Brazil)	293	2.38%
8.	French	289	2.35%
9.	German (Germany)	260	2.12%
10.	German	212	1.73%


50%
40%
30%
20%
10%
0%
18-24
25-34
35-44
45-54
55-64
65+


male
female
18.9%
81.1%

Designing for Growth In-the-wild

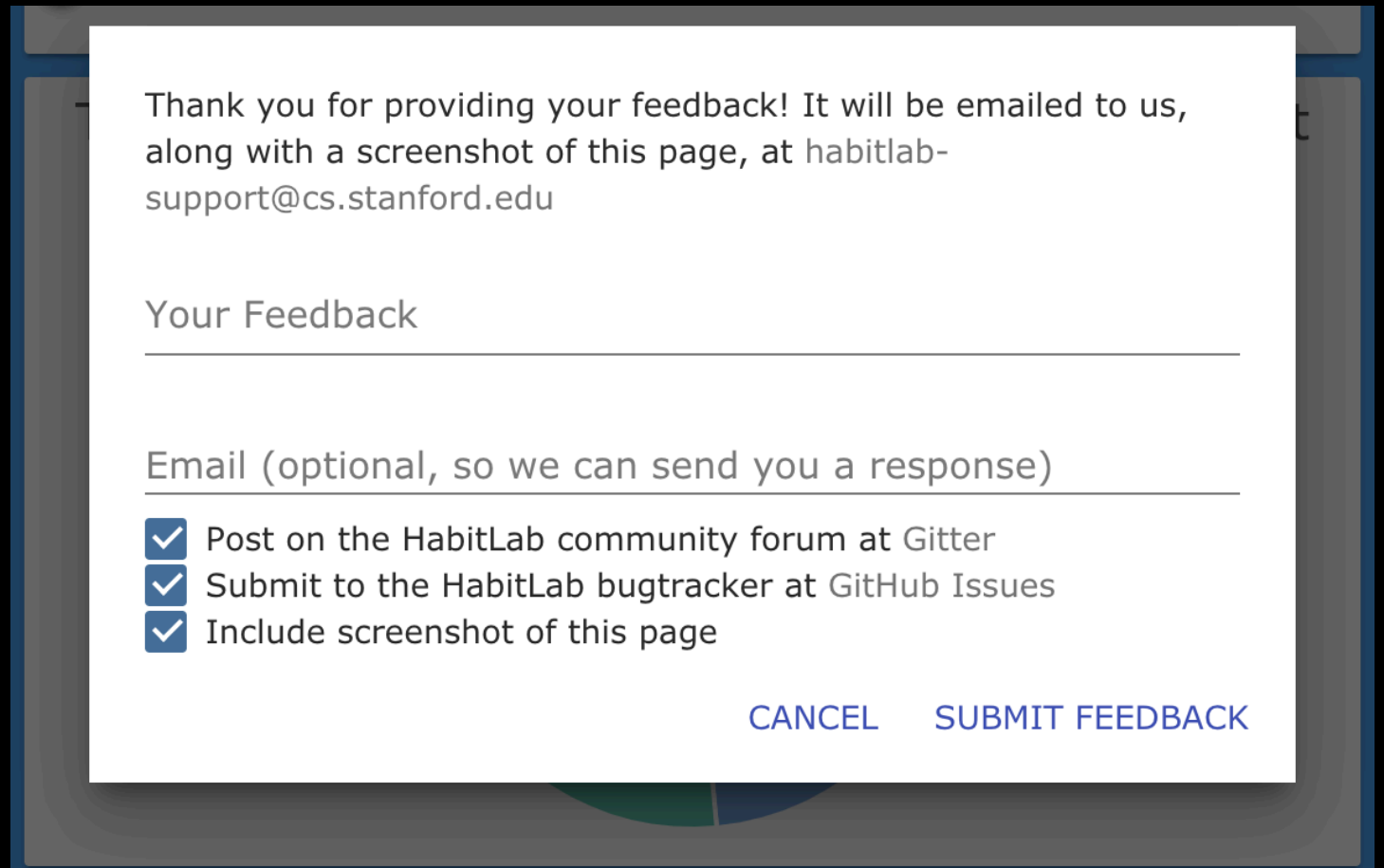
Design for good user experience: polish product, remove bugs, avoid long onboarding surveys and excessive experience sampling

Designing for Growth In-the-wild

A/B test to find good defaults that maximize retention

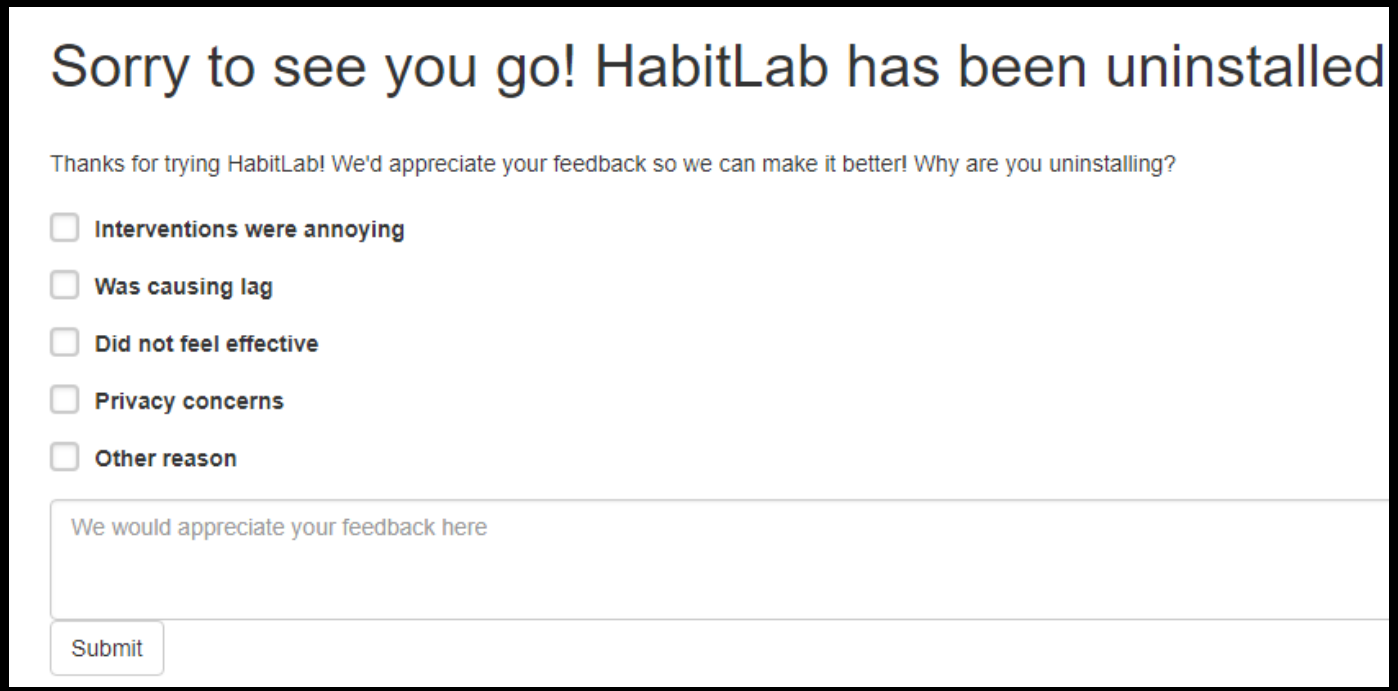
Designing for Growth In-the-wild

Give people plenty of opportunities to leave feedback



Designing for Growth In-the-wild

Give people plenty of opportunities to leave feedback



Designing for Growth In-the-wild

Give people plenty of opportunities to leave feedback

HabitLab "how aggressive" panel 收件箱 ×

[redacted]

上午1:00 (9小时前)

发送至 [redacted]

Hi Professors Kovacs, Wu, and Bernstein,

I've been enjoying HabitLab, and think it's a really nice system. I have one pet peeve though which might lead to my unfortunate disabling of it. Wanted to bring this to your attention and see if a fix could possibly be made.

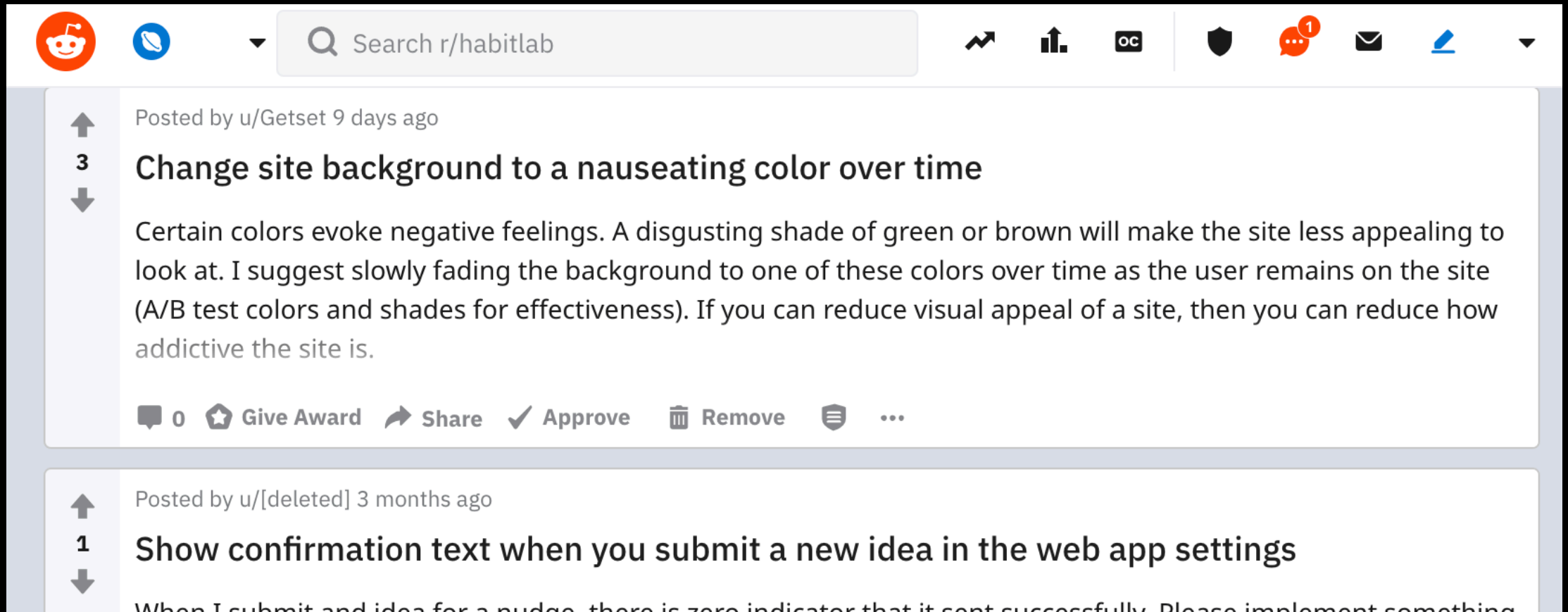
Whenever I go to a "nudged site", this "how aggressive" overlay comes up. I would not like it to. I click on "Light touch" every time, and it's so annoying that I'd sooner remove the Chrome extension than keep doing it every time. I get the thought, that maybe the annoyingness will make me visit those sites less. But if I wanted to not visit them at all, I'd just block them outright using another extension. I want to visit them, but be aware of how much time I'm spending. And I don't want additional tasks to accomplish every time.

Would the "how aggressive" panel triggering-or-not be a setting you could add? I'd really like to keep using this system!

Thank you,

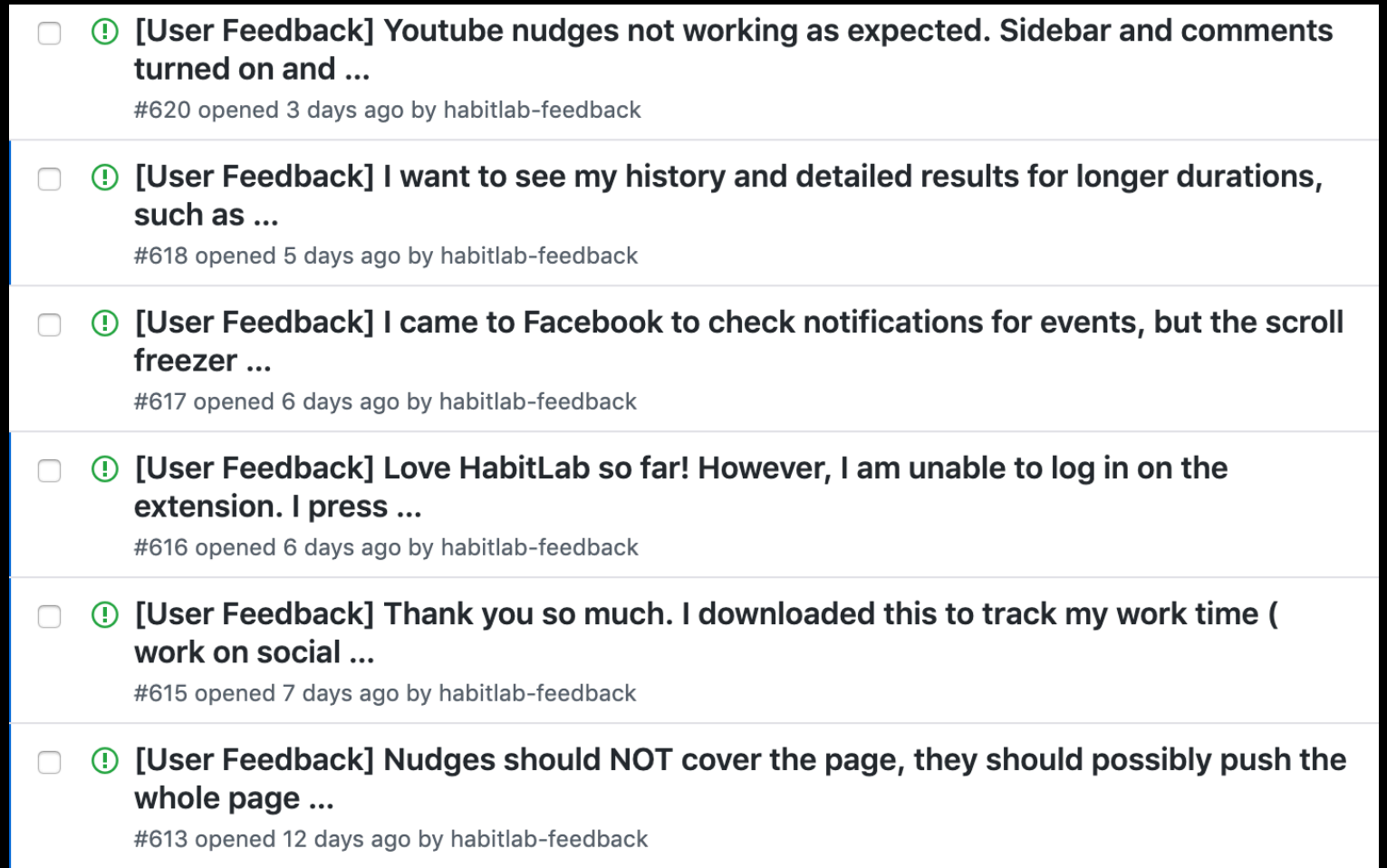
Designing for Growth In-the-wild

Give people plenty of opportunities to leave feedback



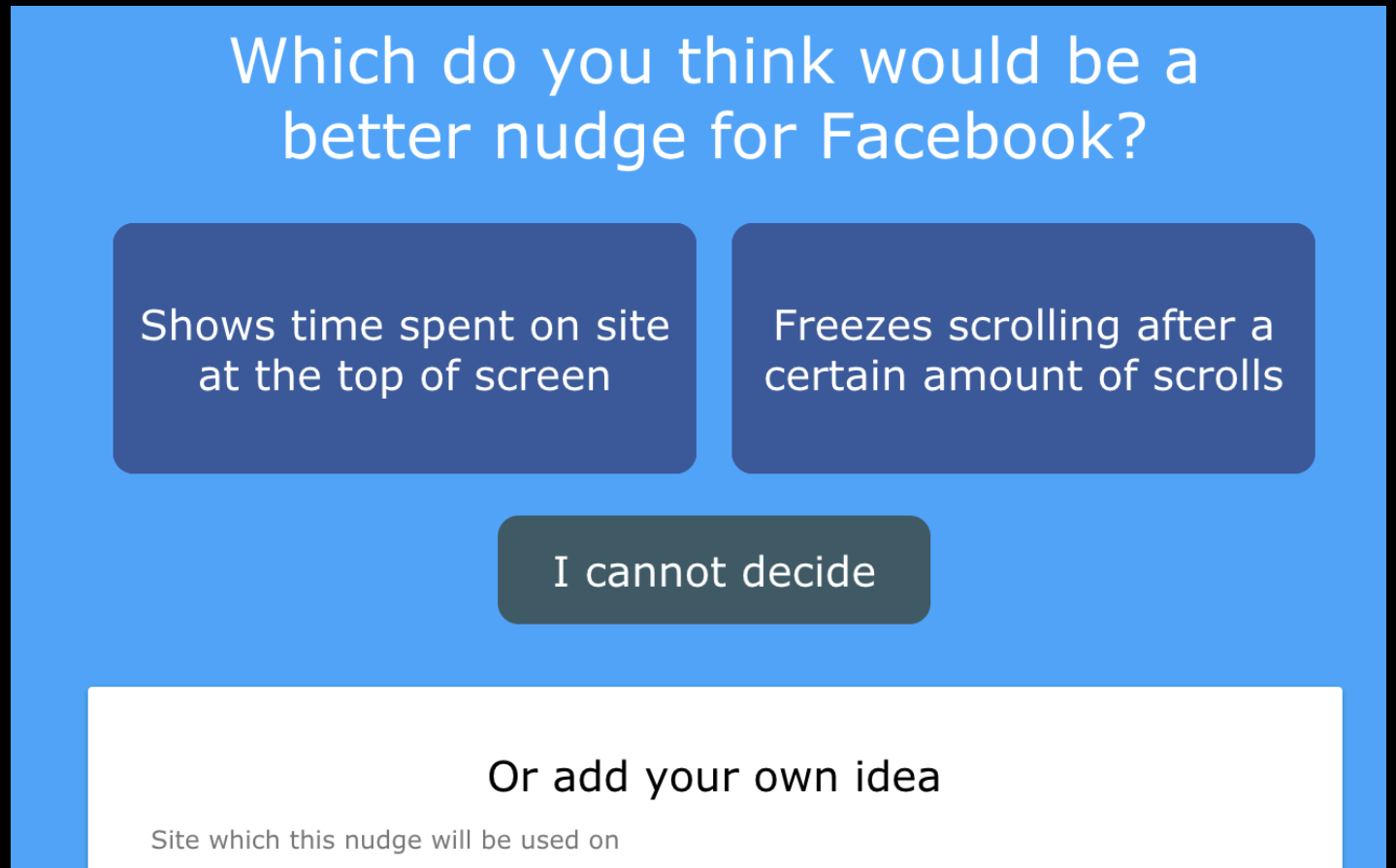
Designing for Growth In-the-wild

Give people plenty of opportunities to leave feedback



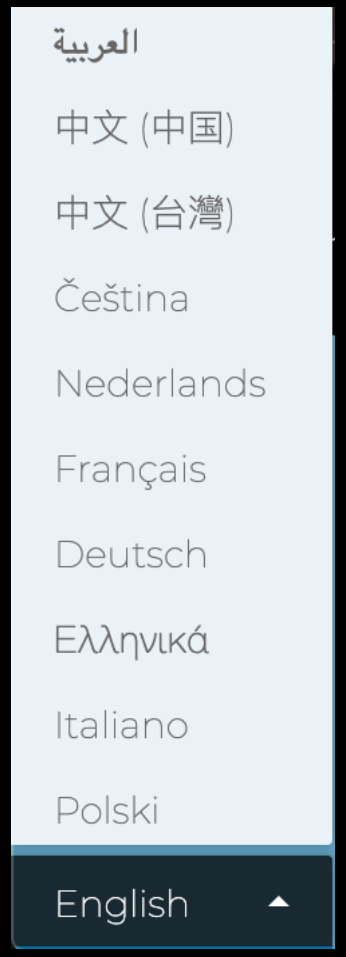
Designing for Growth In-the-wild

Give people opportunities to contribute: ideas



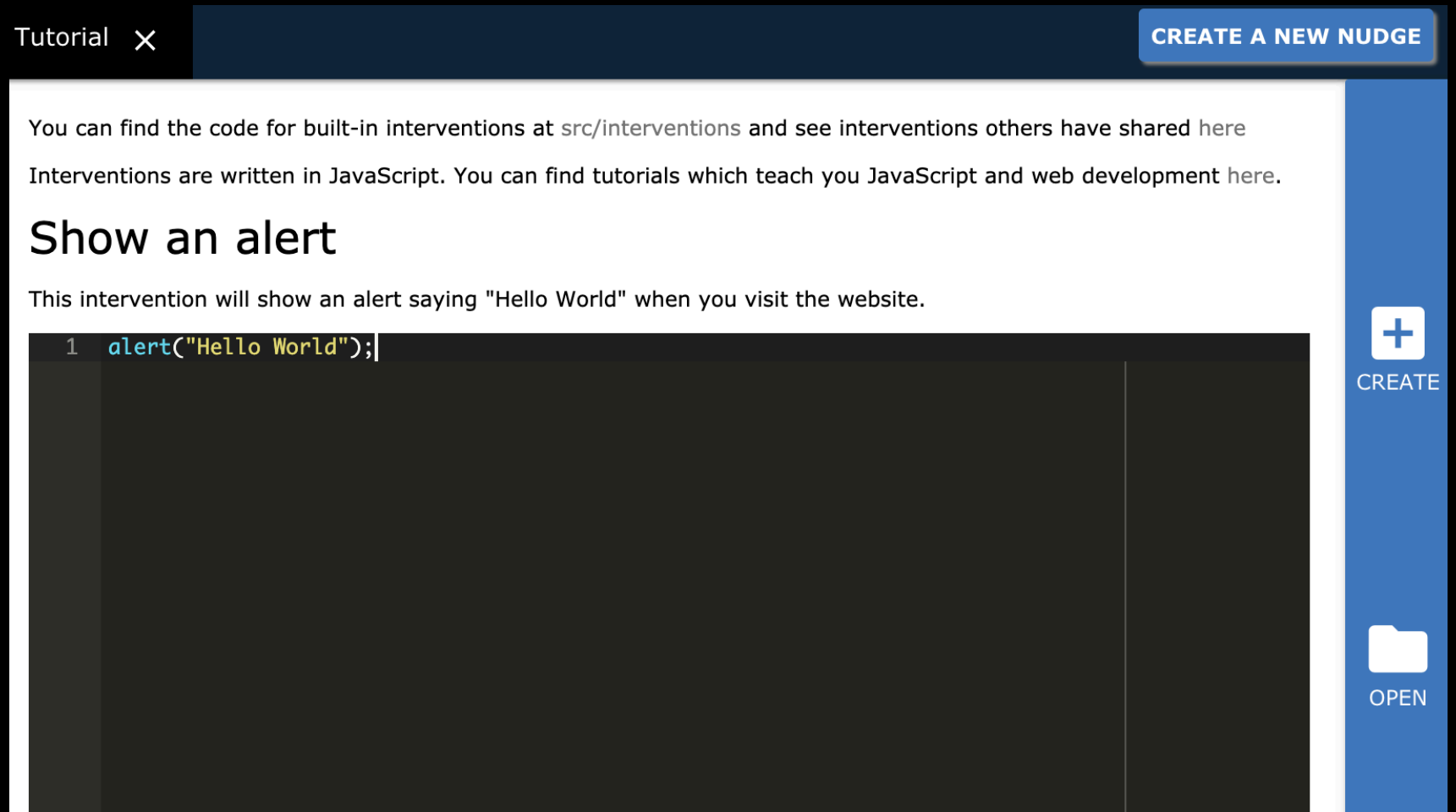
Designing for Growth In-the-wild

Give people opportunities to contribute: internationalization



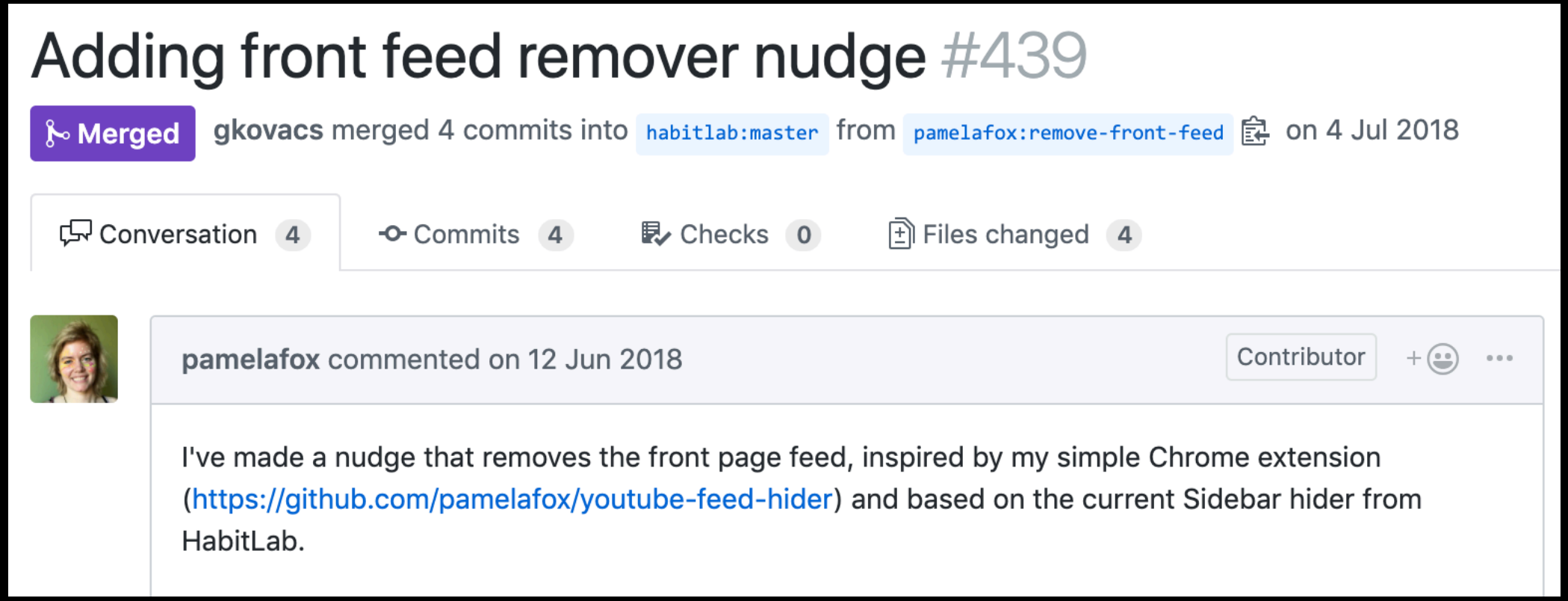
Designing for Growth In-the-wild

Give people opportunities to contribute: code



Designing for Growth In-the-wild

Give people opportunities to contribute: code

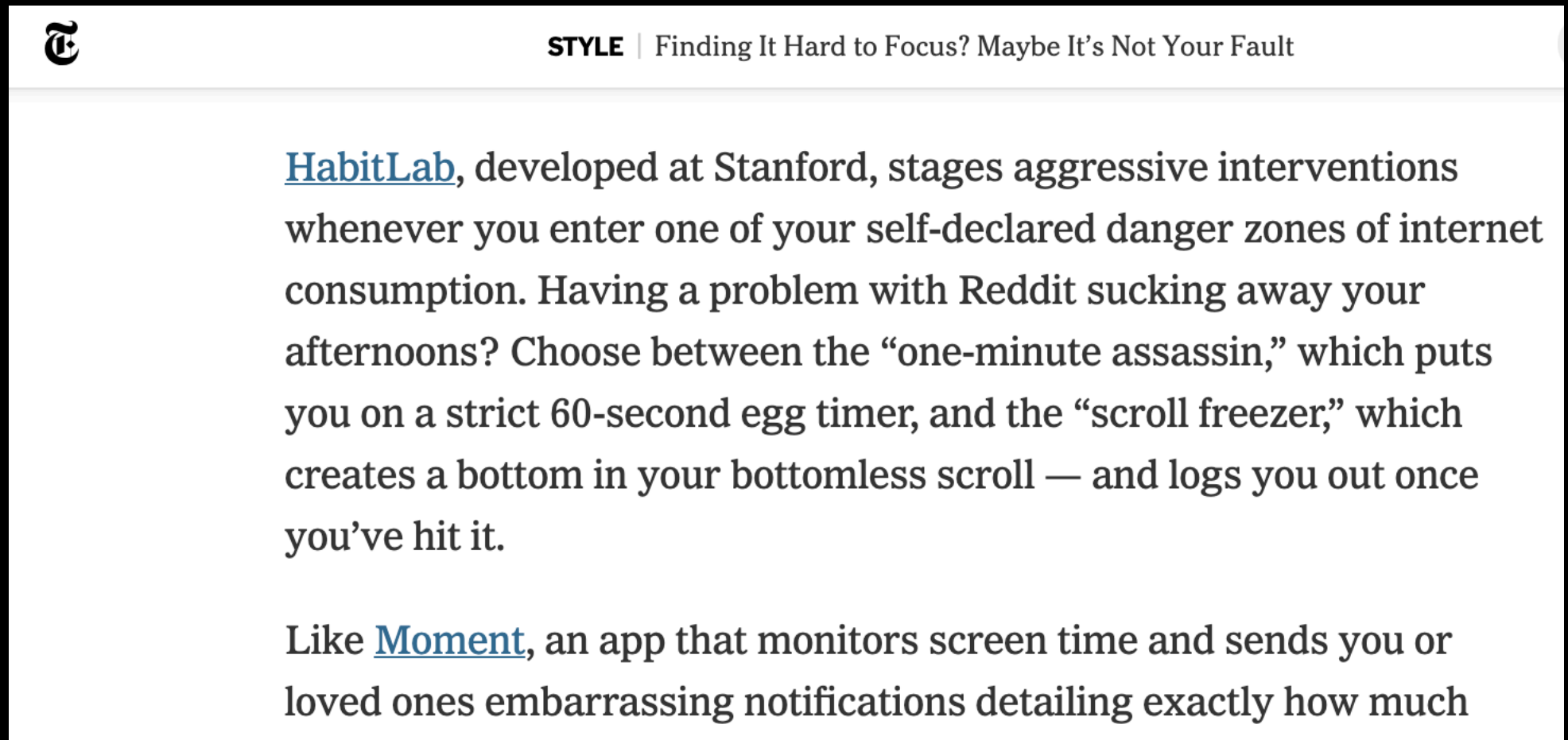


Designing for Growth In-the-wild

Ultimately, a lot depends on press and factors outside your control

Designing for Growth In-the-wild

Ultimately, a lot depends on press and factors outside your control



STYLE | Finding It Hard to Focus? Maybe It's Not Your Fault

HabitLab, developed at Stanford, stages aggressive interventions whenever you enter one of your self-declared danger zones of internet consumption. Having a problem with Reddit sucking away your afternoons? Choose between the "one-minute assassin," which puts you on a strict 60-second egg timer, and the "scroll freezer," which creates a bottom in your bottomless scroll — and logs you out once you've hit it.

Like Moment, an app that monitors screen time and sends you or loved ones embarrassing notifications detailing exactly how much

Designing for Growth In-the-wild

Ultimately, a lot depends on press and factors outside your control

WIRED

The HabitLab Browser Extension Curbs Your Time Wasted on the Web

SHAR
E

SHARE

TWEET

FOR GEZA KOVACS, our collective time-wasting on the web makes for precious data. A PhD candidate in Stanford's human-computer interaction group, Kovacs studies bad browsing habits and researches what can be done to repair them. Like, when you flick open a new tab and reflexively navigate to Facebook, does it help to be reminded that you have other stuff to do today? Would you consider closing the tab if you saw a stopwatch, tick-tocking to remind you of how much time you've lost? And when you close a tab on

Designing for Growth In-the-wild

Ultimately, a lot depends on press and factors outside your control

HabitLab 來自史丹佛大學的時間管理工具，實驗改善上網習慣

作者： Esor Huang 8月 06, 2018

✓ 赞 分享 你和其他 147 位用户赞了

腾讯网 要闻 娱乐 财经 一键登录

这款斯坦福大学的工具，让你远离加班，提升200%效率！

从古至今，人们一直无法摆脱拖延症的困扰。

胡适在他的《胡适留学日记》上，刚说要计划“读完《亨利八世》”，第二天就开始疯狂打起了牌，连着打了好几周。

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition (CSCW 2018)

Behavior change interventions suffer from declined engagement over time

Paul Krebs, James O Prochaska, and Joseph S Rossi. 2010. A meta-analysis of computer-tailored interventions for health behavior change. Preventive medicine 51, 3-4 (2010), 214–221

Novelty effects can provide temporary boosts in engagement

Reza Kormi nouri, Lars Goran Nilson, and Nobuo Ohta. [n. d.]. The novelty effect: Support for the Novelty Encoding Hypothesis. Scandinavian Journal of Psychology

Existing behavior change systems tend to use static interventions

[RQ] Can a strategy of rotating interventions improve effectiveness?

[H1] Static interventions suffer from decreased effectiveness over time


[H2] Rotation will increase intervention effectiveness

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition (CSCW 2018)

Rotating vs static intervention strategies

Study 1: Within subjects

Study 2: Between subjects

Qualitative feedback

Study 3: Improving users' mental models about rotating interventions

Compare rotating and static intervention strategies, in terms of:

Effectiveness of interventions over time (daily time on sites)

Attrition rates (time until uninstall)

Within-subjects design, 217 participants

Within-subjects design, 217 participants

Conditions: on some days, users saw the same intervention (static), on others, interventions changed each visit (rotation)

Within-subjects design, 217 participants

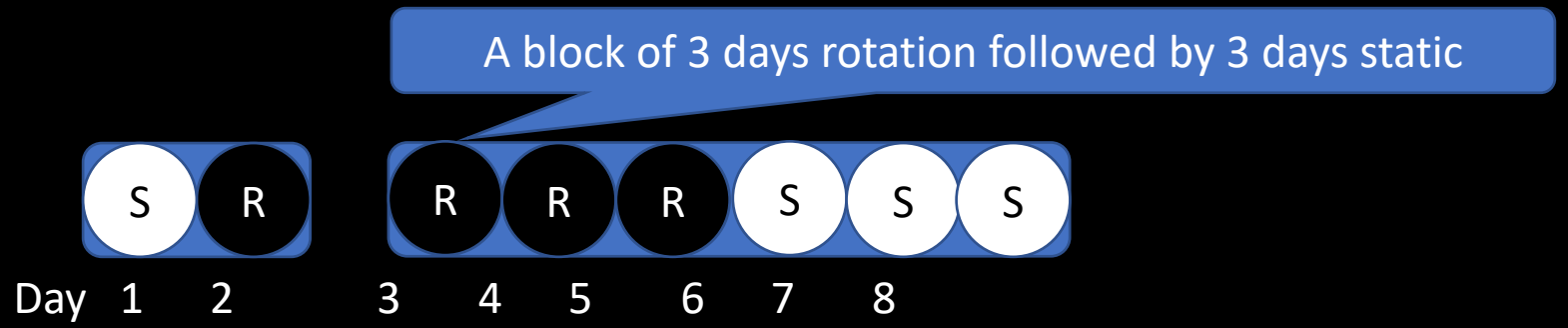
Conditions: on some days, users saw the same intervention (static), on others, interventions changed each visit (rotation)

Conditions were organized into blocks of 1, 3, 5, 7 days

Within-subjects design, 217 participants

Conditions: on some days, users saw the same intervention (static), on others, interventions changed each visit (rotation)

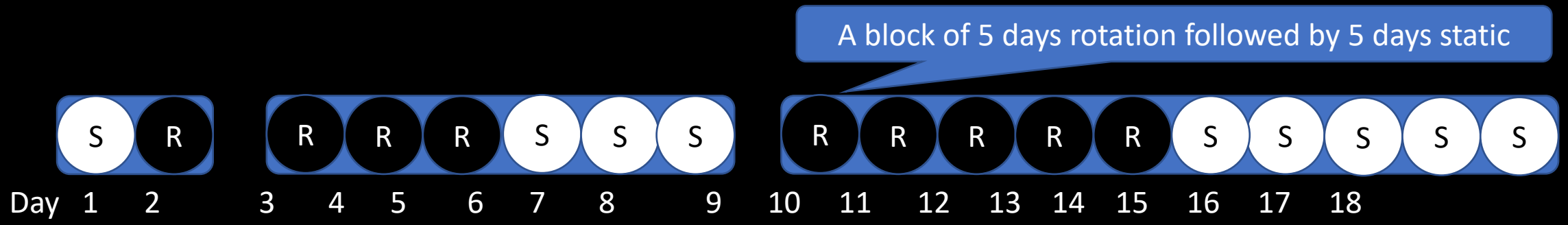
Conditions were organized into blocks of 1, 3, 5, 7 days



Within-subjects design, 217 participants

Conditions: on some days, users saw the same intervention (static), on others, interventions changed each visit (rotation)

Conditions were organized into blocks of 1, 3, 5, 7 days



Do static interventions decline in effectiveness over time?

Linear Mixed Model

Fixed effects:

#days intervention seen

Random effects:

User ID, Domain

Dependent variable:

Time spent on domain that day (log)

Time spent on sites increases over time with static interventions (decline in effectiveness)

	Log time spent per day (dependent variable)
# days static intervention seen	0.225 ($p < 0.05$)
Intercept	4.759
Observations	124

Interpretation (via exponentiation):

Day 1: 116 seconds per site
Day 2: 146 seconds per site
Day 3: 183 seconds per site

Does rotation reduce time spent?

Linear Mixed Model

Fixed effects:

Condition (static or rotation), Block length

Random effects:

User ID, Domain

Dependent variable:

Time spent on domain that day (log)

Daily time on sites reduced in the rotation condition

	Log time spent per day (Dependent variable)
Rotation (baseline: static)	-0.417 ($p < 0.05$)
Intercept	4.981
Observations	370

Interpretation (via exponentiation):

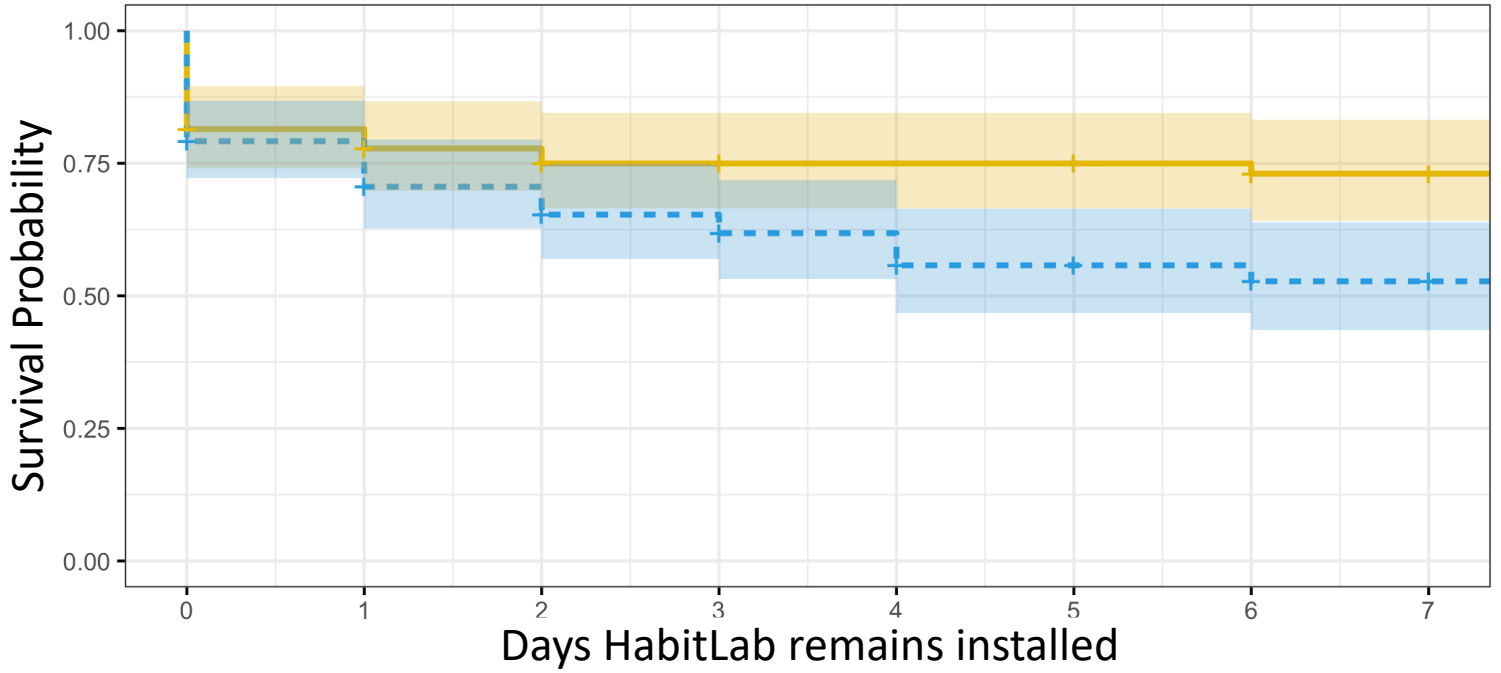
Static: 146 seconds per site daily
Rotation: 96 seconds per site daily

Does rotation increase attrition?

Cox hazard regression

Predicting survival probability as a function of condition (static or rotation), within the first block

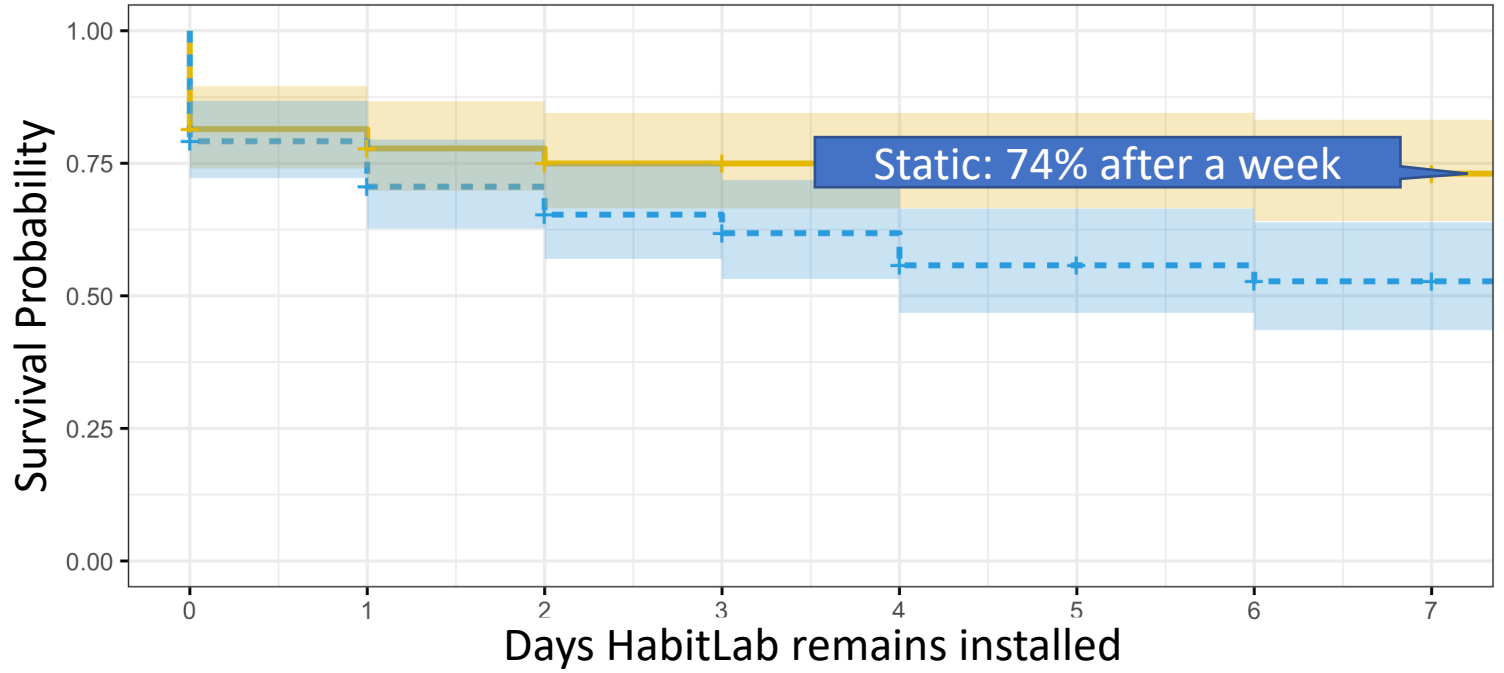
Rotating interventions increases attrition



Static interventions
Rotating interventions

Log hazard ratio for rotation condition:
0.544 ($p < 0.05$)

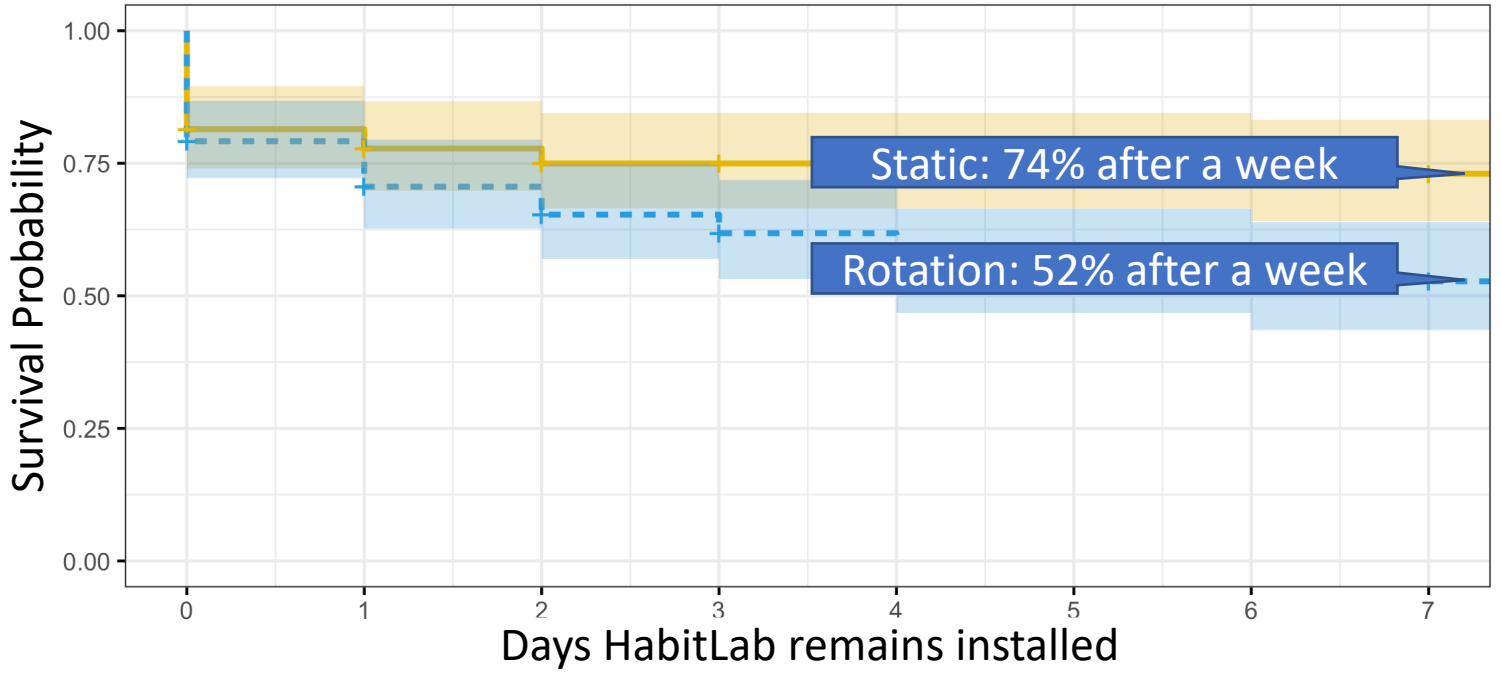
Rotating interventions increases attrition



Static interventions
Rotating interventions

Log hazard ratio for rotation condition:
0.544 ($p < 0.05$)

Rotating interventions increases attrition



Static interventions
Rotating interventions

Log hazard ratio for rotation condition:
0.544 ($p < 0.05$)

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

Qualitative feedback

Study 3: Improving users' mental models about rotating interventions

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

Qualitative feedback

Study 3: Improving users' mental models about rotating interventions

Does rotating between more interventions increase attrition?

Does rotating between more interventions increase attrition?

Between-subjects design, 409 participants, 5 weeks

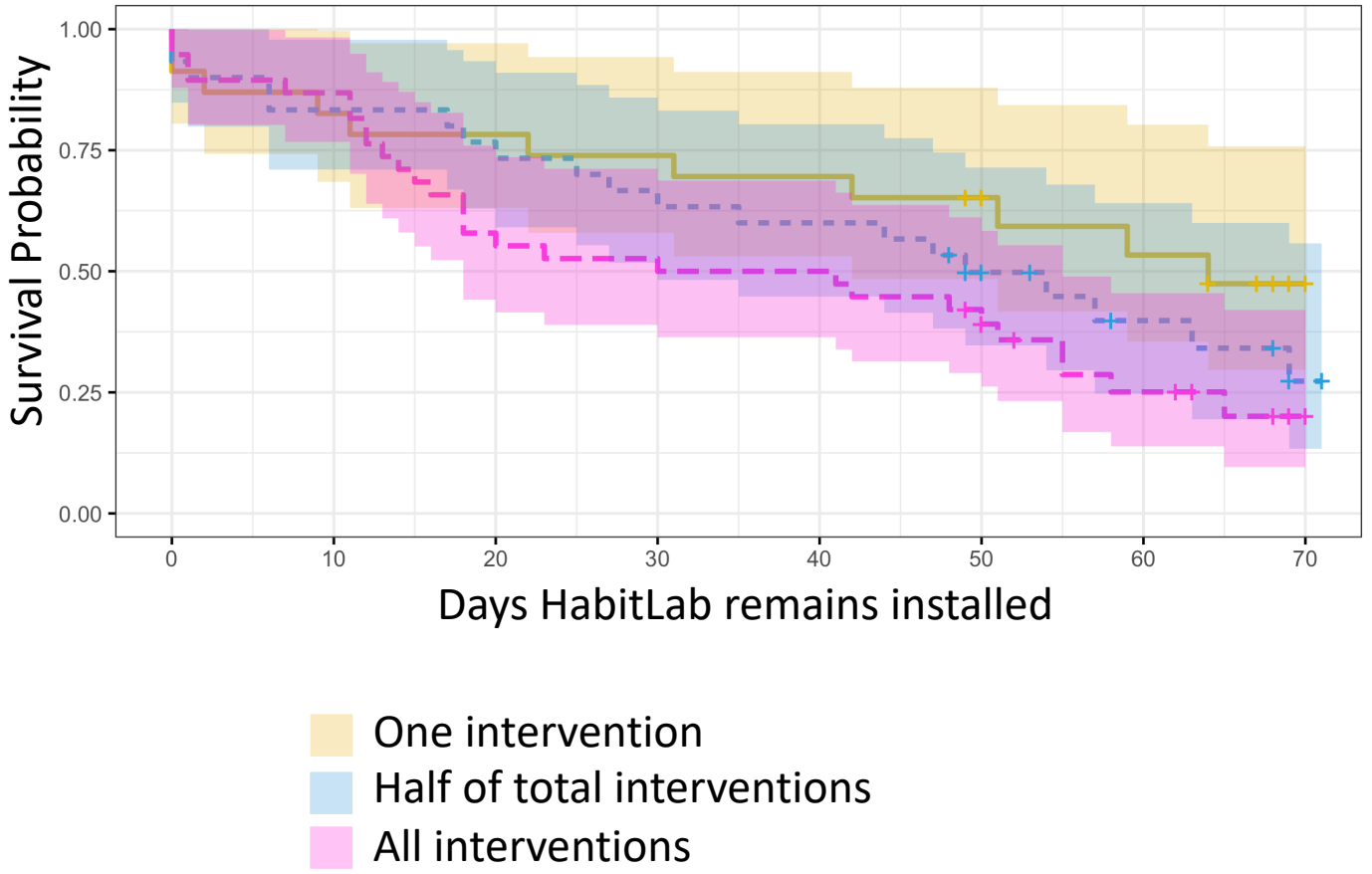
Does rotating between more interventions increase attrition?

Between-subjects design, 409 participants, 5 weeks

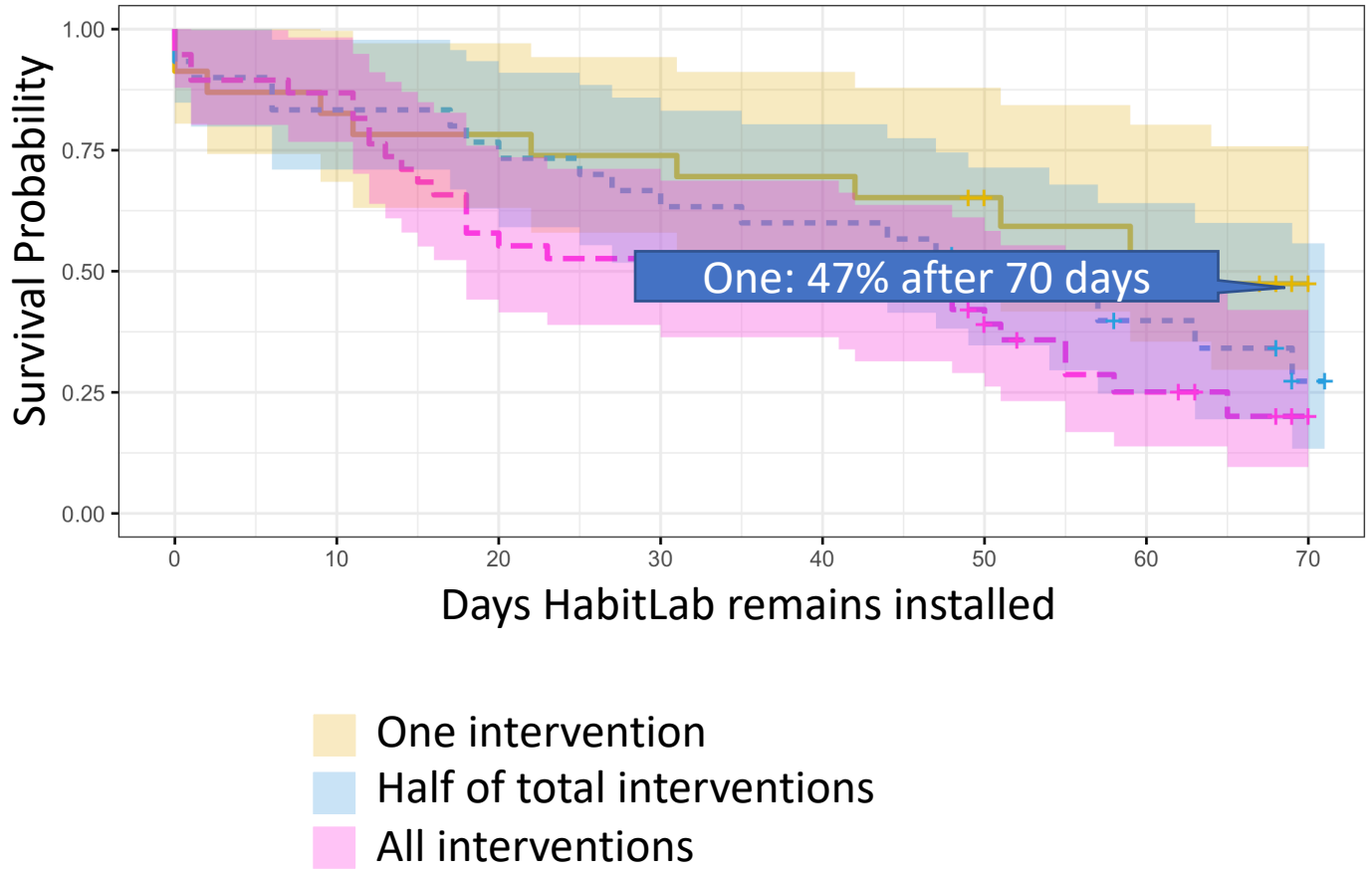
Conditions differ in number of interventions being rotated:

- One intervention per site
- Half of all available interventions per site (ie, 4 on Facebook)
- All available interventions per site (ie, 8 on Facebook)

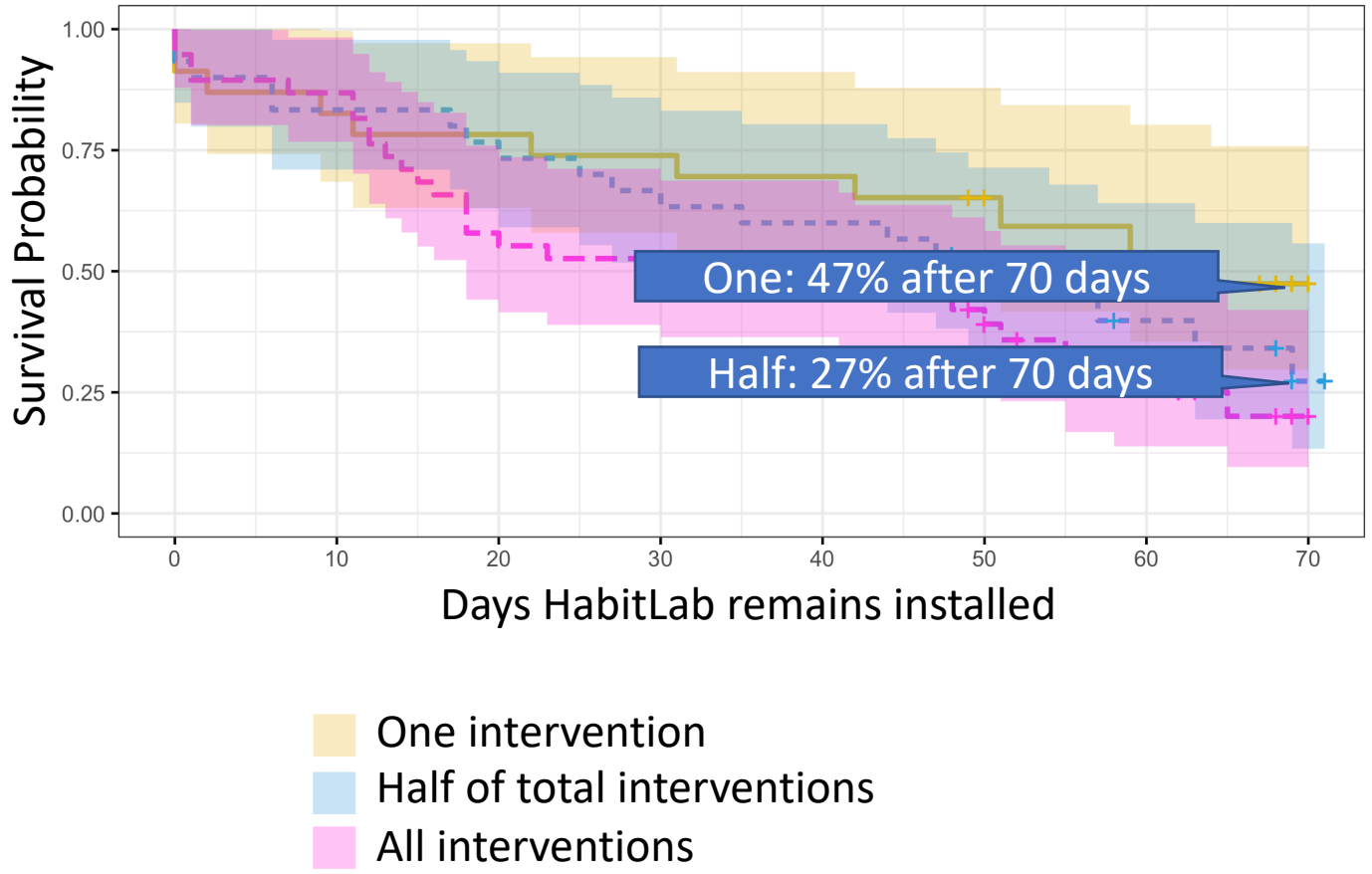
Rotating between interventions increases attrition



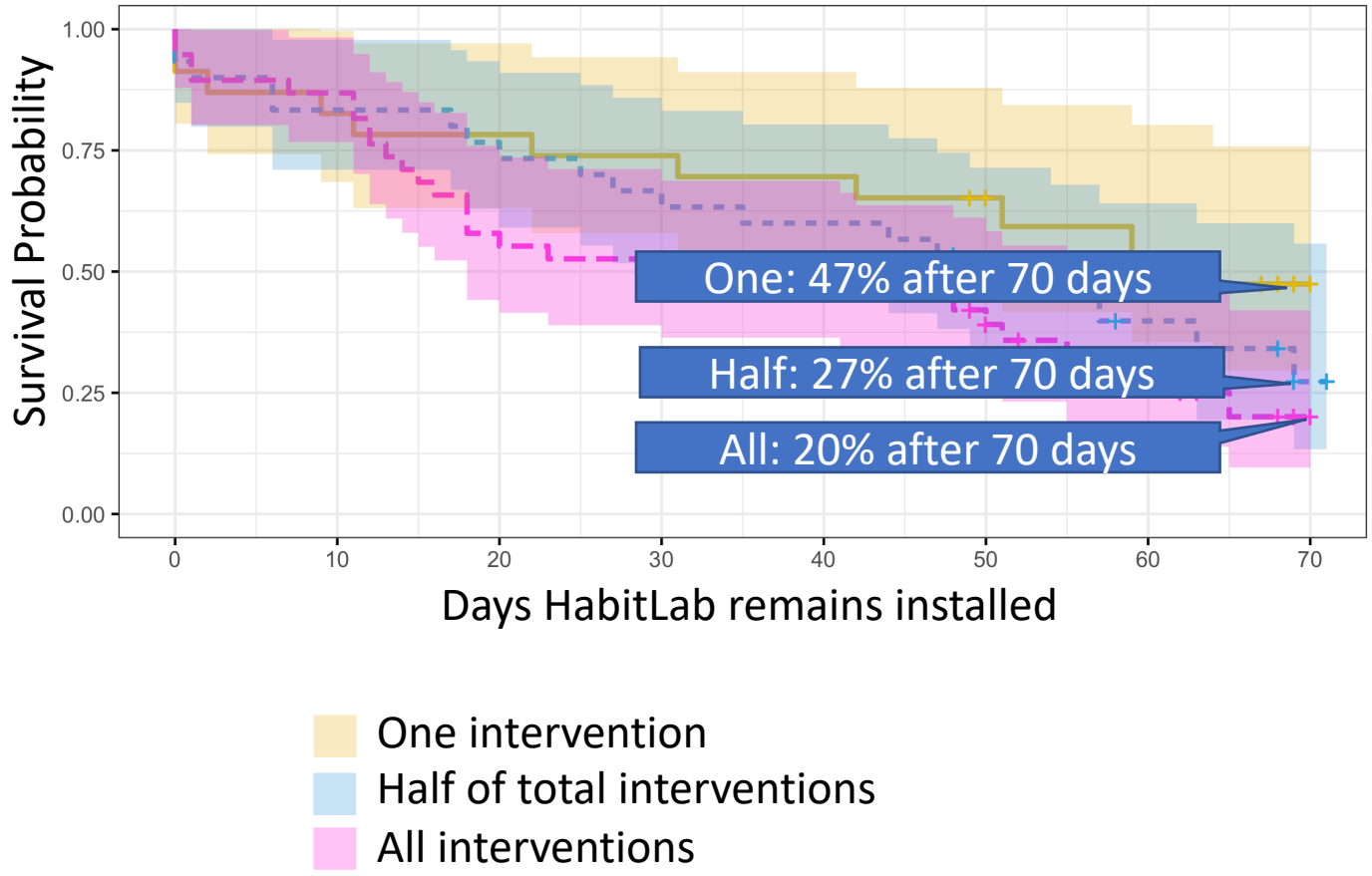
Rotating between interventions increases attrition



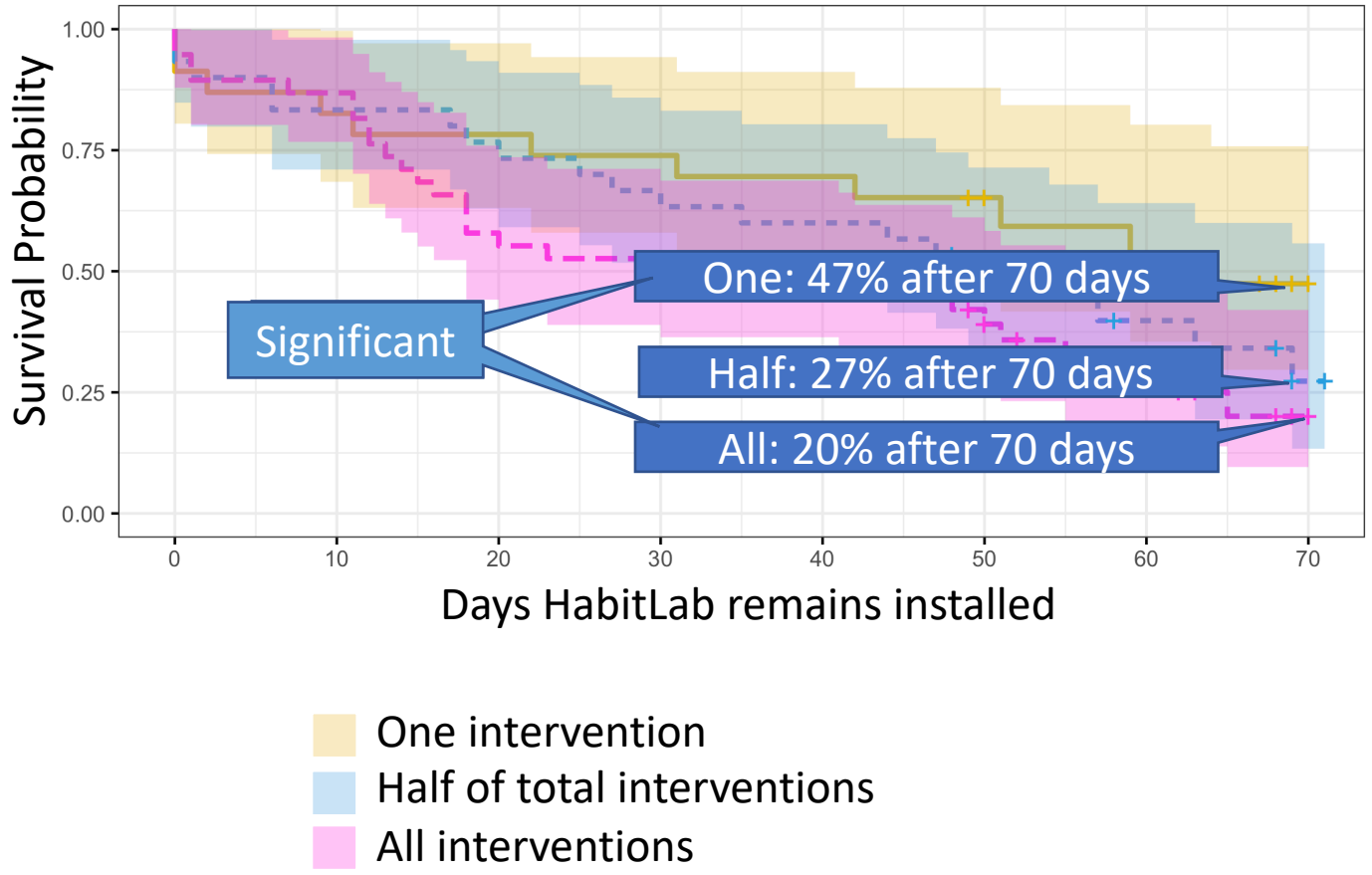
Rotating between interventions increases attrition



Rotating between interventions increases attrition



Rotating between interventions increases attrition



HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

- Rotation itself causes attrition

Qualitative feedback

Study 3: Improving users' mental models about rotating interventions

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

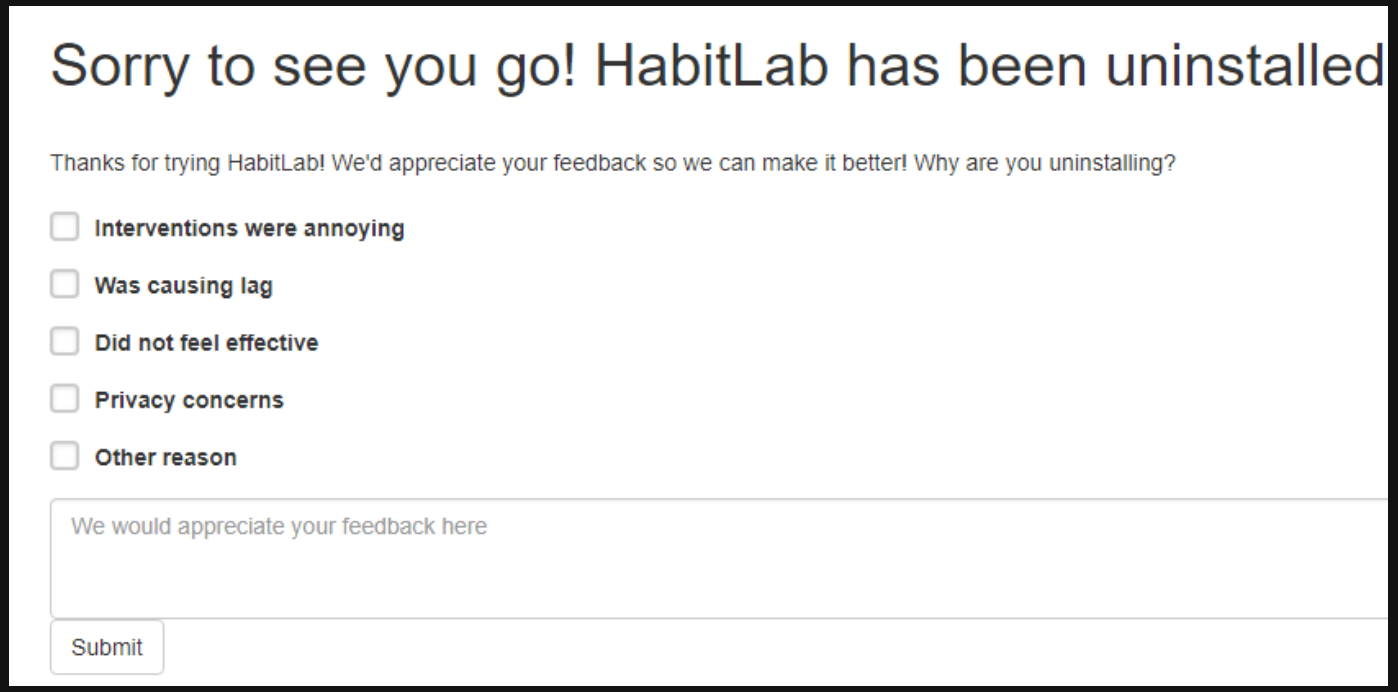
Study 2: Between subjects

- Rotation itself causes attrition

Qualitative feedback

Study 3: Improving users' mental models about rotating interventions

Collected uninstall feedback to understand reasons for attrition



Collected uninstall feedback to understand reasons for attrition

Incorrect mental models

Didn’t seem what I was expected. Installed two minutes ago and removed it

Dissatisfaction with particular interventions

Mostly it was the bar covering up facebook message indicators

Why did rotating interventions increase attrition?

Violation of mental models

Users lack sense of control

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

- Rotation itself causes attrition

Qualitative feedback

- Rotation violates mental models

Study 3: Improving users' mental models about rotating interventions

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

- Rotation itself causes attrition

Qualitative feedback

- Rotation violates mental models

Study 3: Improving users' mental models about rotating interventions

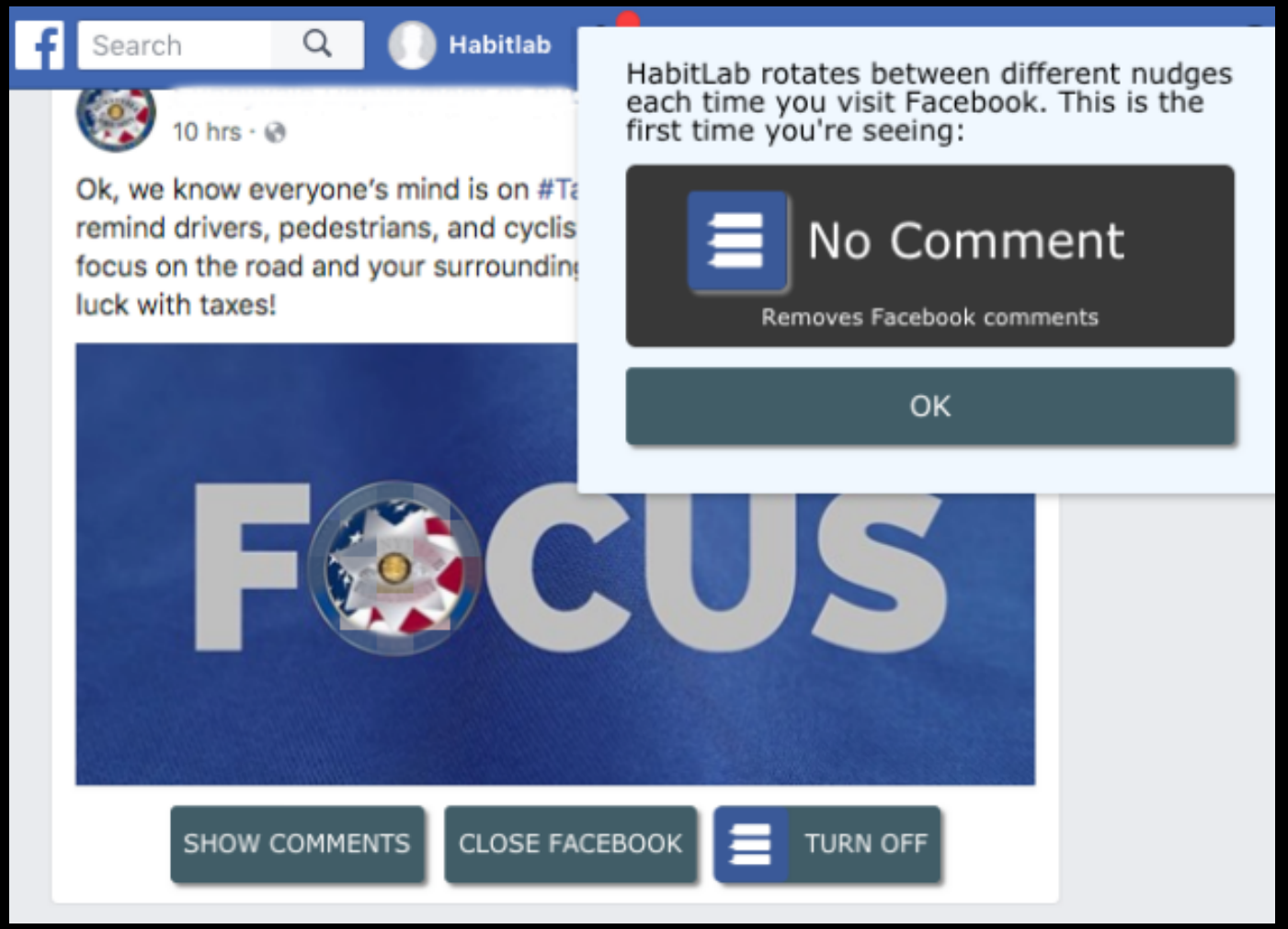
Can we reduce attrition when intervention rotation happens?

Developed 2 dialogs shown when an intervention is first seen

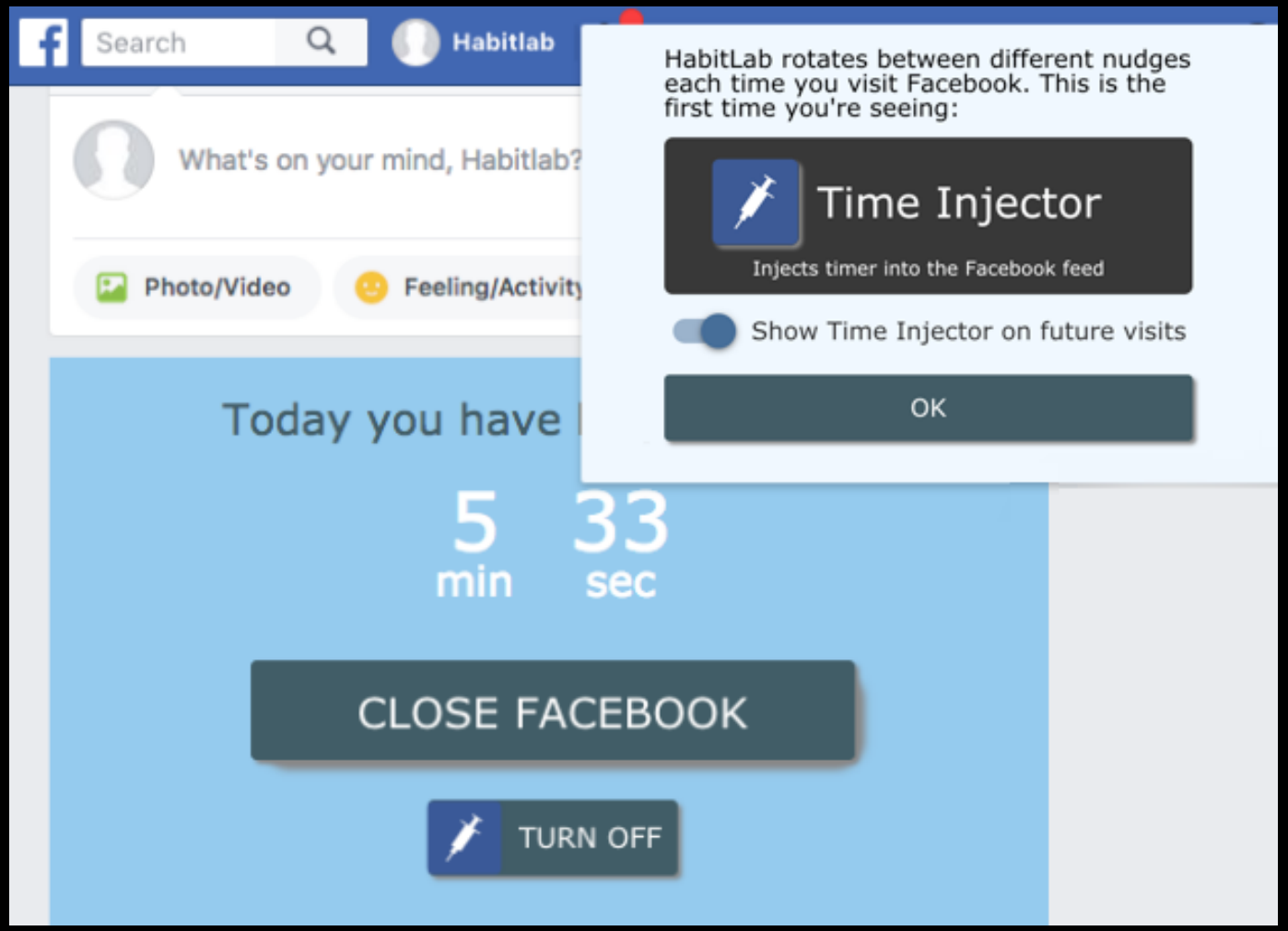
Mental model design

User control design

Mental model design



User control design



Between subjects design, 282 participants, 10 days

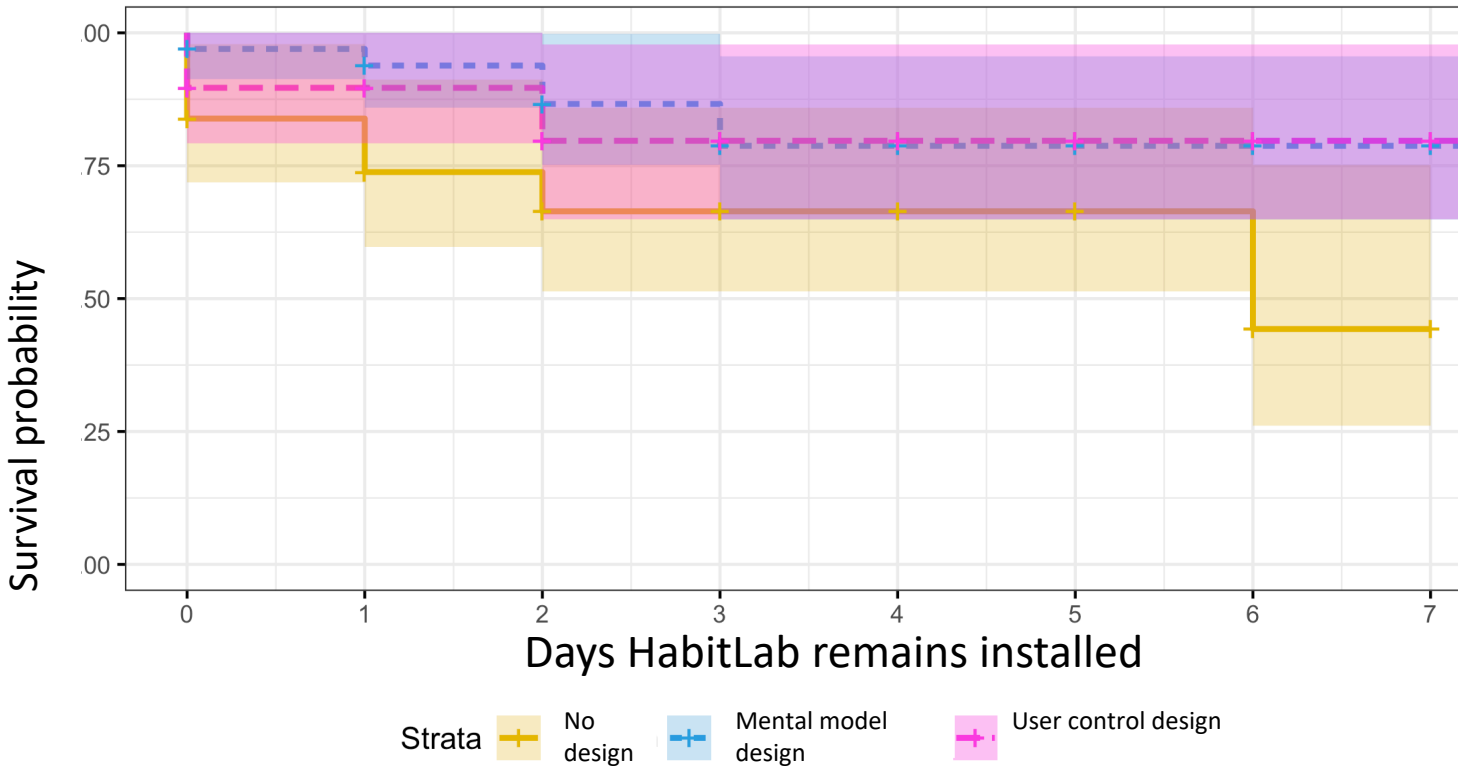
Conditions differ according to which design is shown when an intervention is seen for the first time:

None: No design shown

Mental model design

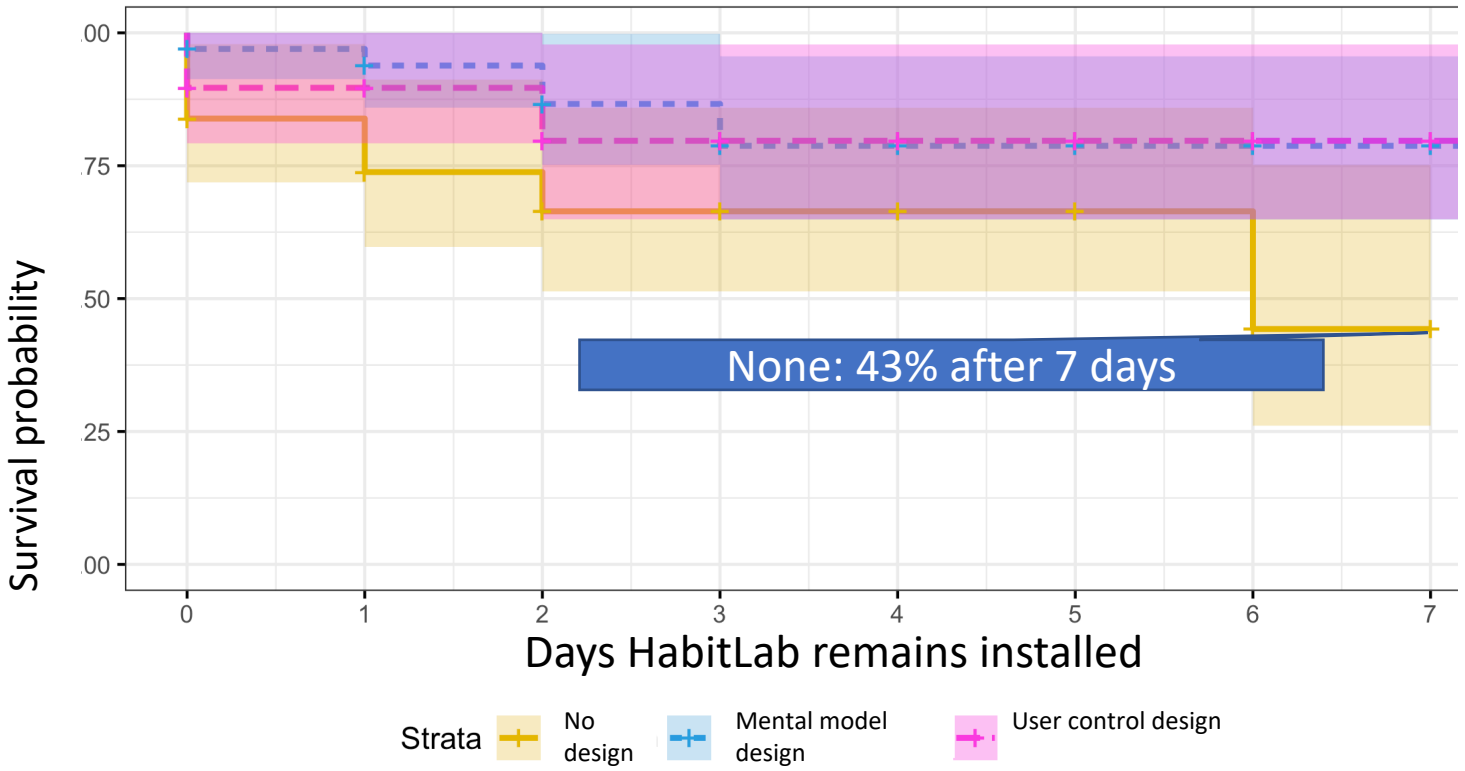
User control design

Mental model design reduces attrition by half



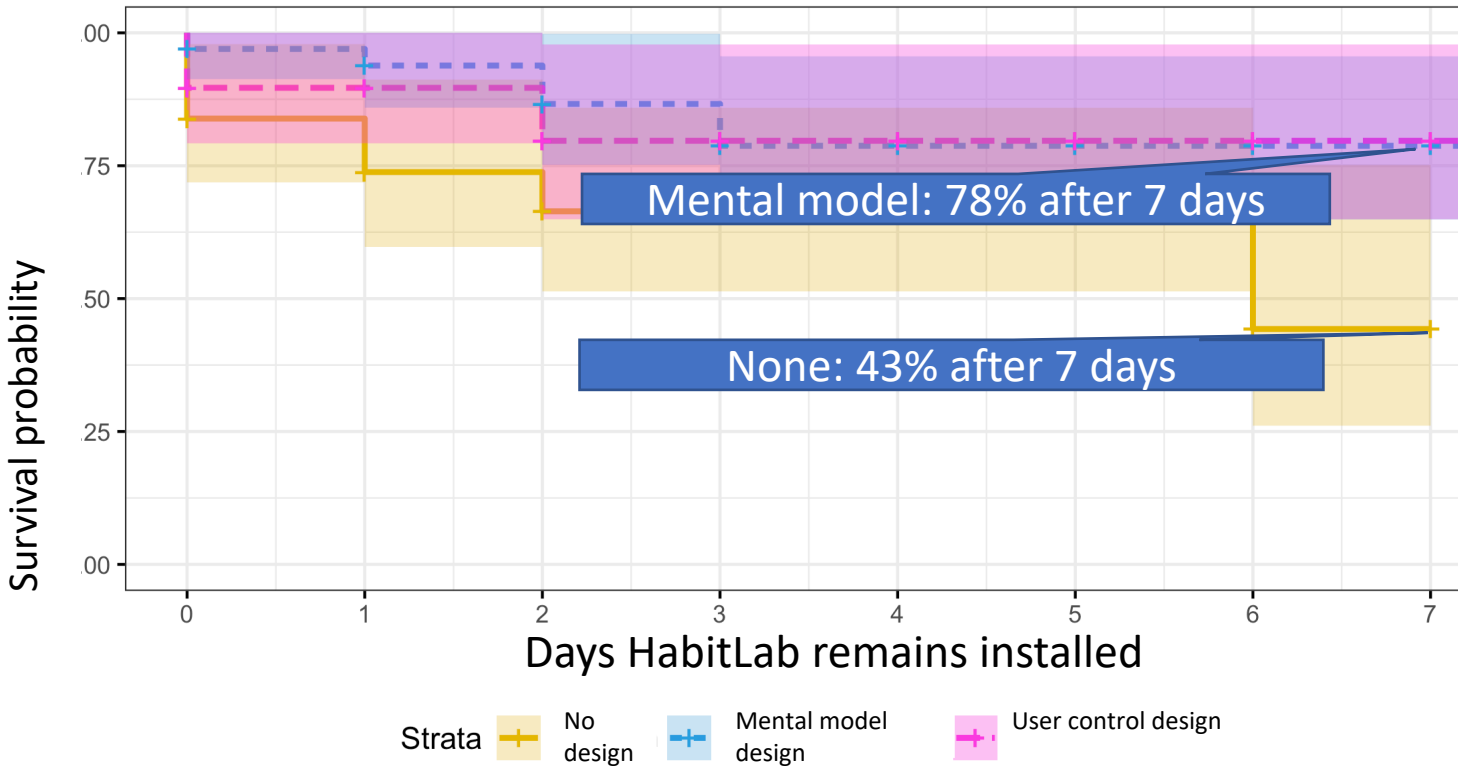
Design	Log hazard ratio
Mental model	-1.015 ($p < 0.05$)
User control	-0.869

Mental model design reduces attrition by half



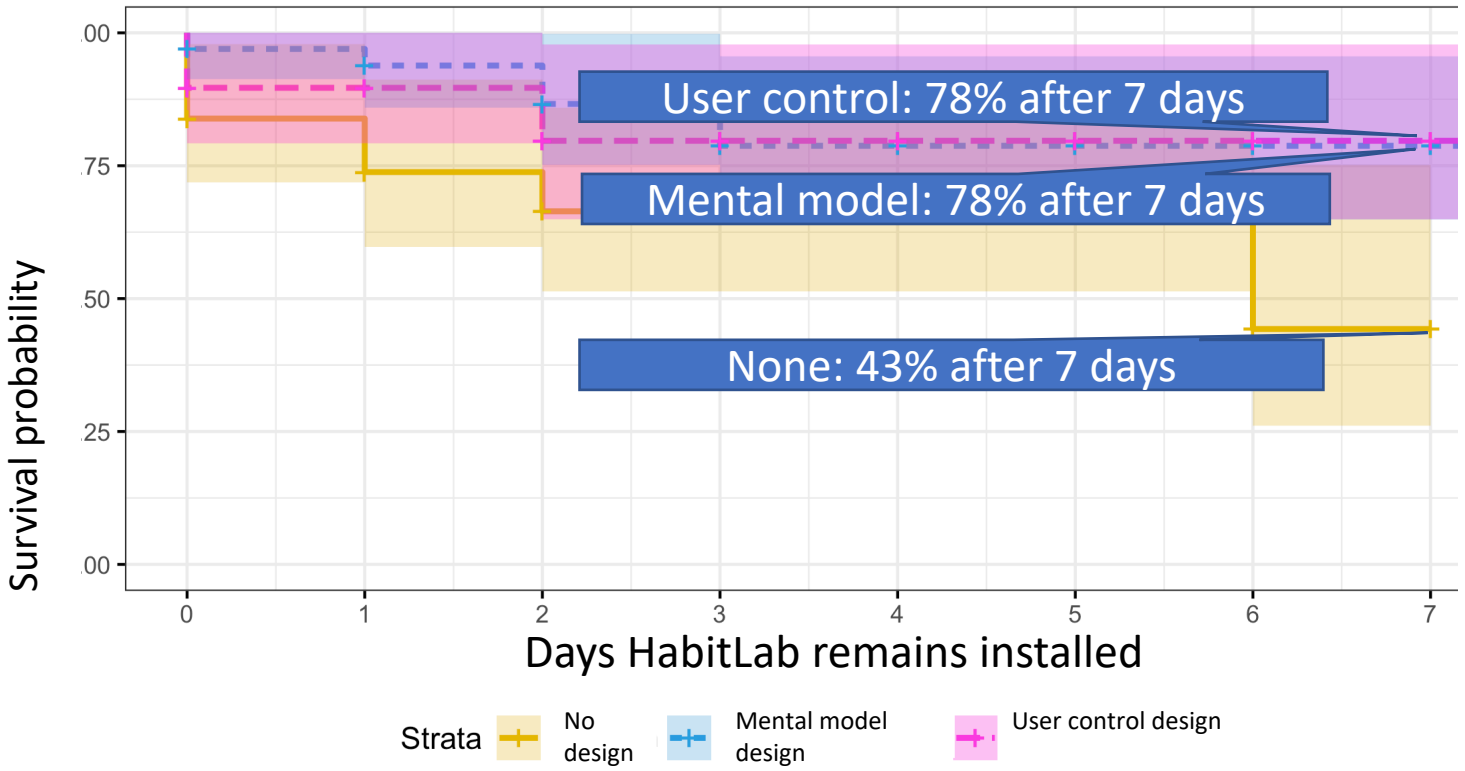
Design	Log hazard ratio
Mental model	-1.015 ($p < 0.05$)
User control	-0.869

Mental model design reduces attrition by half



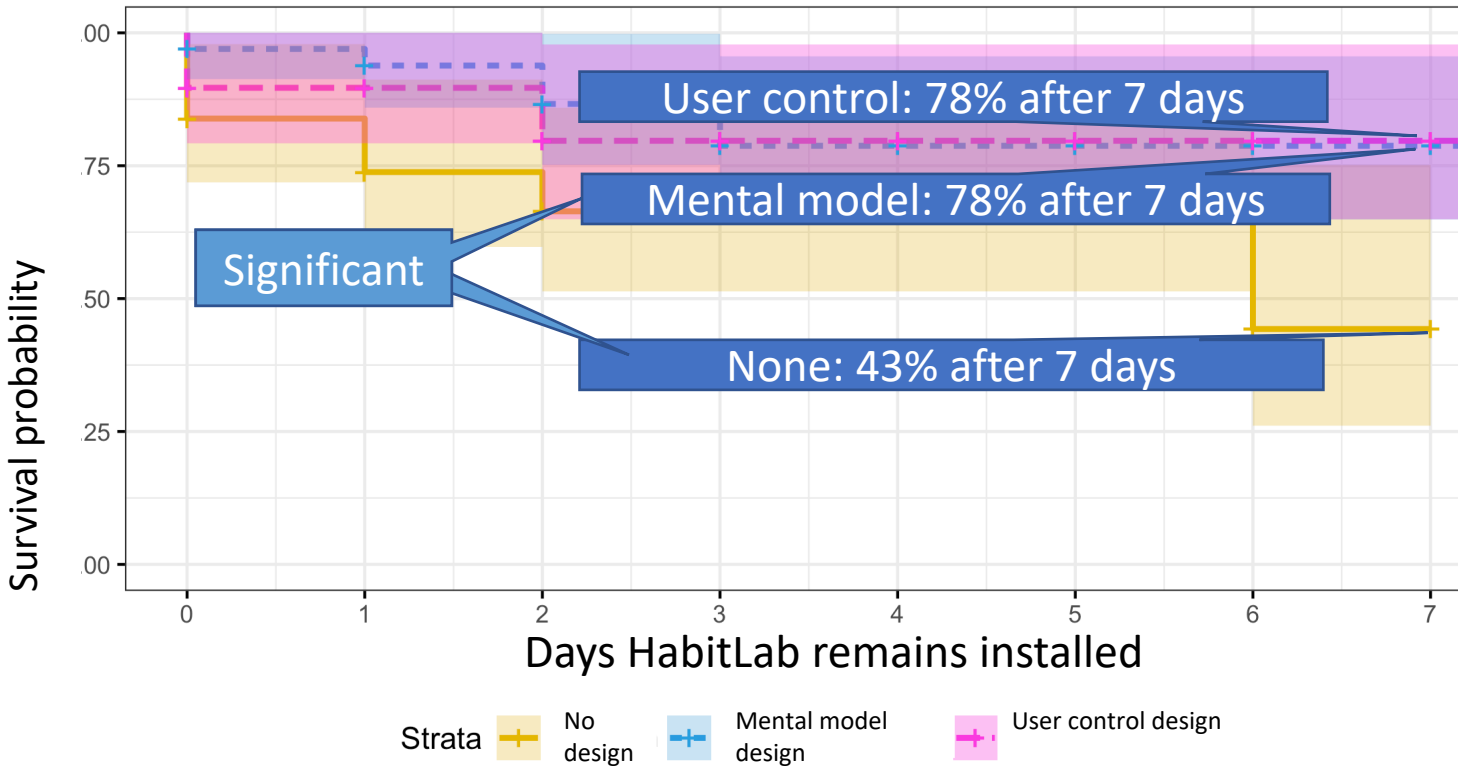
Design	Log hazard ratio
Mental model	-1.015 ($p < 0.05$)
User control	-0.869

Mental model design reduces attrition by half



Design	Log hazard ratio
Mental model	-1.015 ($p < 0.05$)
User control	-0.869

Mental model design reduces attrition by half



Design	Log hazard ratio
Mental model	-1.015 ($p < 0.05$)
User control	-0.869

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating vs static intervention strategies

Study 1: Within subjects

- Rotating interventions improves effectiveness but increases attrition

Study 2: Between subjects

- Rotation itself causes attrition

Qualitative feedback

- Rotation violates mental models

Study 3: Improving users' mental models about rotating interventions

- Our design halves attrition

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition (CSCW 2018)

- Static interventions decline in effectiveness over time
- Rotating interventions improves effectiveness but increases attrition
- Attrition may be due to incorrect mental models and lack of control
- We can reduce attrition with a simple design that improves users' mental models

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It? (CHI 2019)

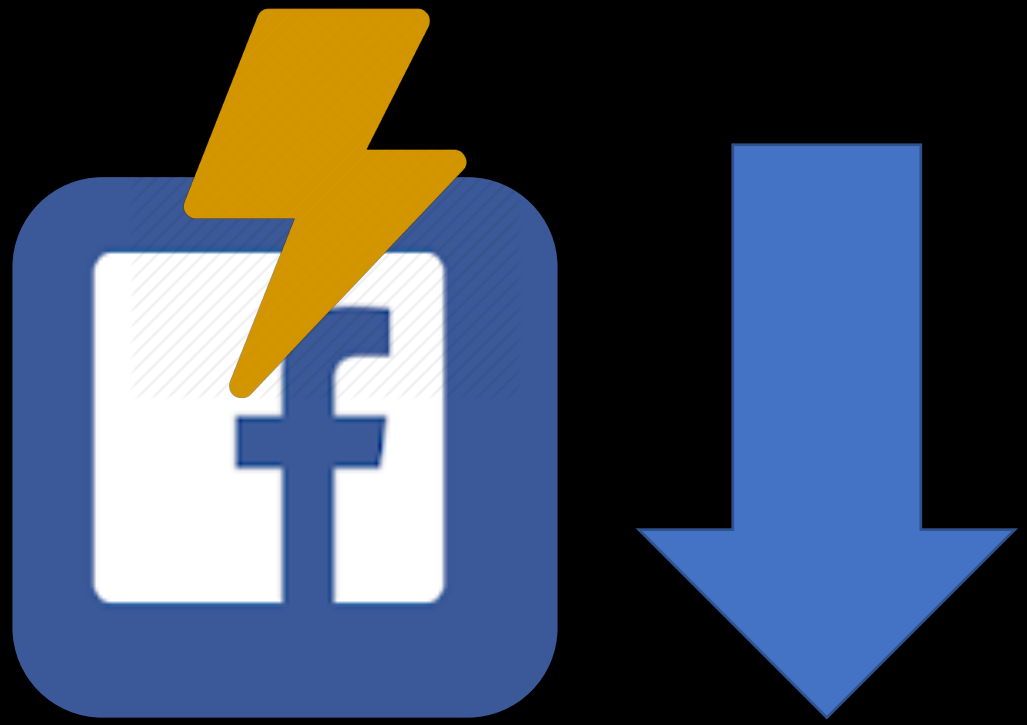
Users use productivity tools to help them reduce time online

Is that time actually saved, or just redirected to other unproductive activities?

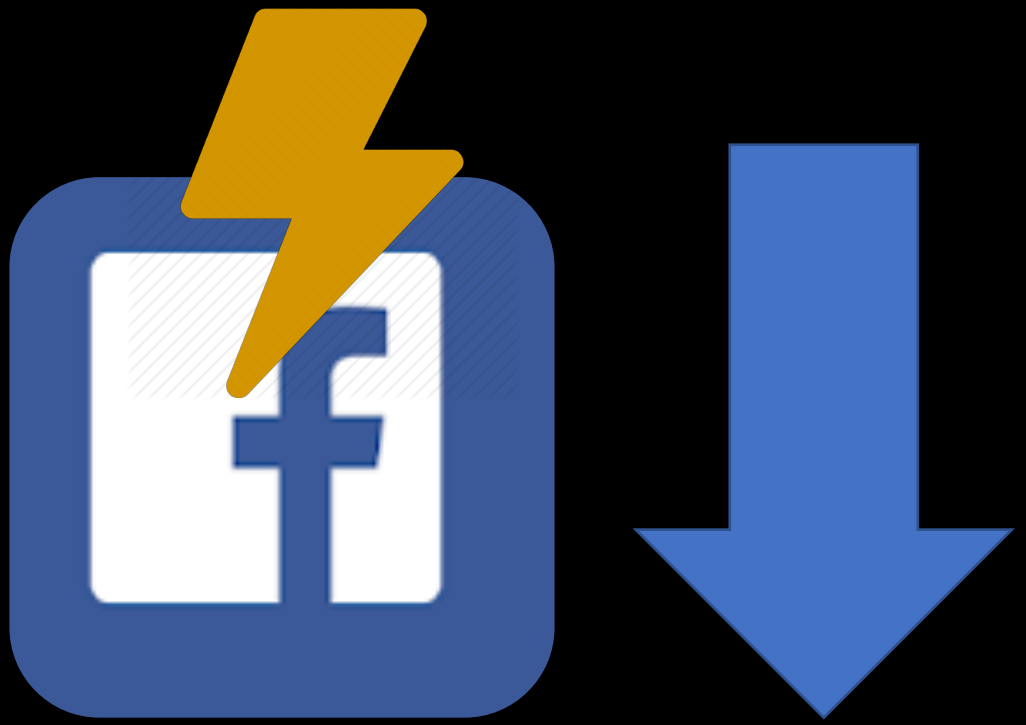
We often assume that intervention effects are isolated

We often assume that intervention effects are isolated

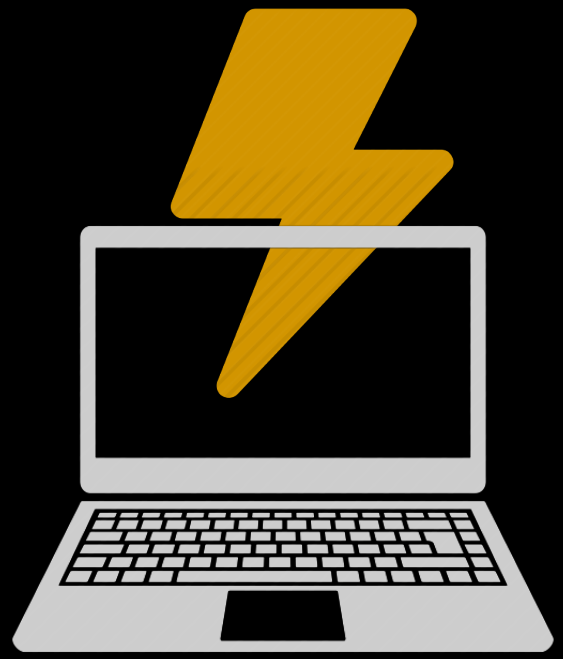
We often assume that intervention effects are isolated

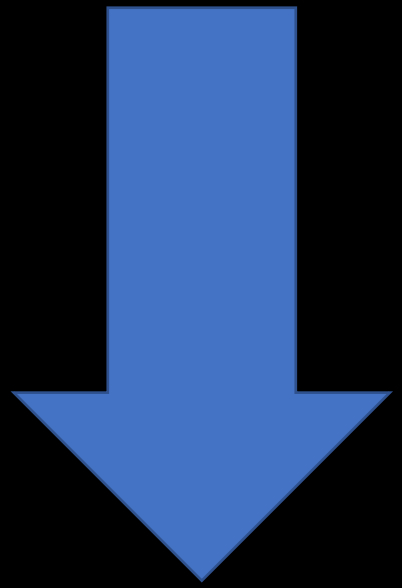


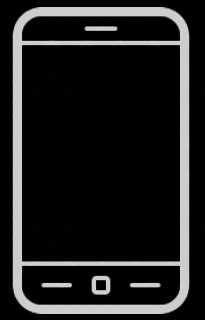
We often assume that intervention effects are isolated

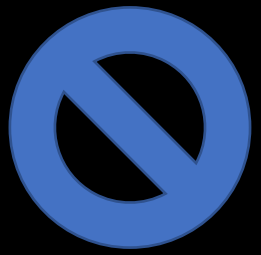


We often assume that intervention effects are isolated









What if the time you saved is just shifted elsewhere?

We have a limited supply of willpower

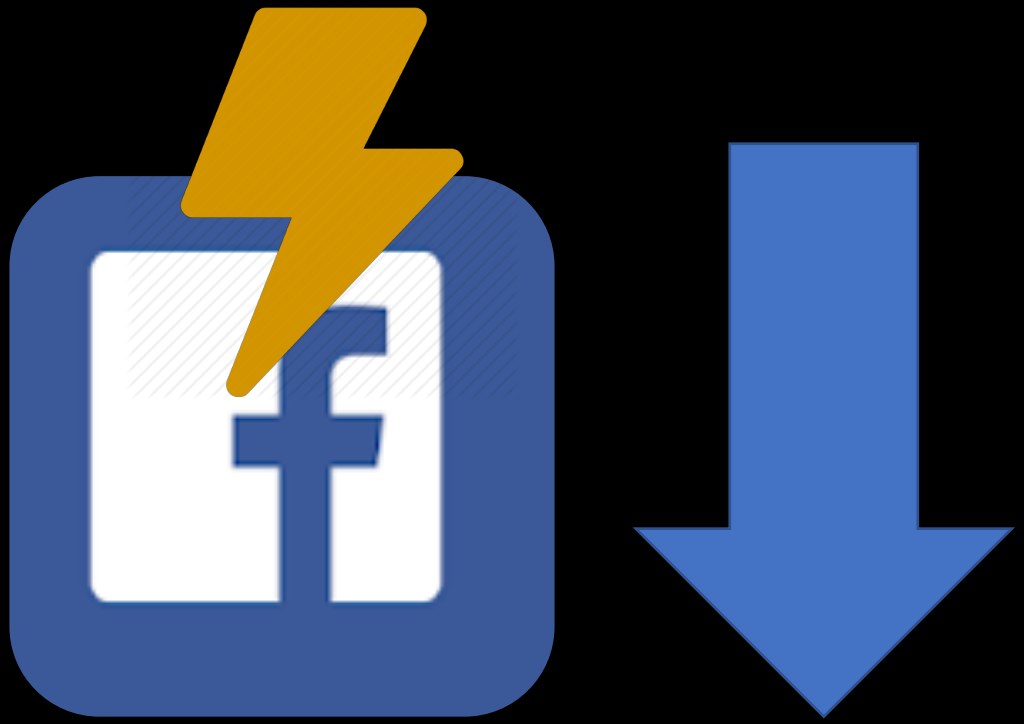
We need breaks and downtime

Laura Dabbish, Gloria Mark, and Víctor M González. 2011. Why do I keep interrupting myself?: Environment, Habit and Self-interruption. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 3127–3130.

Roy E Baumeister, Ellen Bratslavsky, Mark Muraven, and Dianne M Tice. 1998. Ego Depletion: Is the Active Self a Limited Resource? Journal of Personality and Social Psychology 74, 5 (1998), 1252–1265.

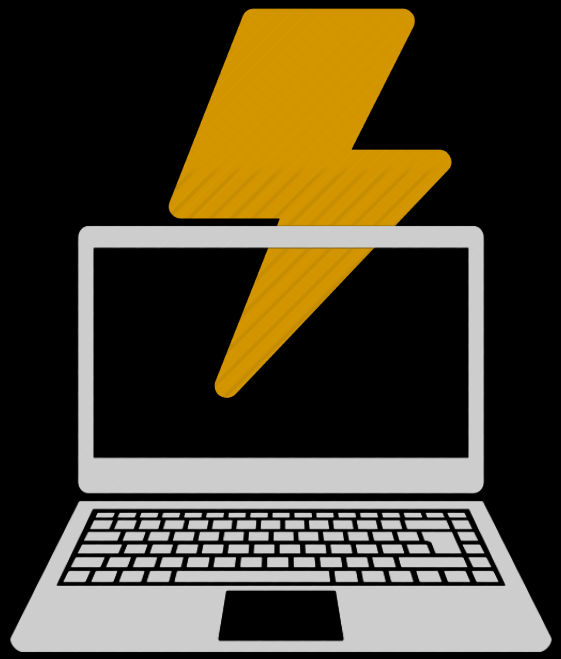
Is there a conservation-of-procrastination effect?

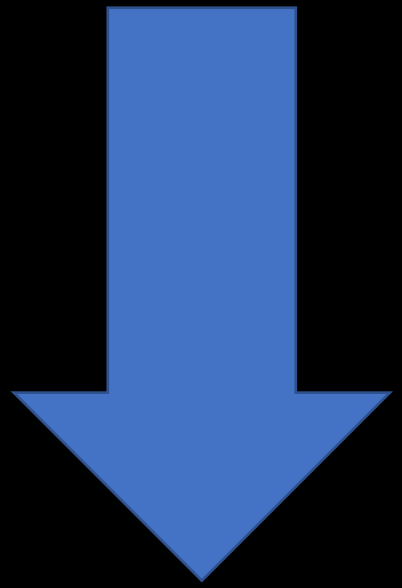
Does reducing time on one site or app increase time on others?

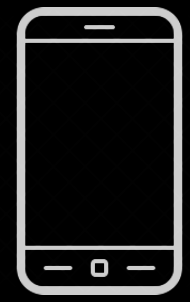


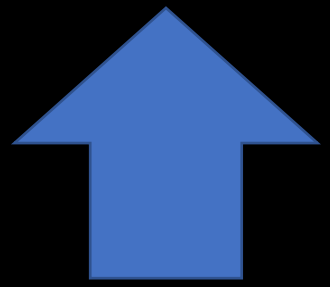
Is there a conservation-of-procrastination effect?

Does reducing time on one device increase time on others?





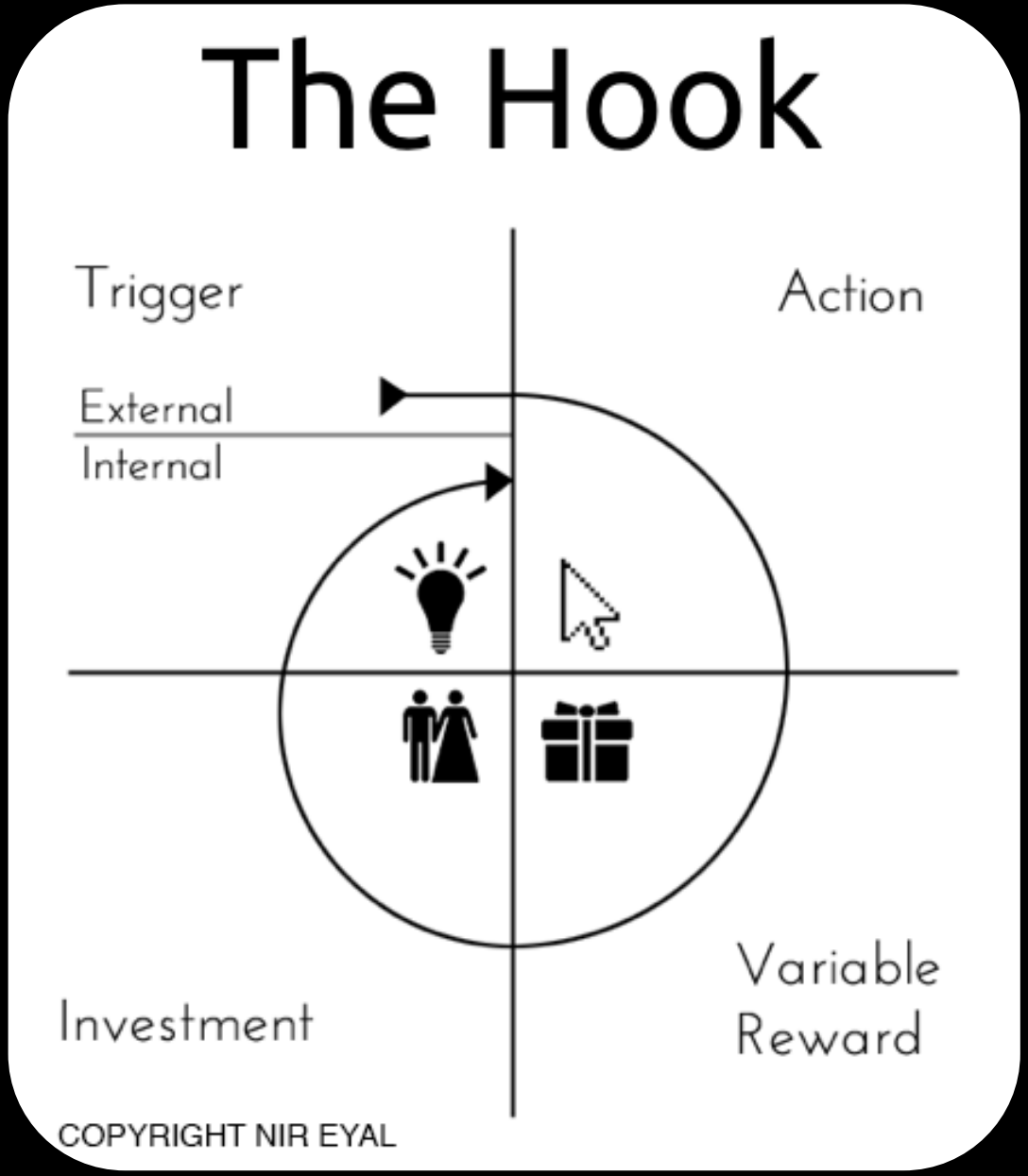




Do interventions have benefits outside the apps they were targeting?

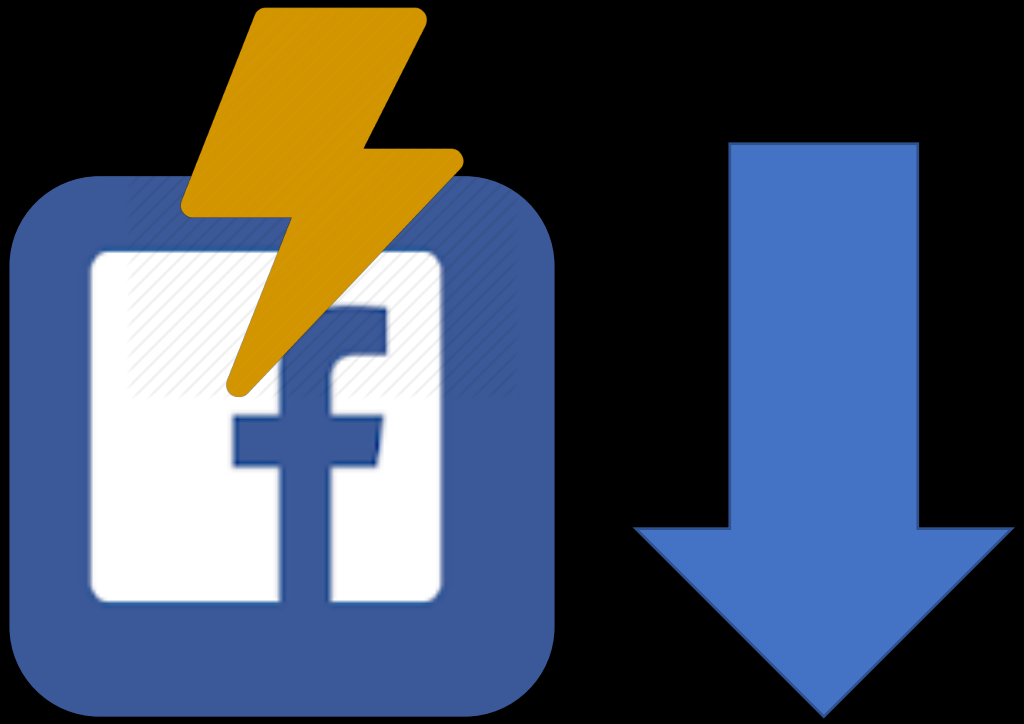
Apps are designed to be habit-building

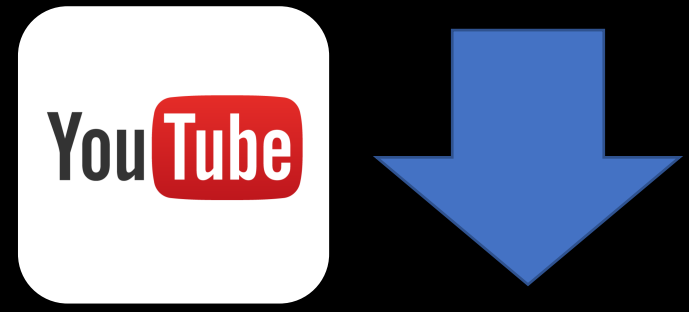
Results in habit loop of constantly visiting sites or checking phones



Does breaking habit loops result in further decreases in time spent elsewhere?

Does reducing time on one app or site decrease time on others?



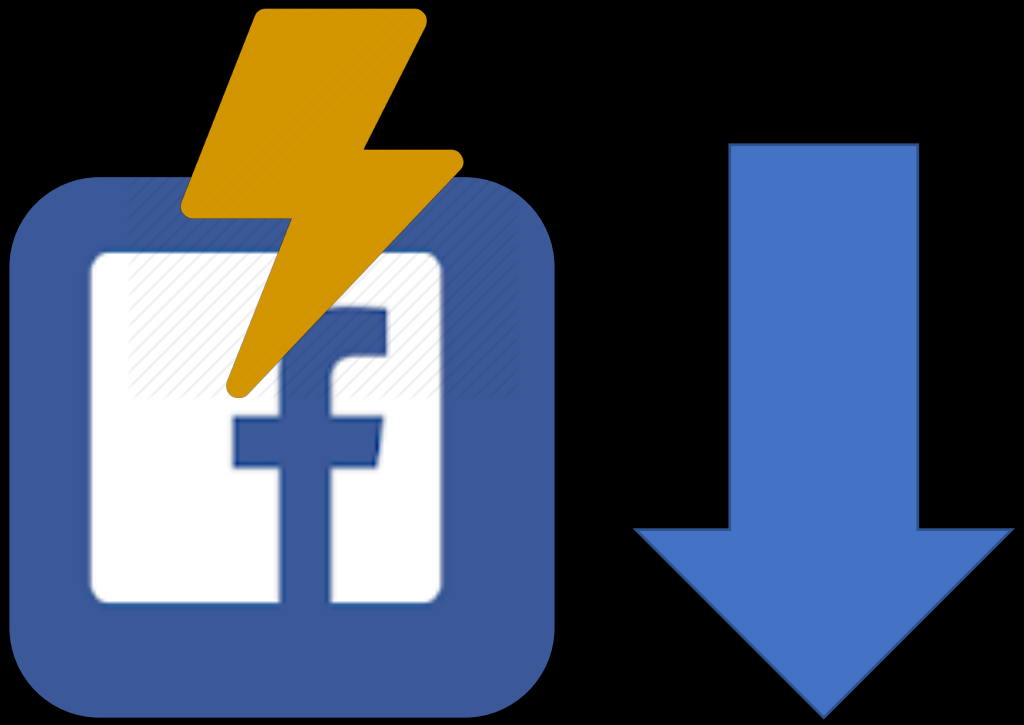


Does breaking habit loops result in further decreases in time spent elsewhere?

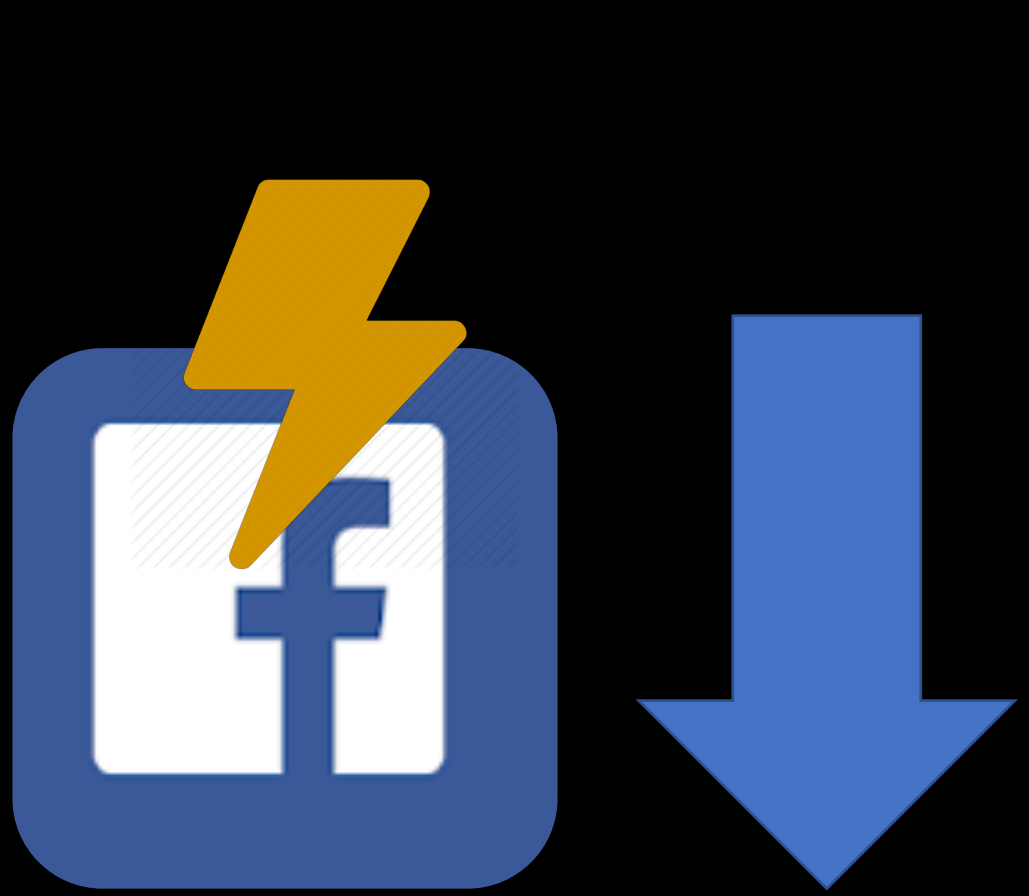
Does reducing time on one device decrease time on the other?

RQ1: Do interventions on one site or app influence time spent on other sites and apps?

RQ1: Do interventions on one site or app influence time spent on other sites and apps?



RQ1: Do interventions on one site or app influence time spent on other sites and apps?

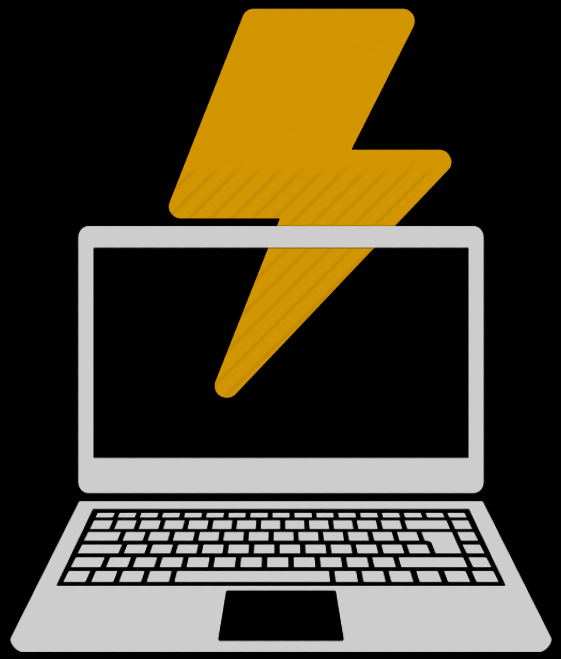


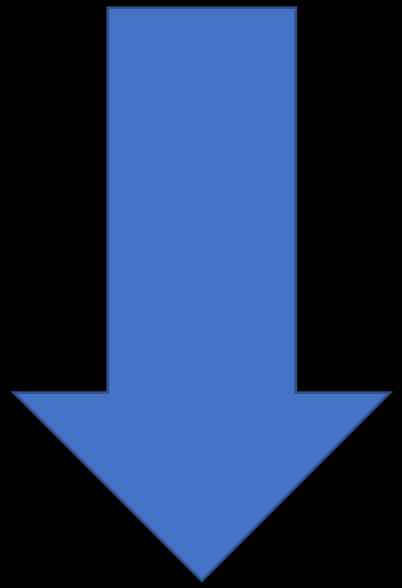
Isolation

Redistribution

Reduction

RQ2: Do interventions on one device influence time spent on other devices?

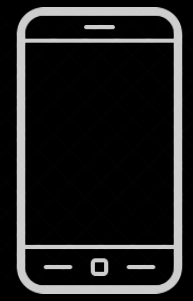


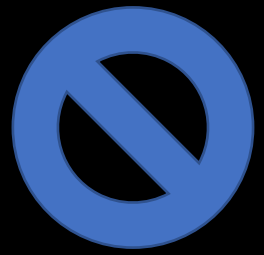


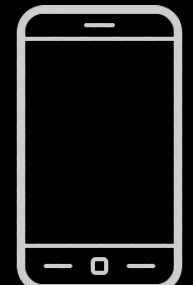
Isolation

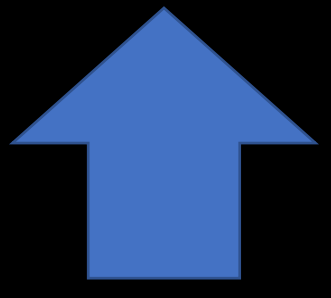
Redistribution

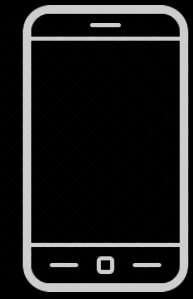
Reduction

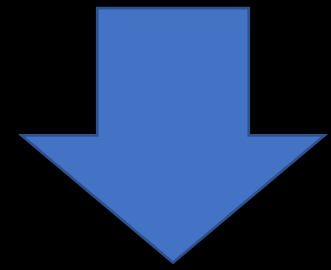












HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It? (CHI 2019)

Are interventions effective at reducing time on the focal goal?

RQ1: Is time redistributed to other sites on the same platform?

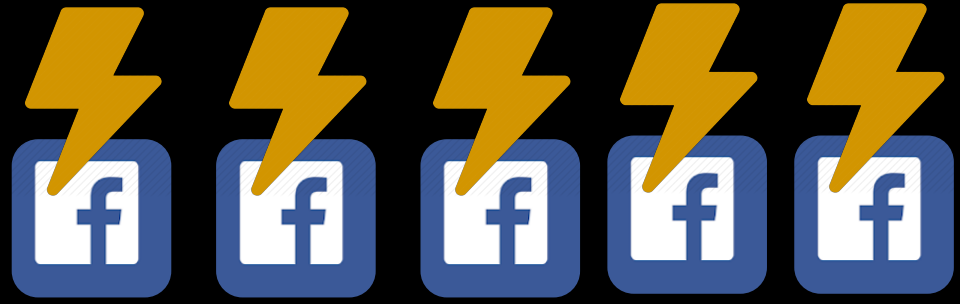
RQ2: Is time redistributed across devices?

Are interventions effective? (Method)

For each goal, we randomly assign it to one of 2 conditions each week:

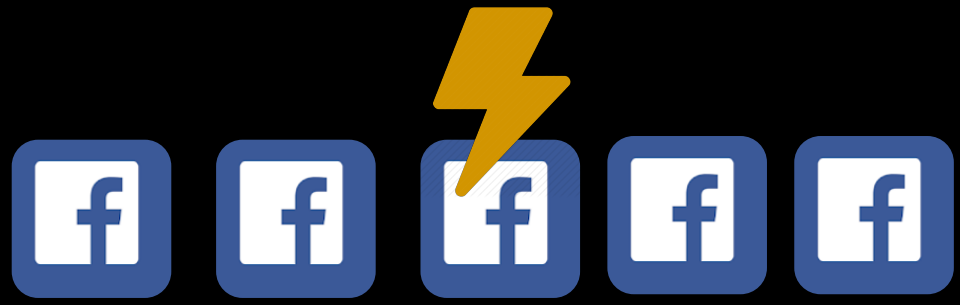
Frequent

An intervention is shown every visit
(each site visit for browser, each app visit on android)



Infrequent

An intervention is shown on 20% of visits



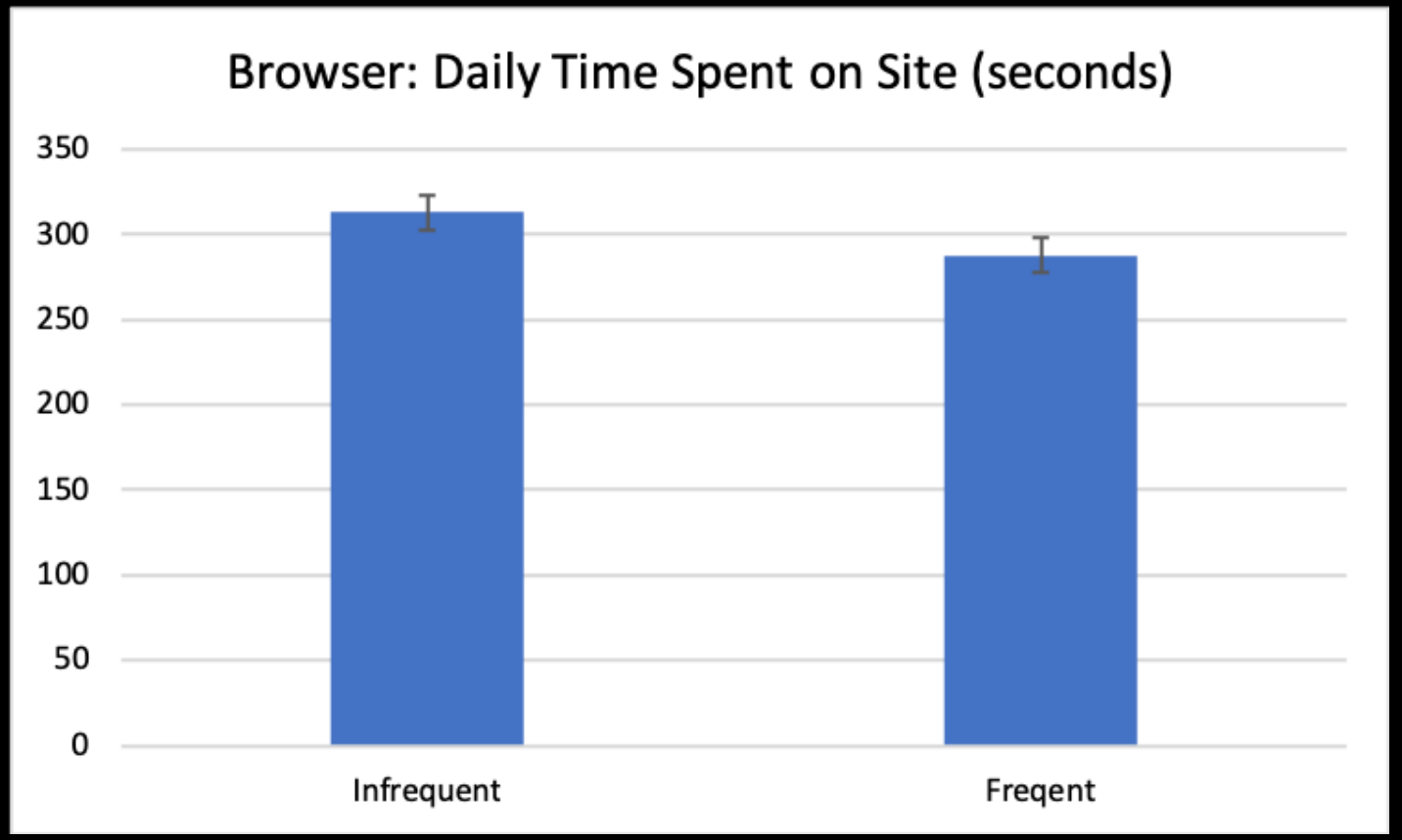
Are interventions effective? (Method)

- Compare daily time spent on days in the frequent vs infrequent conditions, for each goal

Are interventions effective? (Method)

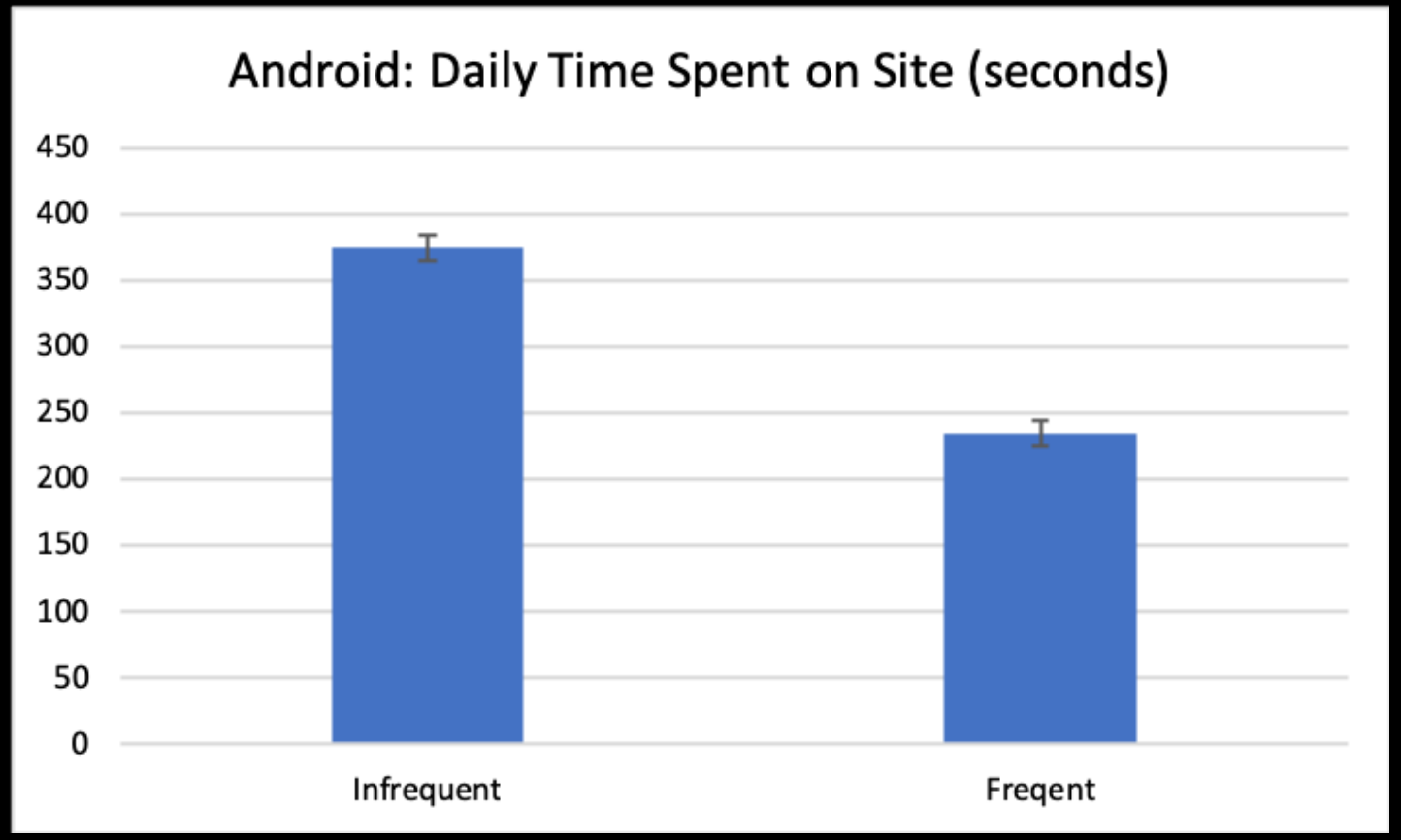
- Compare daily time spent on days in the frequent vs infrequent conditions, for each goal
- 5.8 weeks with 1034 users on browser (n=22,462 days), and 876 users on mobile (n=26,273 days)

Interventions are effective on both platforms



7.3% reduction in daily time spent on browser version, on frequent weeks. Statistically significant ($p < 0.001$)

Interventions are effective on both platforms



37.2% reduction in daily time spent on android version, on frequent weeks. Statistically significant ($p < 0.001$)

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Are interventions effective at reducing time on the focal goal?

- Effective on both browser + mobile

RQ1: Is time redistributed to other sites on the same platform?

RQ2: Is time redistributed across devices?

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Are interventions effective at reducing time on the focal goal?

- Effective on both browser + mobile

RQ1: Is time redistributed to other sites on the same platform?

RQ2: Is time redistributed across devices?

Is time redistributed within platform? (Method)

We developed a metric of how intense interventions are this day (intensity)

- Percentage of sessions on a goal that triggered an intervention

Is time redistributed within platform? (Method)

We developed a metric of how intense interventions are this day (intensity)

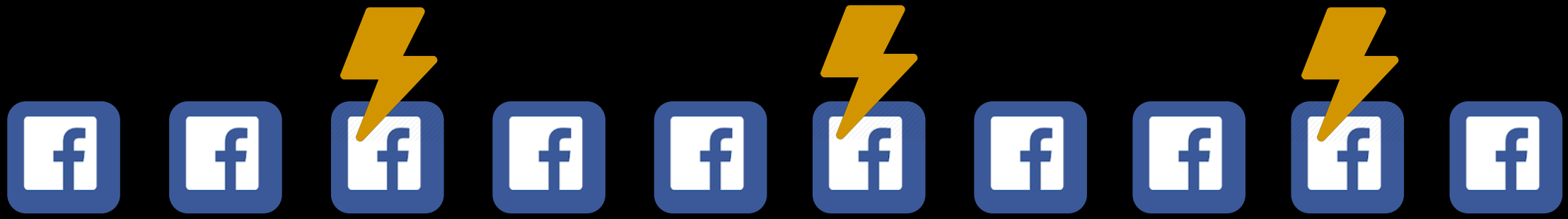
- Percentage of sessions on a goal that triggered an intervention.
- E.g. if visited Facebook 10x, and saw 0 interventions, then intensity=0

Intensity=0

Is time redistributed within platform? (Method)

We developed a metric of how intense interventions are this day (intensity)

- Percentage of sessions on a goal that triggered an intervention.
- E.g. if visited Facebook 10x, and saw 3 interventions, then intensity=0.3

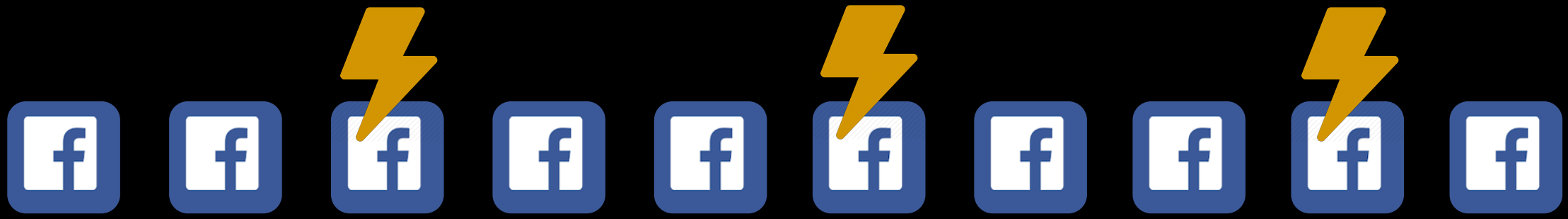


Intensity=0.3

Is time redistributed within platform? (Method)

We developed a metric of how intense interventions are this day (intensity)

- Percentage of sessions on a goal that triggered an intervention.
- Verified that on days where intensity is higher, overall total time on goal sites is significantly lower on both platforms



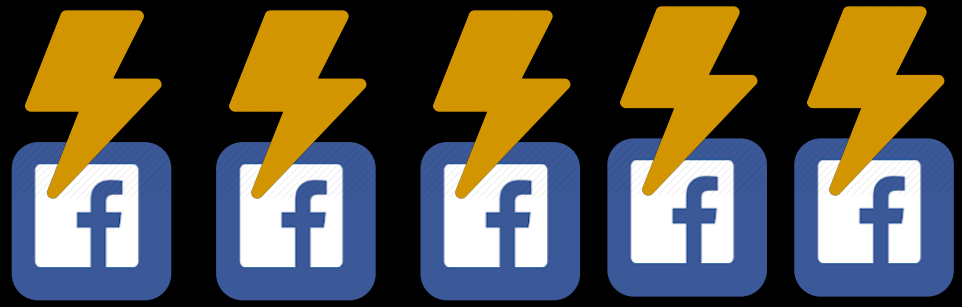
Intensity=0.3

Is time redistributed within platform? (Method)

To manipulate intensity, we randomly assign each goal to have either frequent or infrequent interventions each week, resulting in a continuous intensity value from 0 to 1

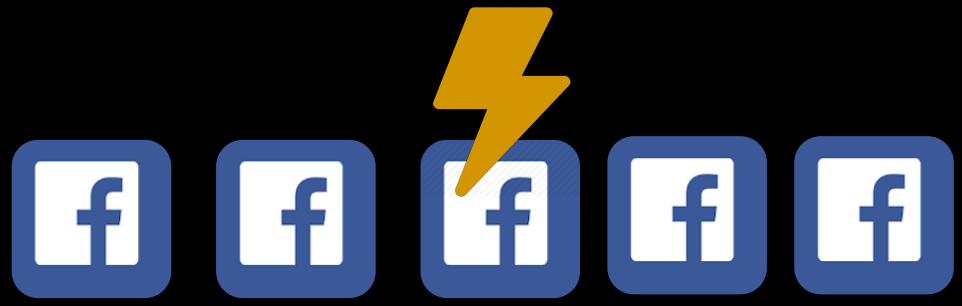
Frequent

An intervention is shown every visit
(each site visit for browser, each app visit on android)



Infrequent

An intervention is shown on 20% of visits

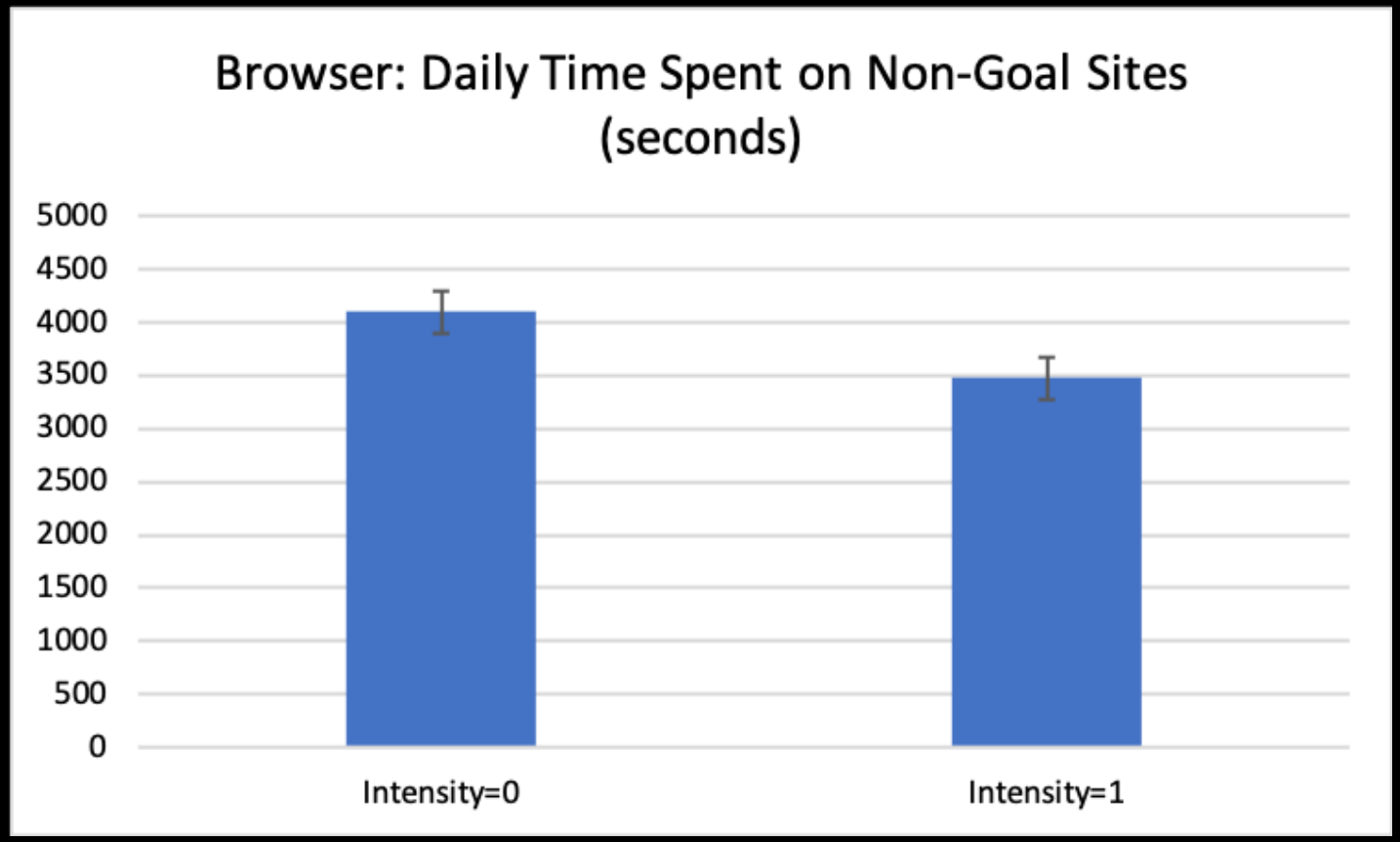


Is time redistributed within platform? (Method)

On days when intensity is higher, what is the effect on the time spent on non-goal apps and sites?

LMM **(Linear Mixed Model) structure**	
Dependent variable	Total time on non-goal sites (log)
Fixed effects	Intensity
Random effects	User

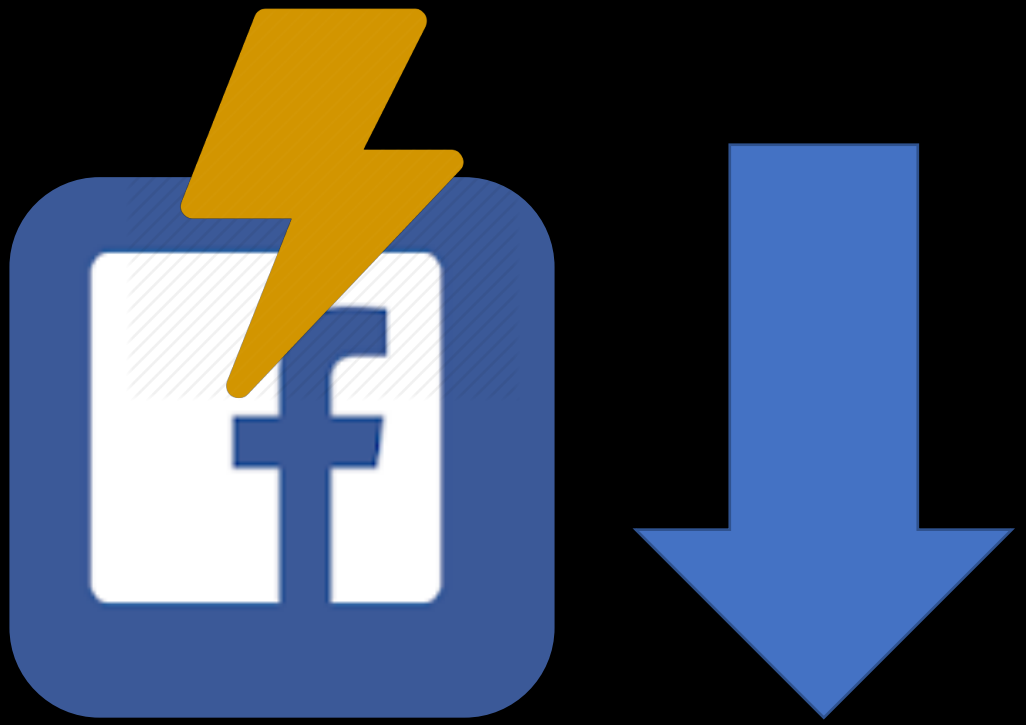
Browser: reduction of time spent on other sites when intensity is higher



15% reduction in time spent when intensity increases from 0→1 ($p < 0.0001$)

Mobile: No significant effect of time on one app on other apps

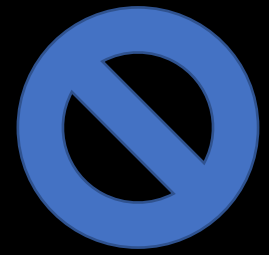
RQ1: Do interventions on one site/app influence time spent on other sites/apps?

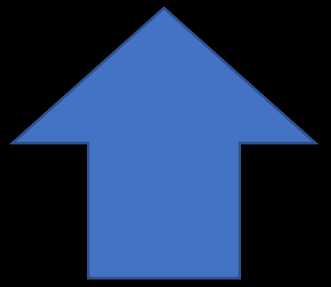


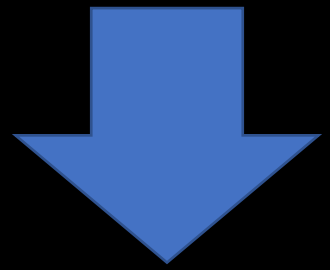
Isolation

Conservation

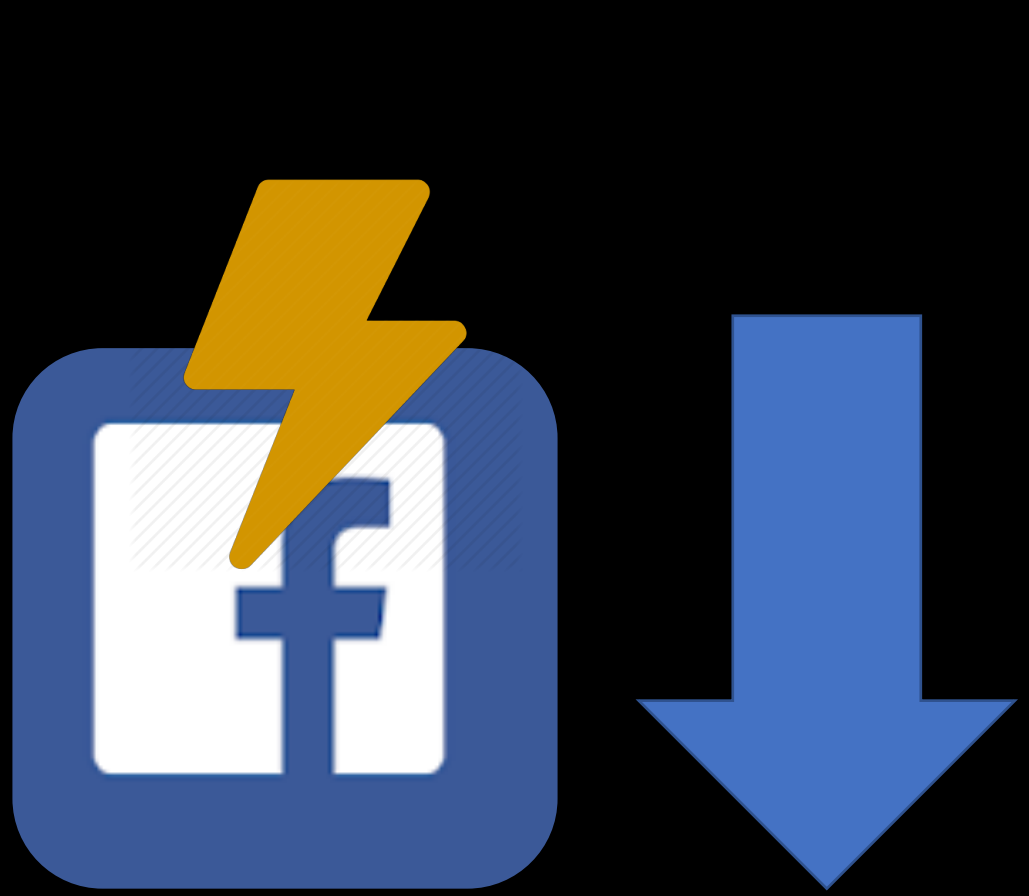
Reduction







RQ1: Do interventions on one site/app influence time spent on other sites/apps?

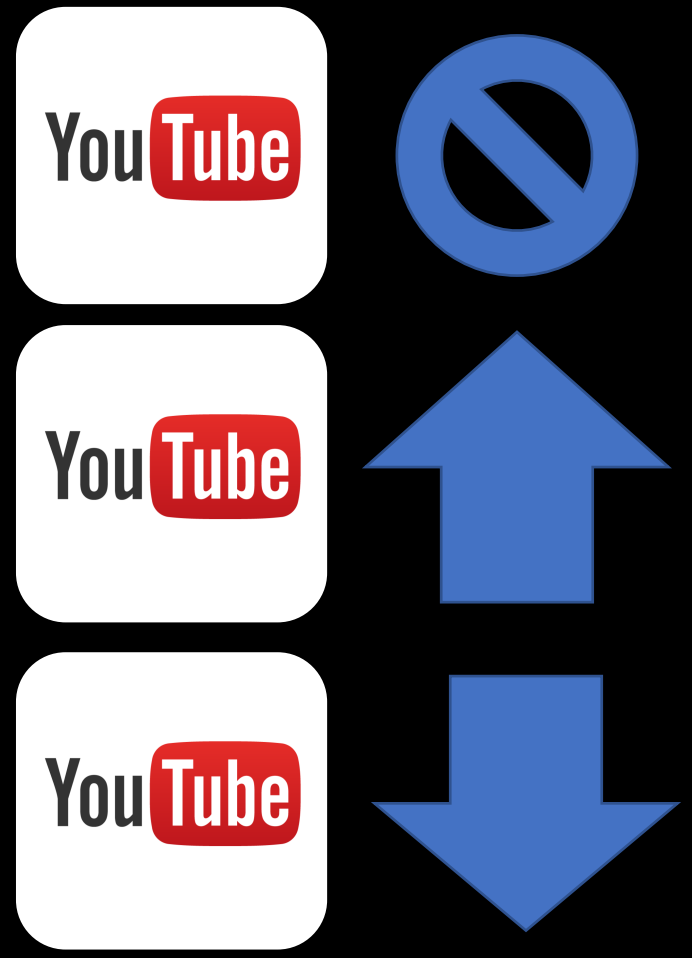


Isolation

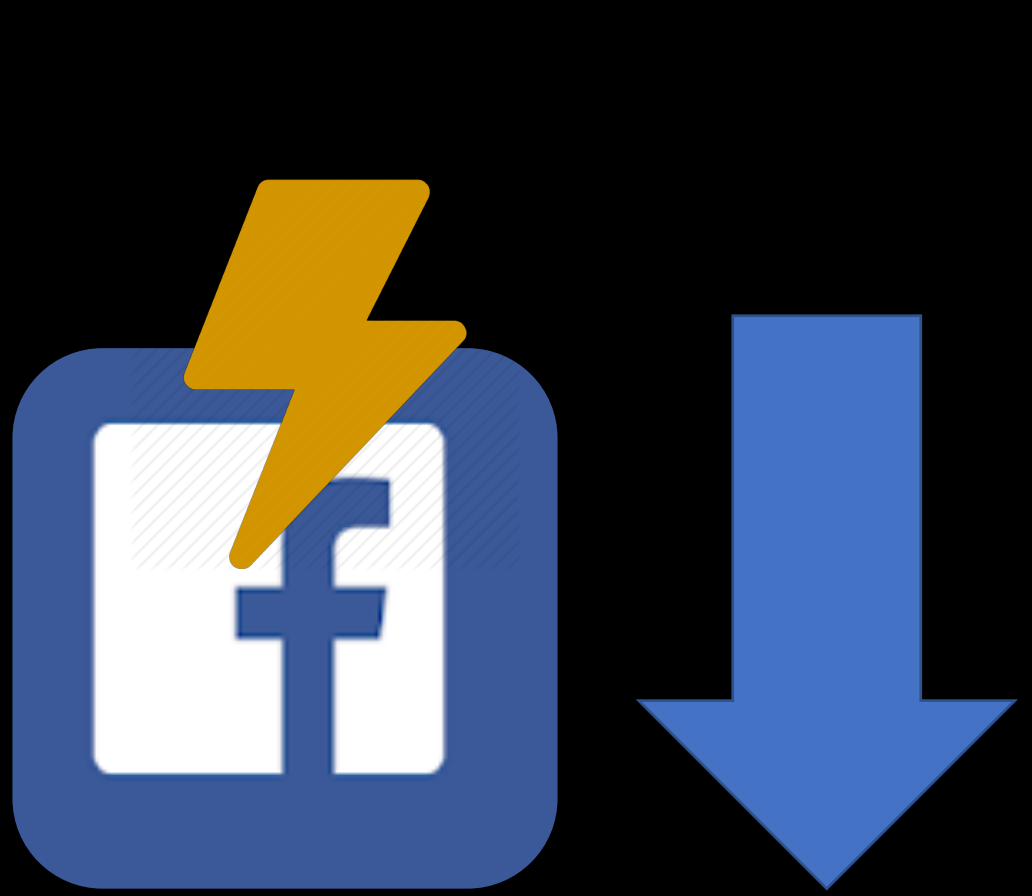
Conservation

Reduction

Browser



RQ1: Do interventions on one site/app influence time spent on other sites/apps?



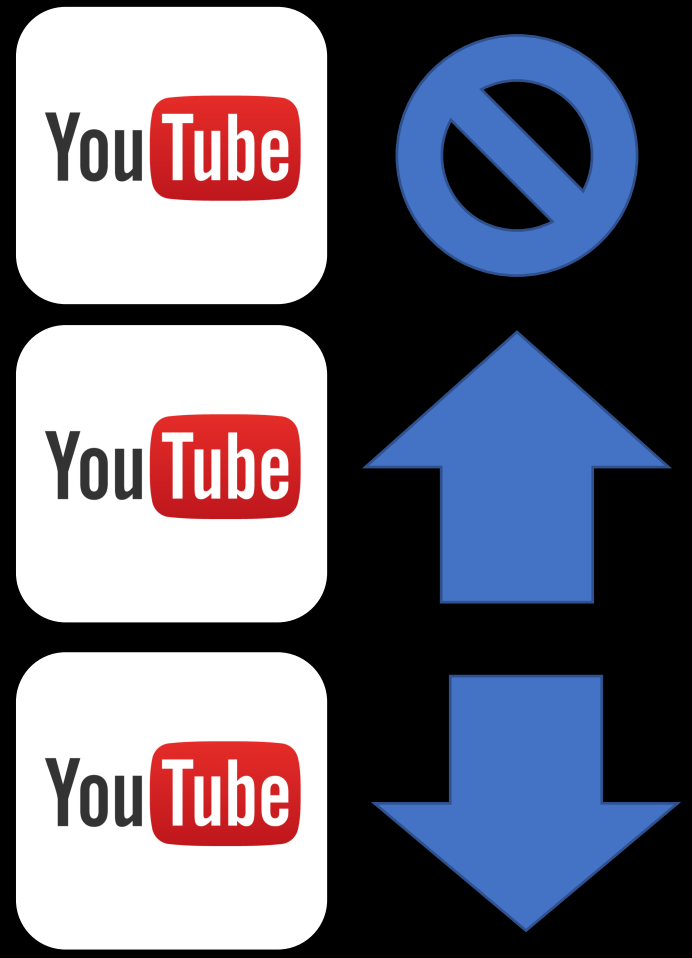
Isolation

Mobile

Conservation

Reduction

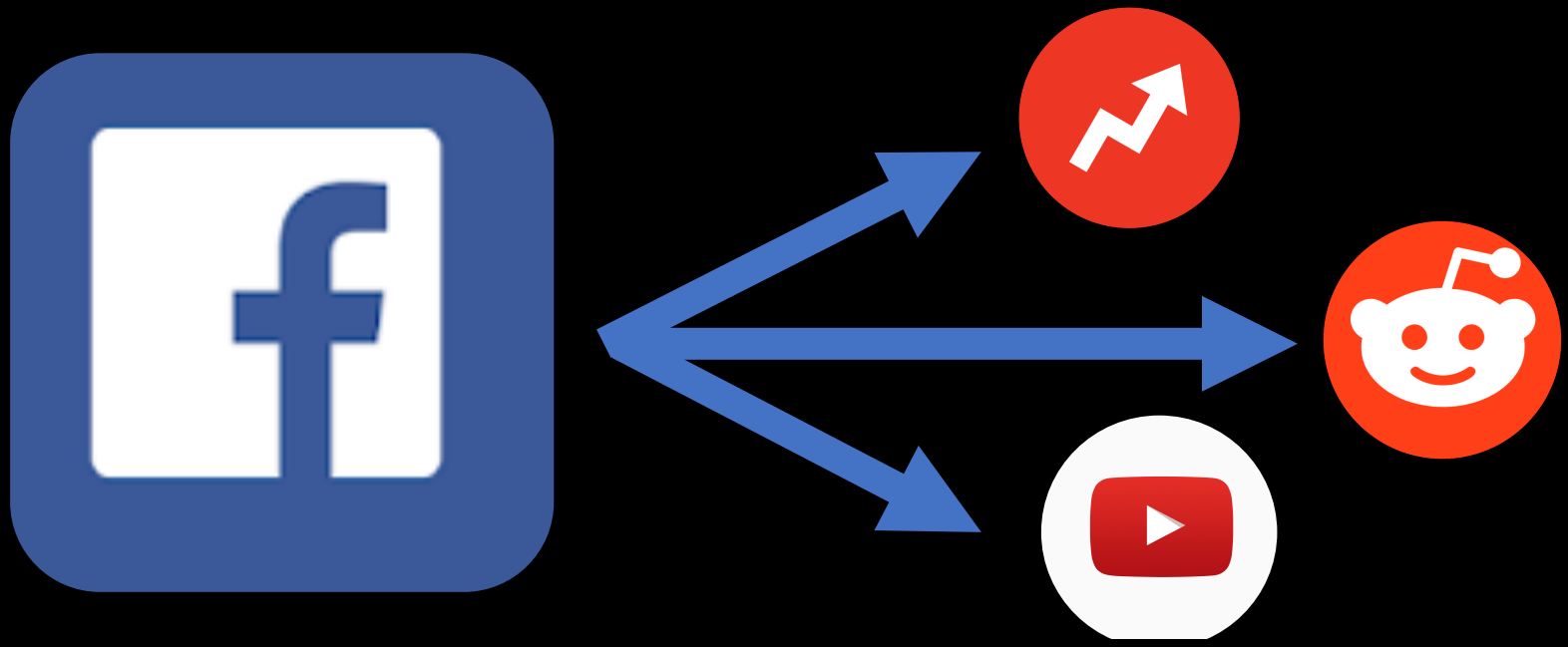
Browser



Why was there a reduction effect on browser?

Aggregator sites such as Facebook often link to other domains

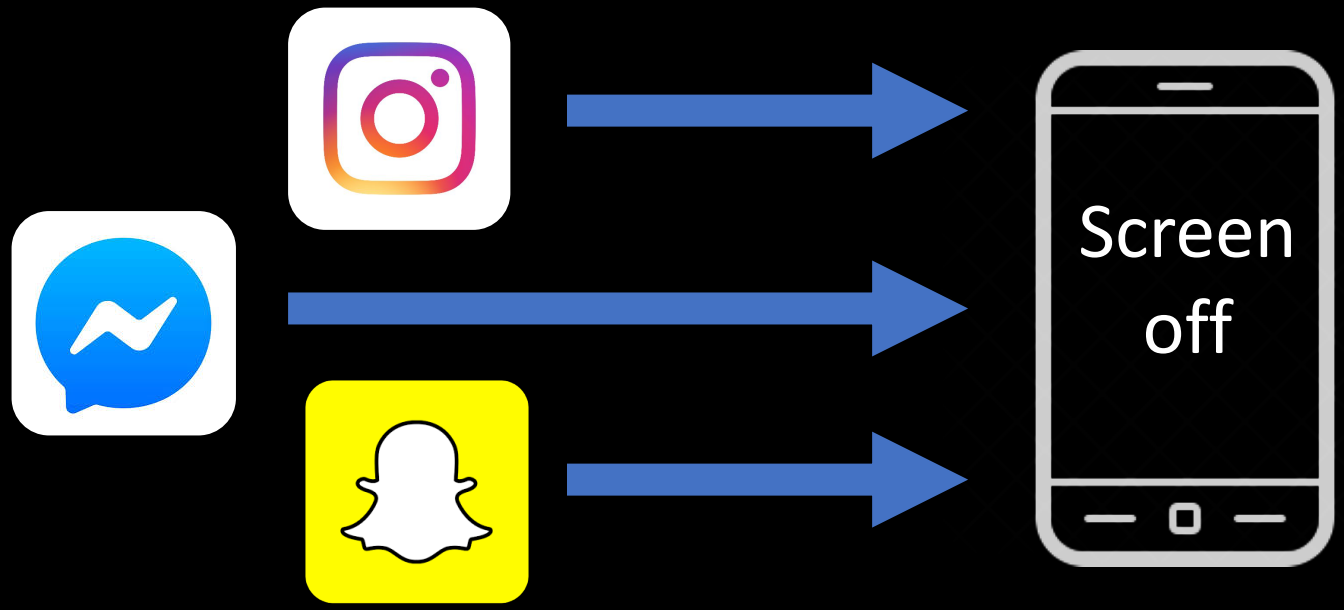
By reducing visits and time on Facebook, we reduce time on other domains



Why was there no effect on mobile?

Mobile goal apps were mostly messaging-oriented, not aggregators

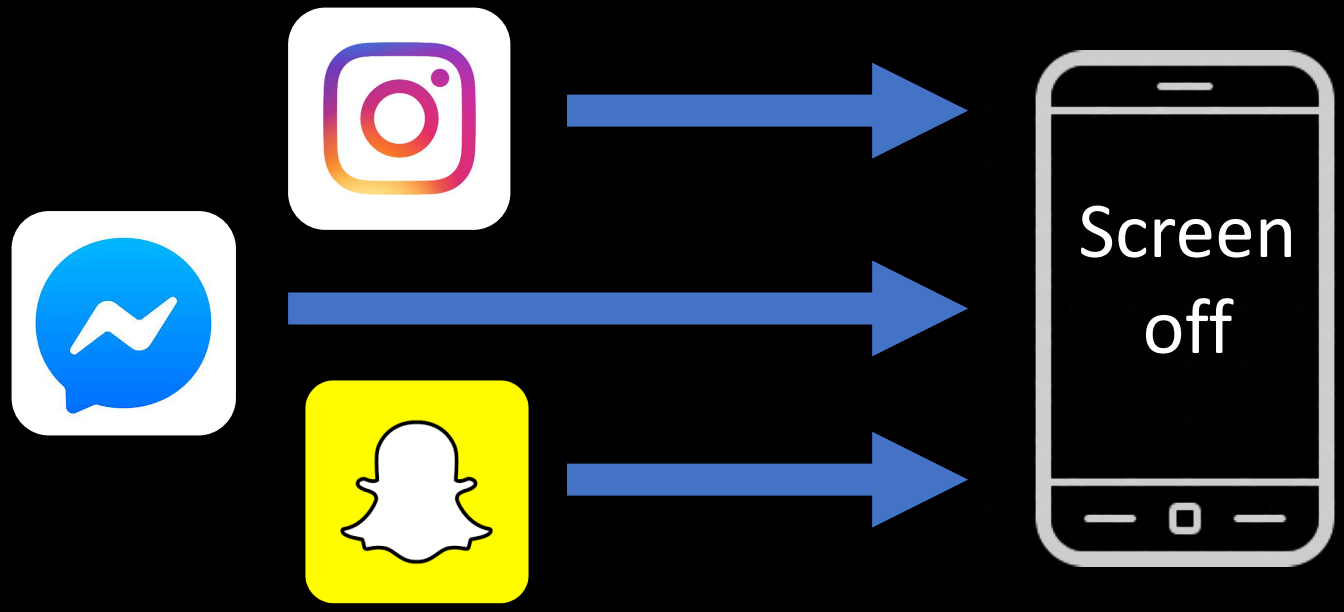
Sessions were short and followed by turning off the screen



Oulasvirta, Antti, et al. "Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2005.

Why was there no effect on mobile?

Many mobile apps such as Facebook embed an in-app browser, so visiting external links remains within the same app



HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Are interventions effective at reducing time on the focal goal?

- Effective on both browser + mobile

RQ1: Is time redistributed to other sites on the same platform?

- Reducing time on one site reduces time elsewhere on browser (reduction) but not mobile (isolation)

RQ2: Is time redistributed across devices?

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Are interventions effective at reducing time on the focal goal?

- Effective on both browser + mobile

RQ1: Is time redistributed to other sites on the same platform?

- Reducing time on one site reduces time elsewhere on browser (reduction) but not mobile (isolation)

RQ2: Is time redistributed across devices?

Is time redistributed across devices? (Method)

- On days when intensity is higher on one device, what is the effect on total time spent on goal sites on the other device?

Is time redistributed across devices? (Method)

- On days when intensity is higher on one device, what is the effect on total time spent on goal sites on the other device?
- Limited to participants using HabitLab on both platforms (60 participants, n=429 days)

Is time redistributed across devices? (Method)

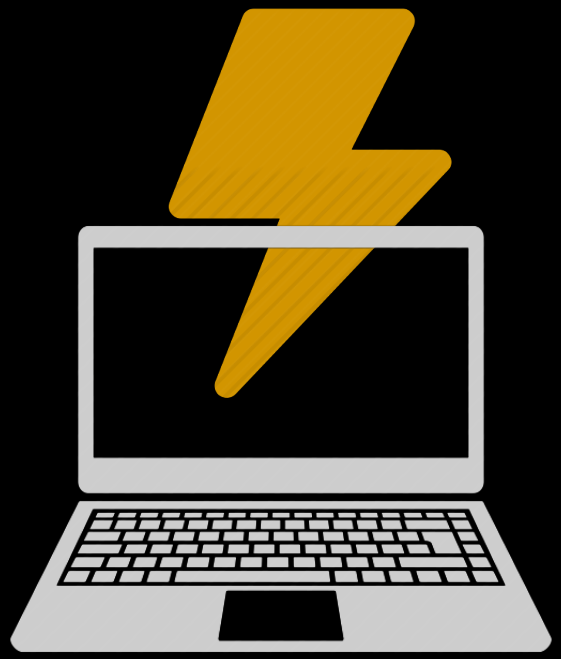
- On days when intensity is higher on one device, what is the effect on total time spent on goal sites on the other device?
- Limited to participants using HabitLab on both platforms (60 participants, n=429 days)

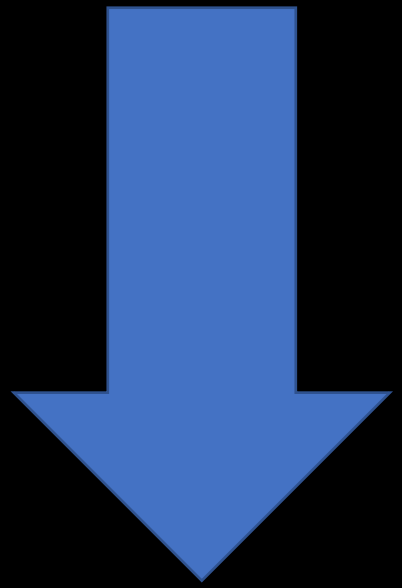
LMM **(Linear Mixed Model) structure**	
Dependent variable	Total time on other device (log)
Fixed effects	Intensity on this device
Random effects	User

Time is not redistributed across devices

- Effects of browser intensity on mobile: No significant effect ($p>.5$)
- Effects of mobile intensity on browser: No significant effect ($p>.5$)

RQ2: Do interventions on one device influence time spent on other devices?

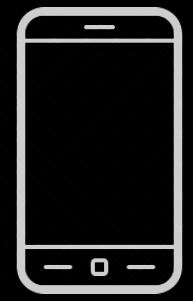


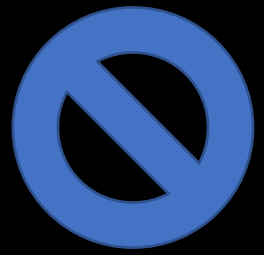


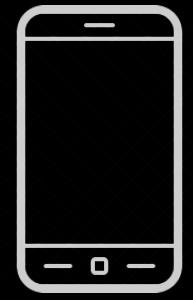
Isolation

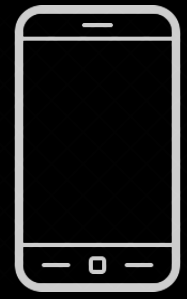
Redistribution

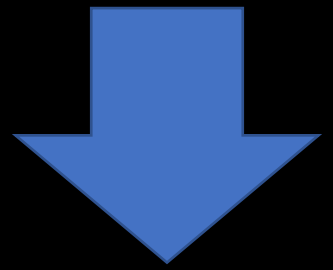
Reduction



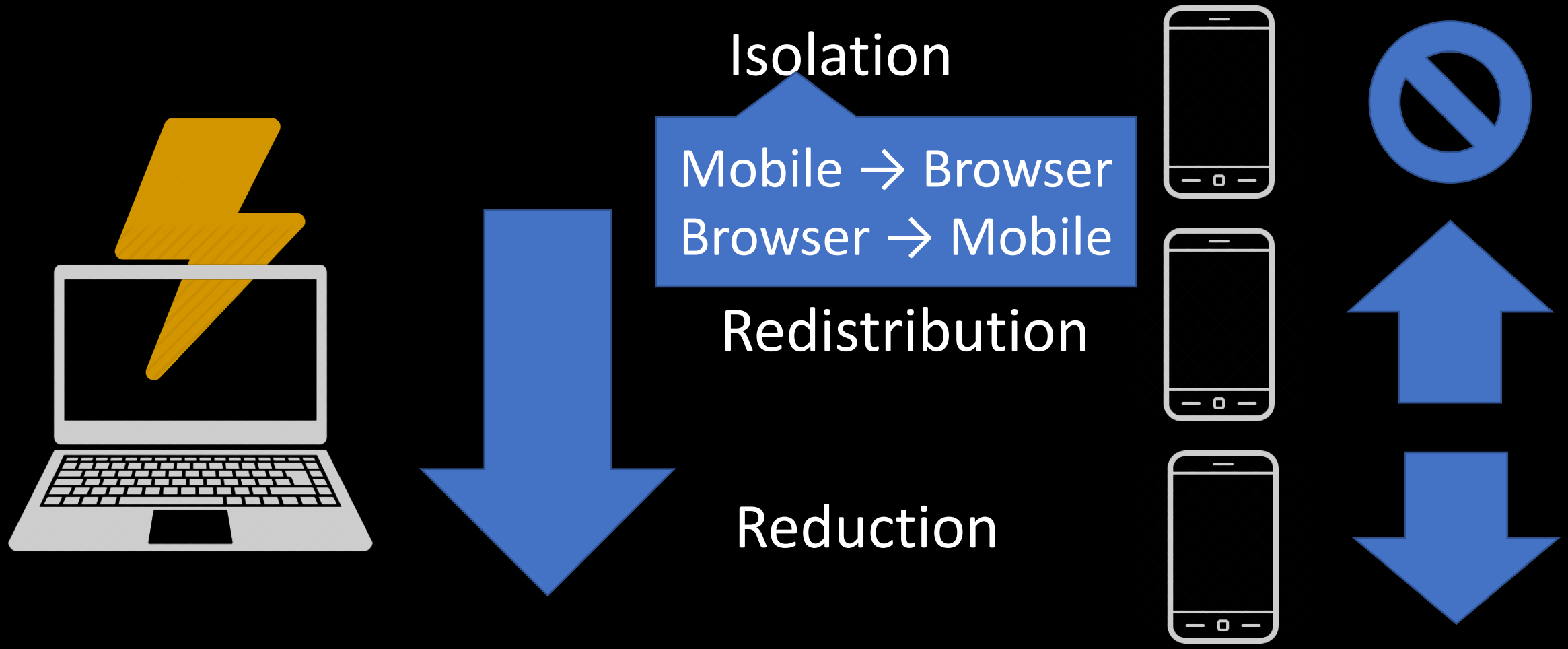






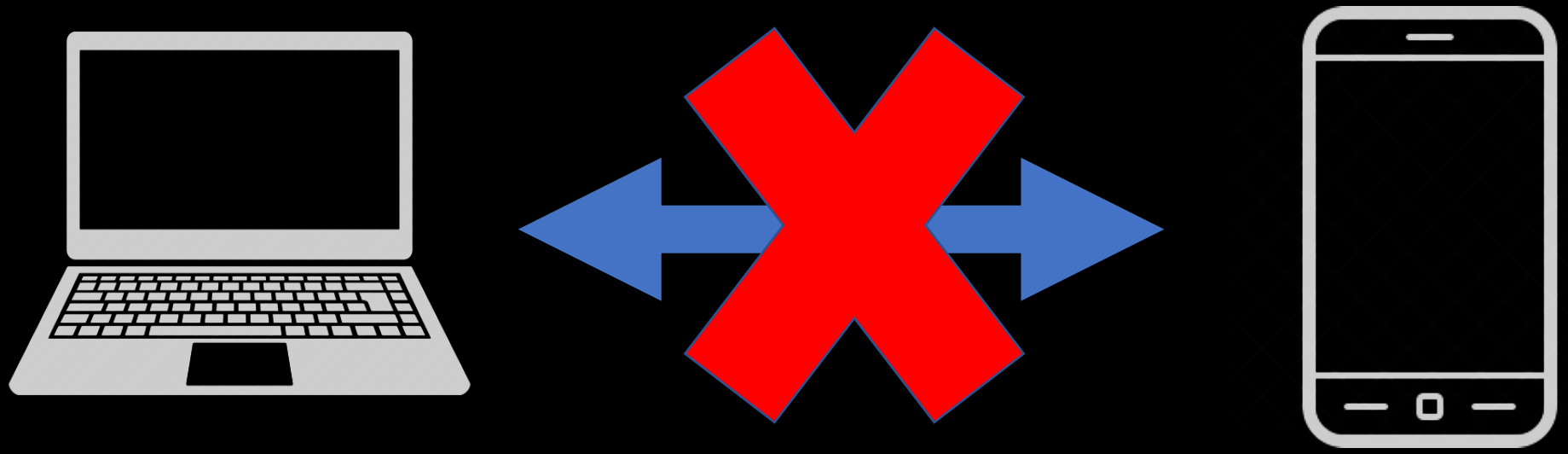


RQ2: Do interventions on one device influence time spent on other devices?



Why were there no cross-device effects?

- Laptops and phones are used in different contexts
- Unlike browsers, there are no cross-device “links” – few apps/sites prompt you to start using the other device



HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Are interventions effective at reducing time on the focal goal?

- Effective on both browser + mobile

RQ1: Is time redistributed to other sites on the same platform?

- Reducing time on one site reduces time elsewhere on browser (reduction) but not mobile (isolation)

RQ2: Is time redistributed across devices?

- Time is not redistributed across devices (isolation hypothesis)

Discussion and implications

- We did not observe negative secondary effects of productivity interventions (on other apps, sites, or devices)
- On browsers, there’s actually a reduction elsewhere, from reducing time on sites (likely due to aggregator sites)

Discussion and implications

- When designing interventions, we should consider effects not just on the targeted behavior, but the workflow as a whole

Limitations

- Only monitoring time on phones and browsers
  - Cannot observe if time is being redistributed to non-digital activities
- Only studied productivity domain
  - “Absence of negative secondary effects” may not generalize to other behavior change domains

- Does reducing time via interventions influence time spent elsewhere?

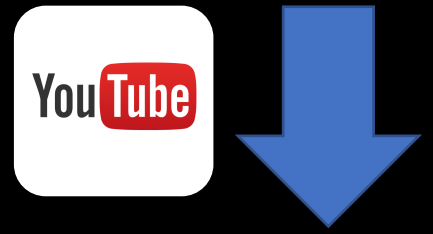
- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser

Isolation

Redistribution

Reduction

Browser



- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser, but not mobile

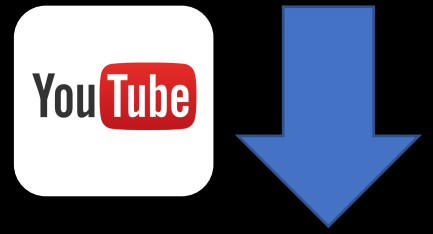
Isolation

Mobile

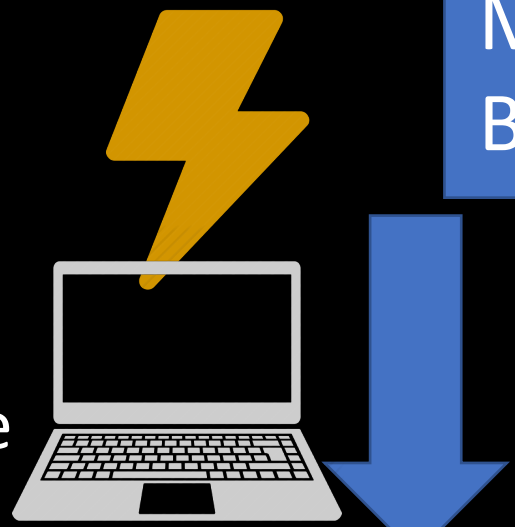
Redistribution

Reduction

Browser



- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser, but not mobile
- Cross-device: no effects



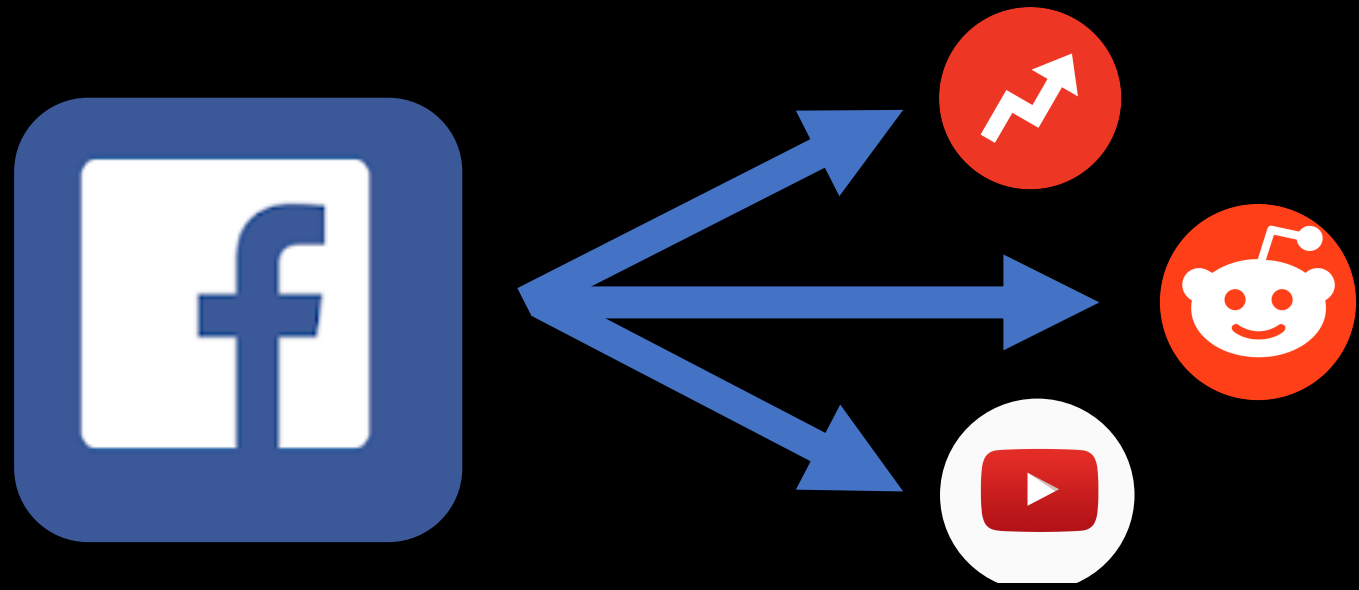
Isolation

Mobile → Browser
Browser → Mobile

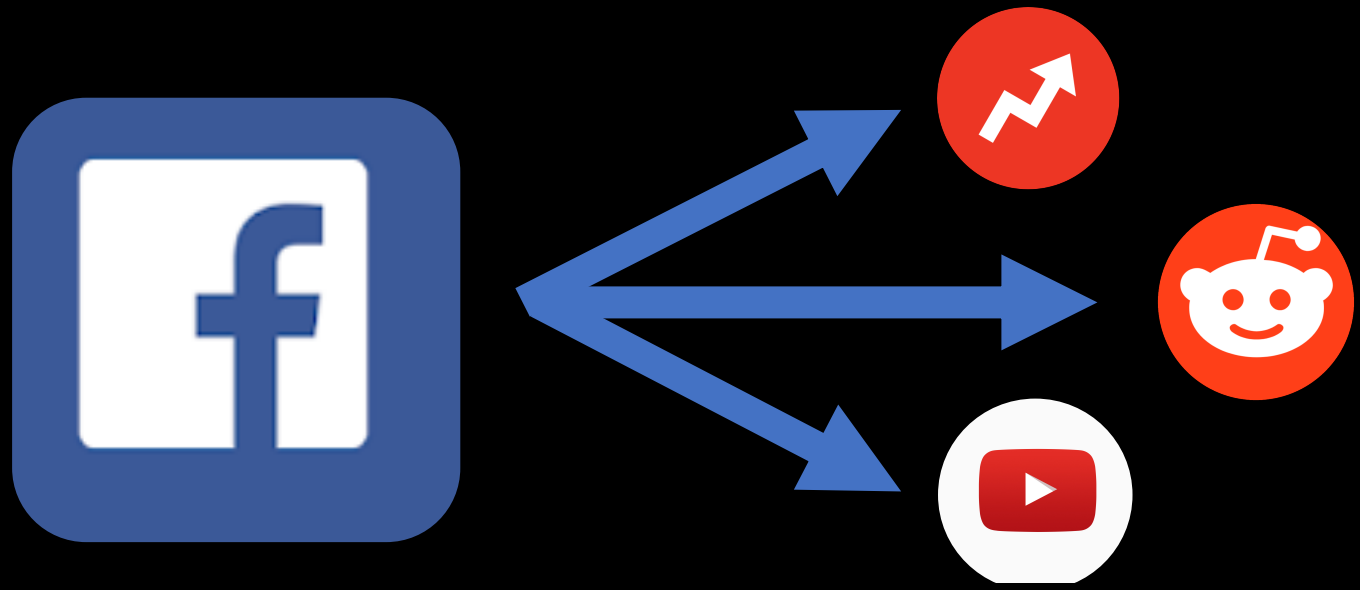
Redistribution

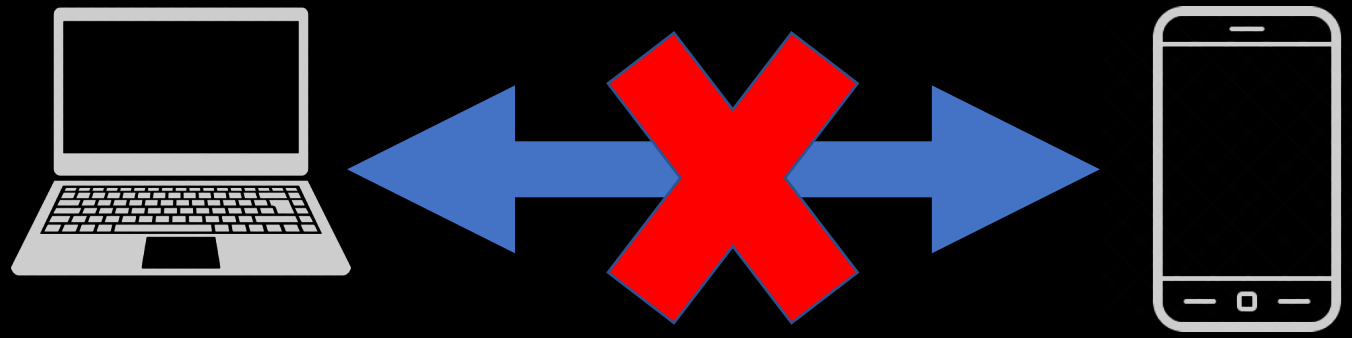
Reduction

- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser, but not mobile
  - Due to aggregator sites driving traffic to other sites
- Cross-device: no effects



- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser, but not mobile
  - Due to aggregator sites driving traffic to other sites
- Cross-device: no effects
  - Devices used in different contexts





HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It? (CHI 2019)

- Does reducing time via interventions influence time spent elsewhere?
- Within-device: reduction on browser, but not mobile
  - Due to aggregator sites driving traffic to other sites
- Cross-device: no effects
  - Devices used in different contexts

HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

How do users' preferences change over time?

Sorry to see you go! HabitLab has been uninstalled

Thanks for trying HabitLab! We'd appreciate your feedback so we can make it better! Why are you uninstalling?

- [ ] **Interventions were annoying**
- [ ] **Was causing lag**
- [ ] **Did not feel effective**
- [ ] **Privacy concerns**
- [ ] **Other reason**

We would appreciate your feedback here

Submit

Users have different expectations for what they want productivity tools to do

In our uninstall feedback form, we had many users who stated interventions were too hard, but others stated that they though the interventions were too easy

Users have different expectations for what they want productivity tools to do

In our uninstall feedback form, we had many users who stated interventions were too hard, but others stated that they though the interventions were too easy

Naive solution: Classify interventions according to difficulty, and ask them during onboarding what difficulty they would prefer

Users have different expectations for what they want productivity tools to do

In our uninstall feedback form, we had many users who stated interventions were too hard, but others stated that they though the interventions were too easy

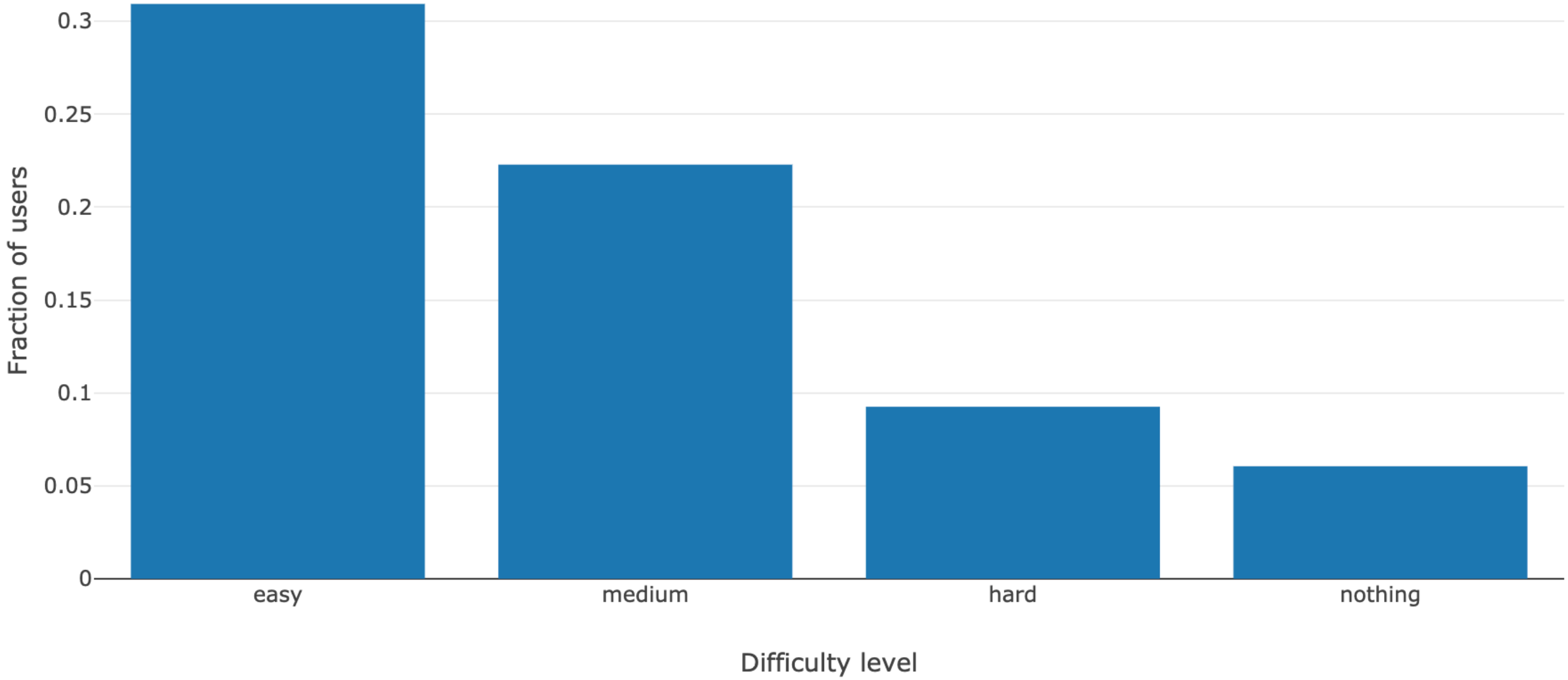
Naive solution: Classify interventions according to difficulty, and ask them during onboarding what difficulty they would prefer

How aggressive do you want HabitLab to be in helping you reduce your time spent online?

- ○ **Don't do anything** - just track time.
- ○ **Light touch** - e.g., show a timer when you visit Facebook.
- ○ **Medium** - e.g., remove your Facebook feed until you click to show it.
- ○ **Heavy handed** - e.g., close the site after 60 seconds.

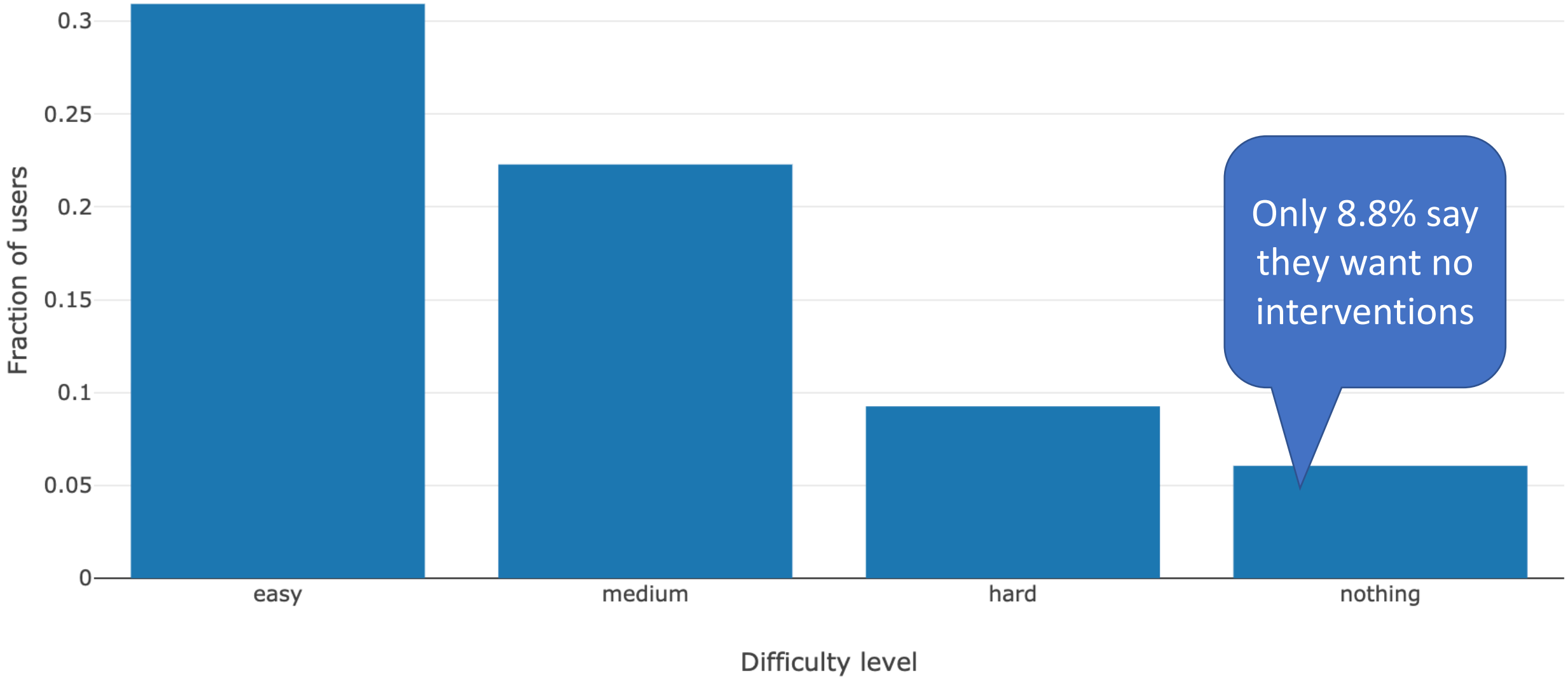
Initial difficulties chosen during onboarding (n = 5114 users)

All pairs statistically significantly different (p < 10^-9, chi-squared tests)



Initial difficulties chosen during onboarding (n = 5114 users)

All pairs statistically significantly different (p < 10^-9, chi-squared tests)



HabitLab: Our in-the-wild behavior change platform

Outcomes

Does effectiveness remain constant over time?

What externalities exist?

Interventions

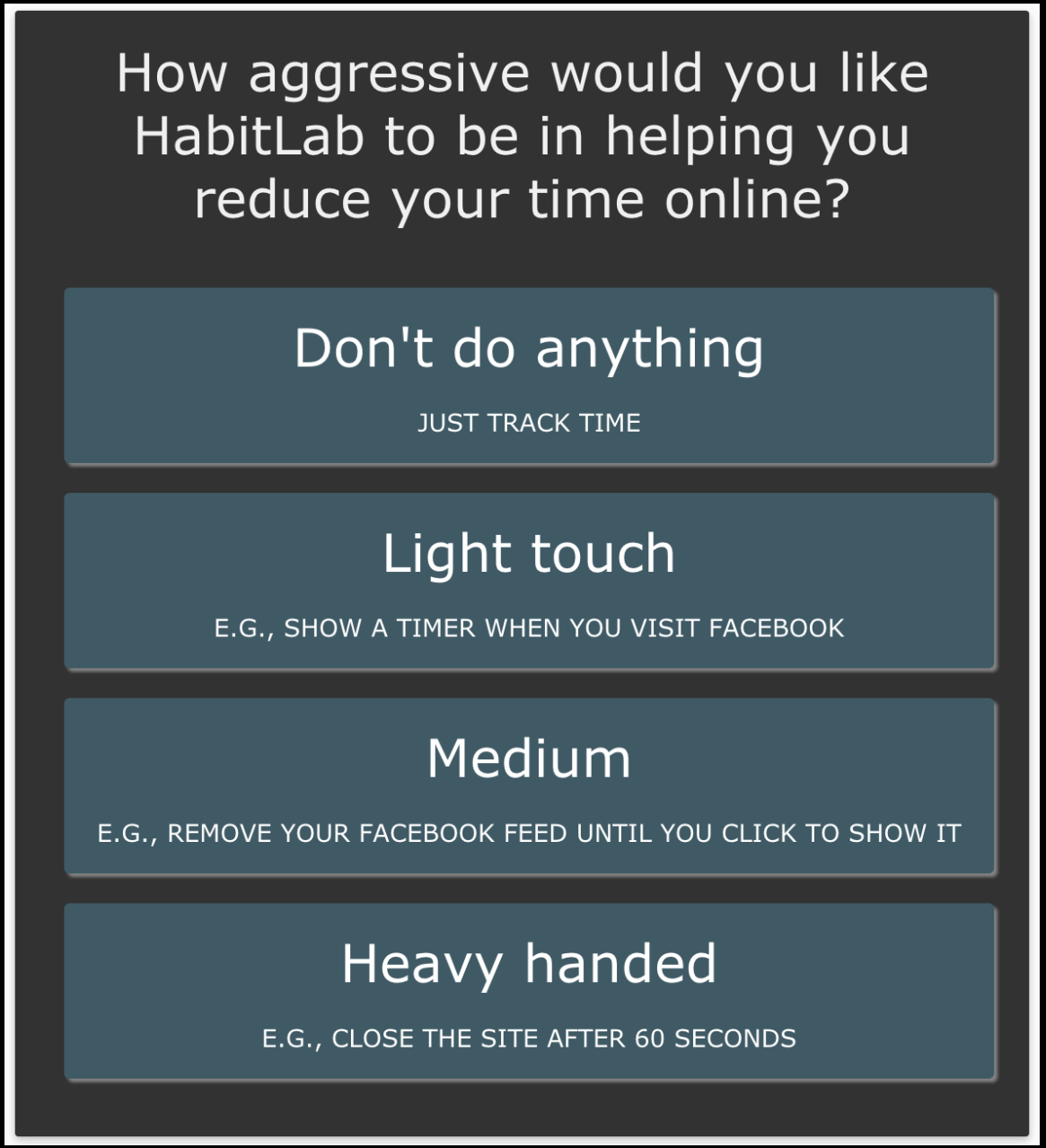
How do users' preferences change over time?

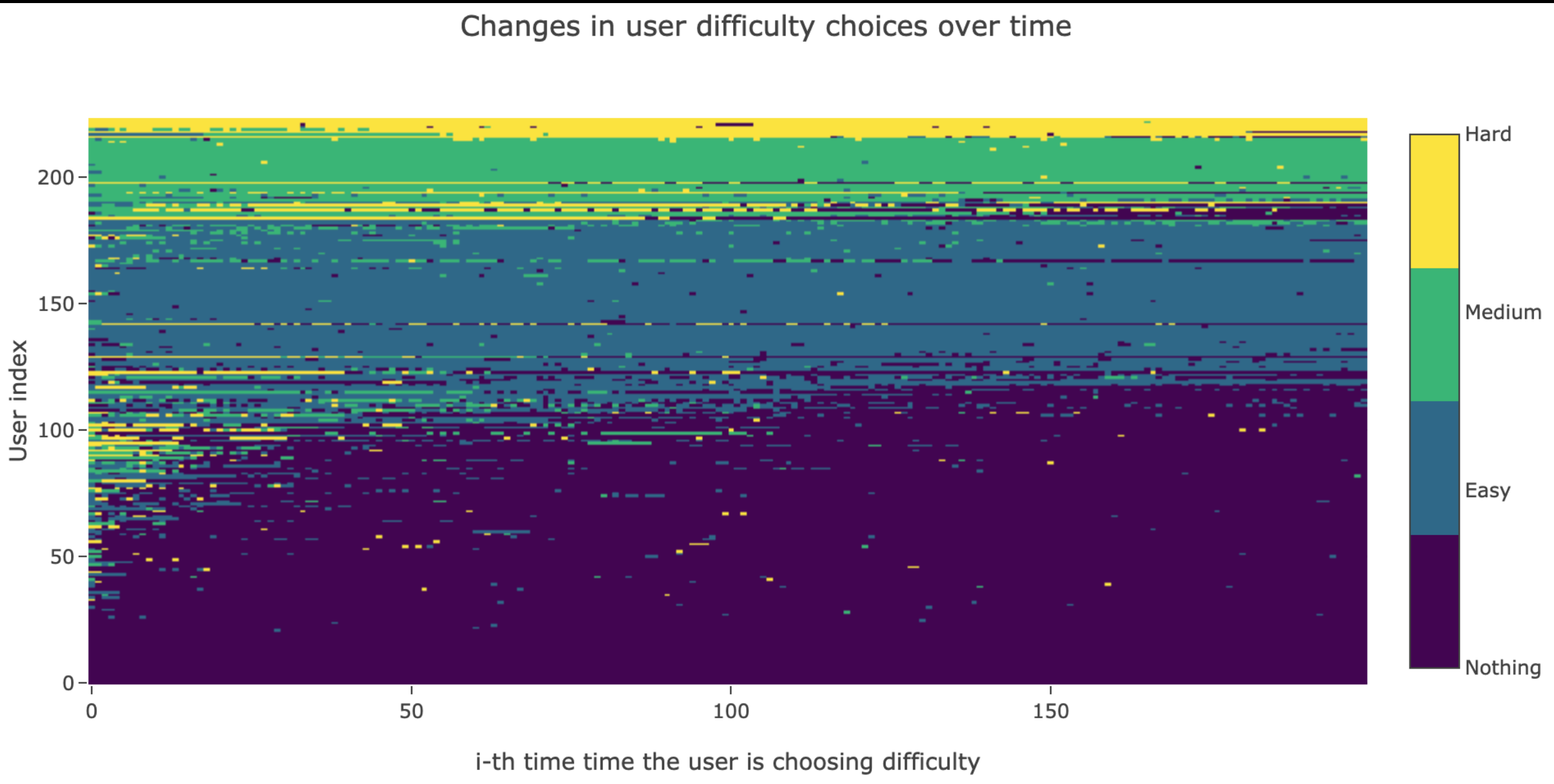
- Users initially are optimistic about how difficult they want their interventions
  - ”No intervention” is least frequently chosen during onboarding

Can users predict during onboarding what interventions will work for them?

Do their preferences change over time?

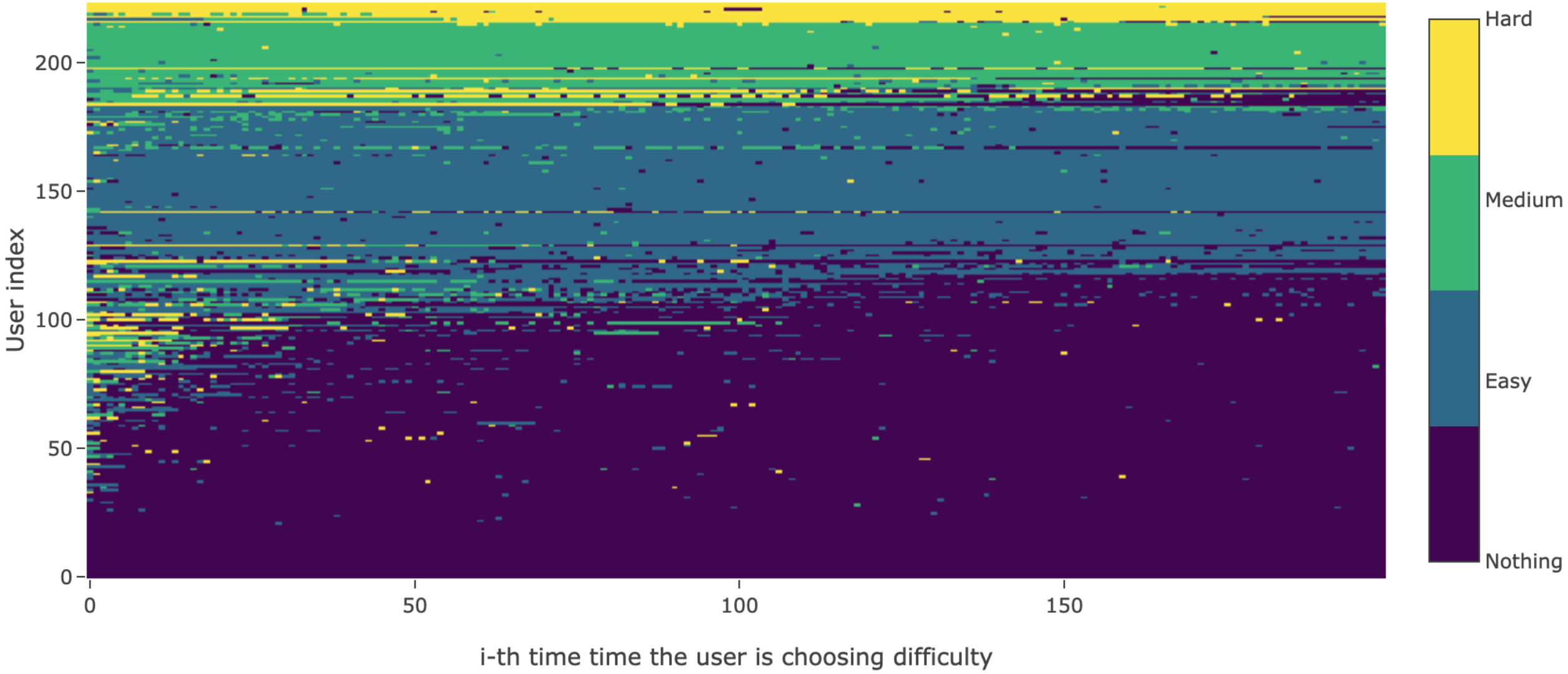
Ask users when they visit site about intervention difficulty preference for this visit

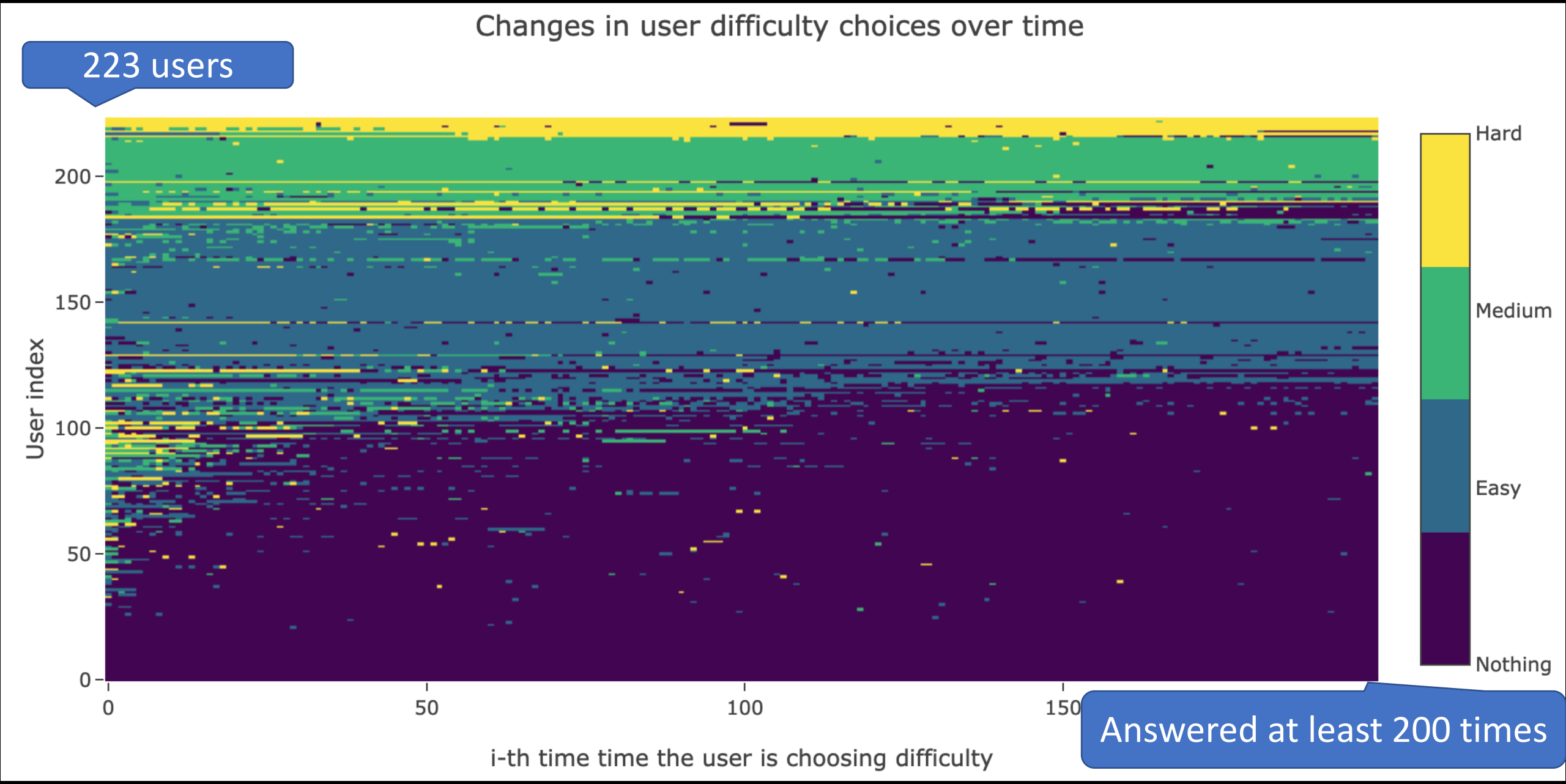


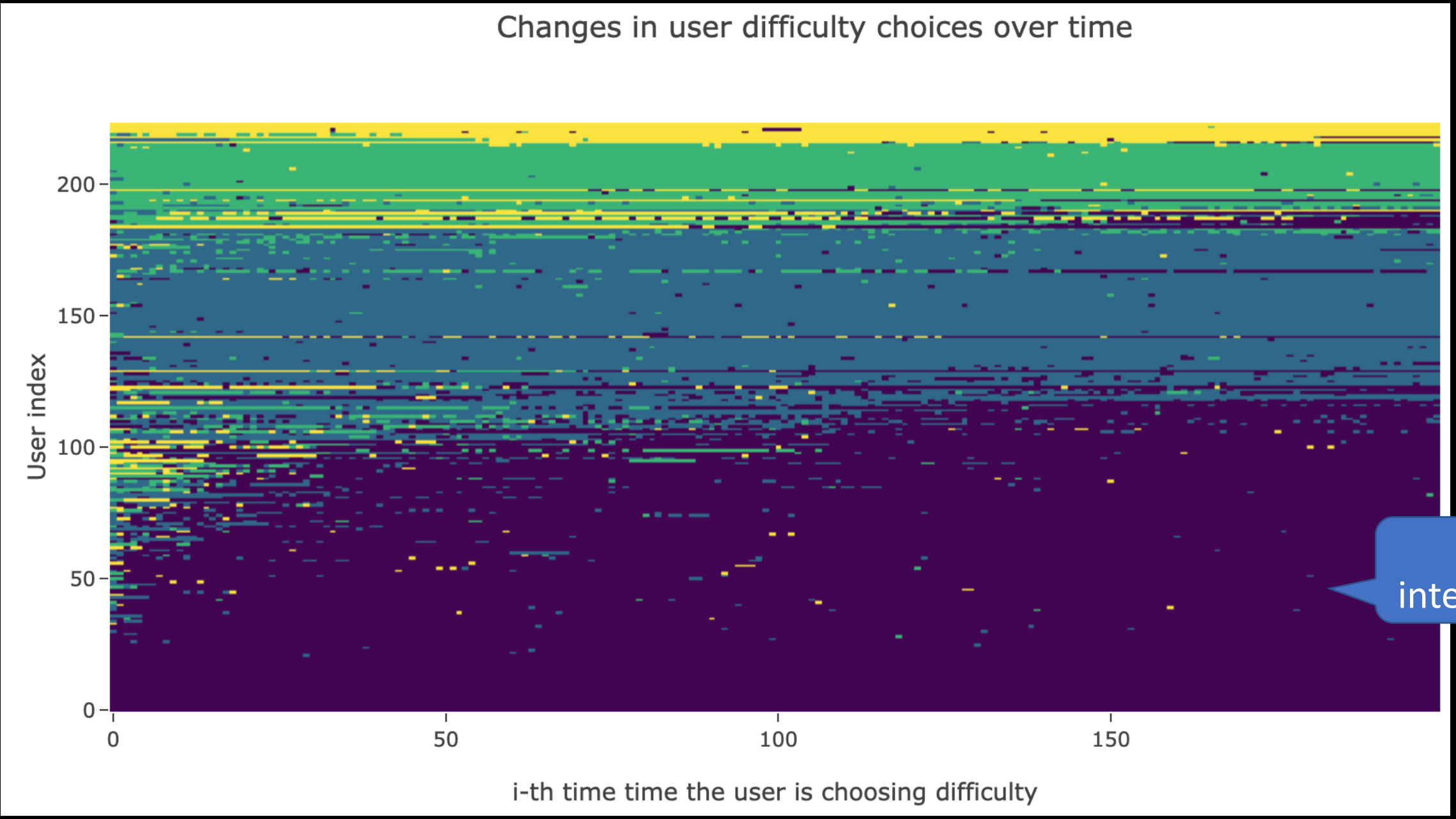

Changes in user difficulty choices over time
User index
0
50
100
150
200
0
50
100
150
i-th time time the user is choosing difficulty
Hard
Medium
Easy
Nothing

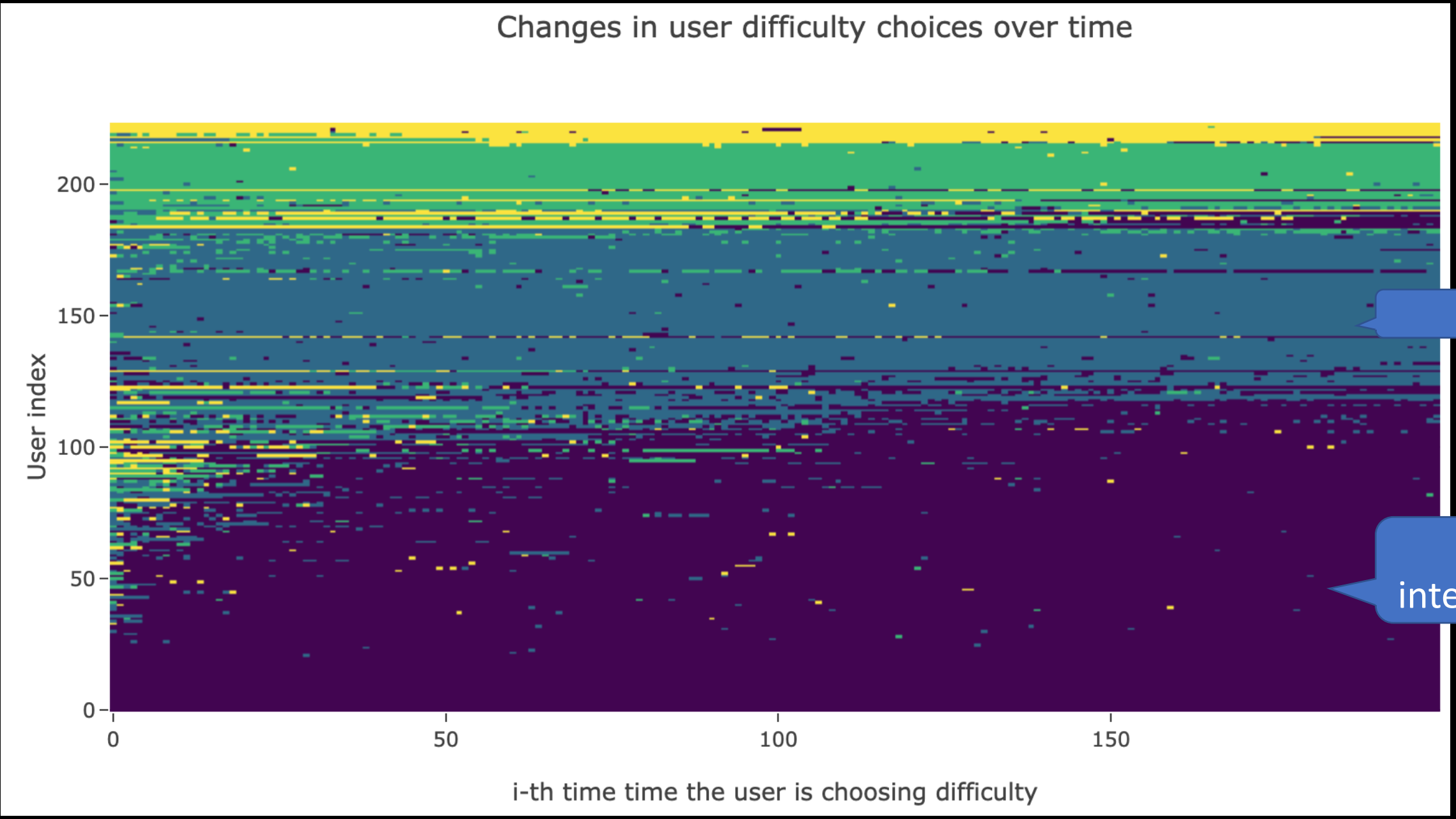
Changes in user difficulty choices over time

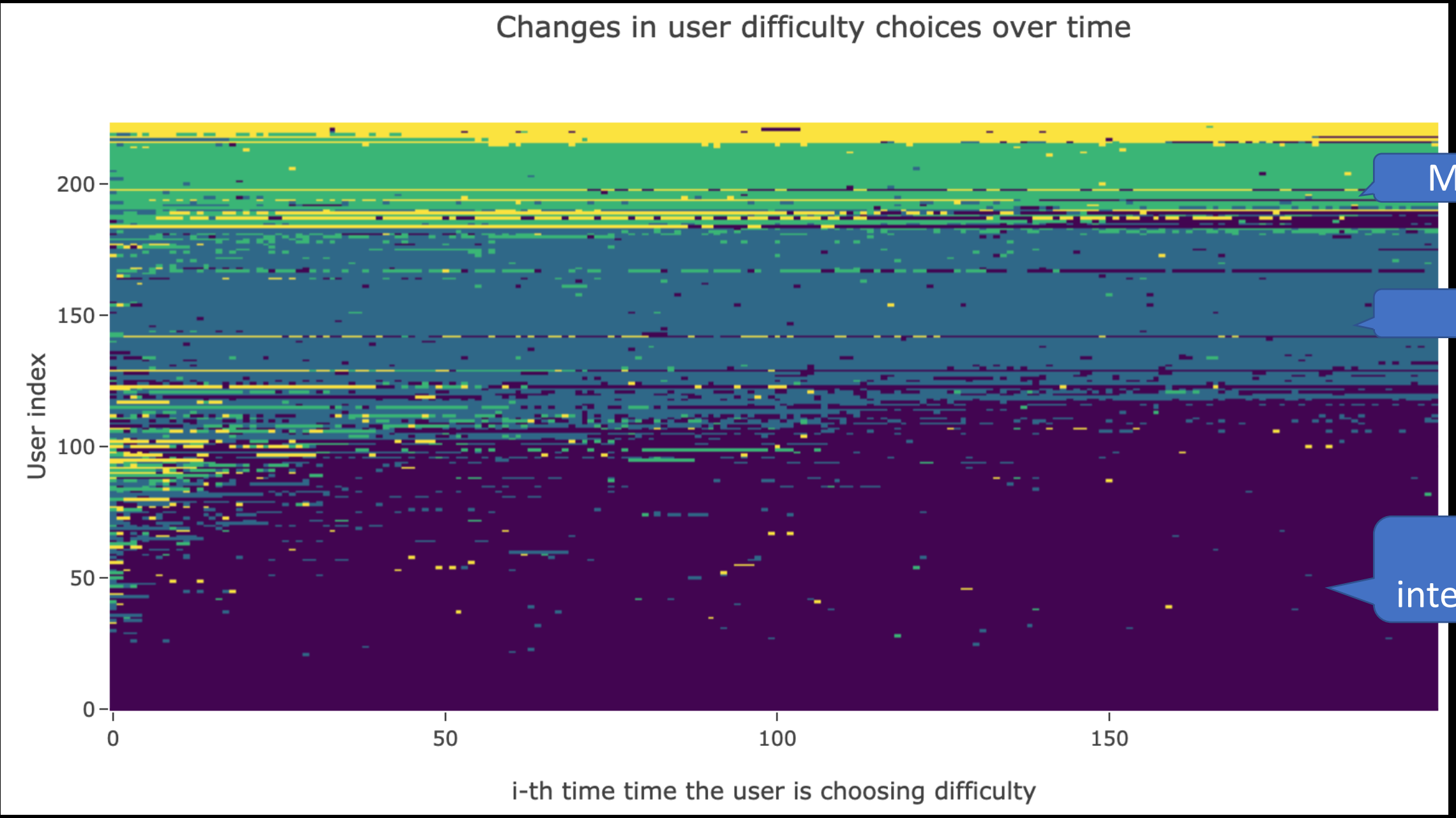
223 users

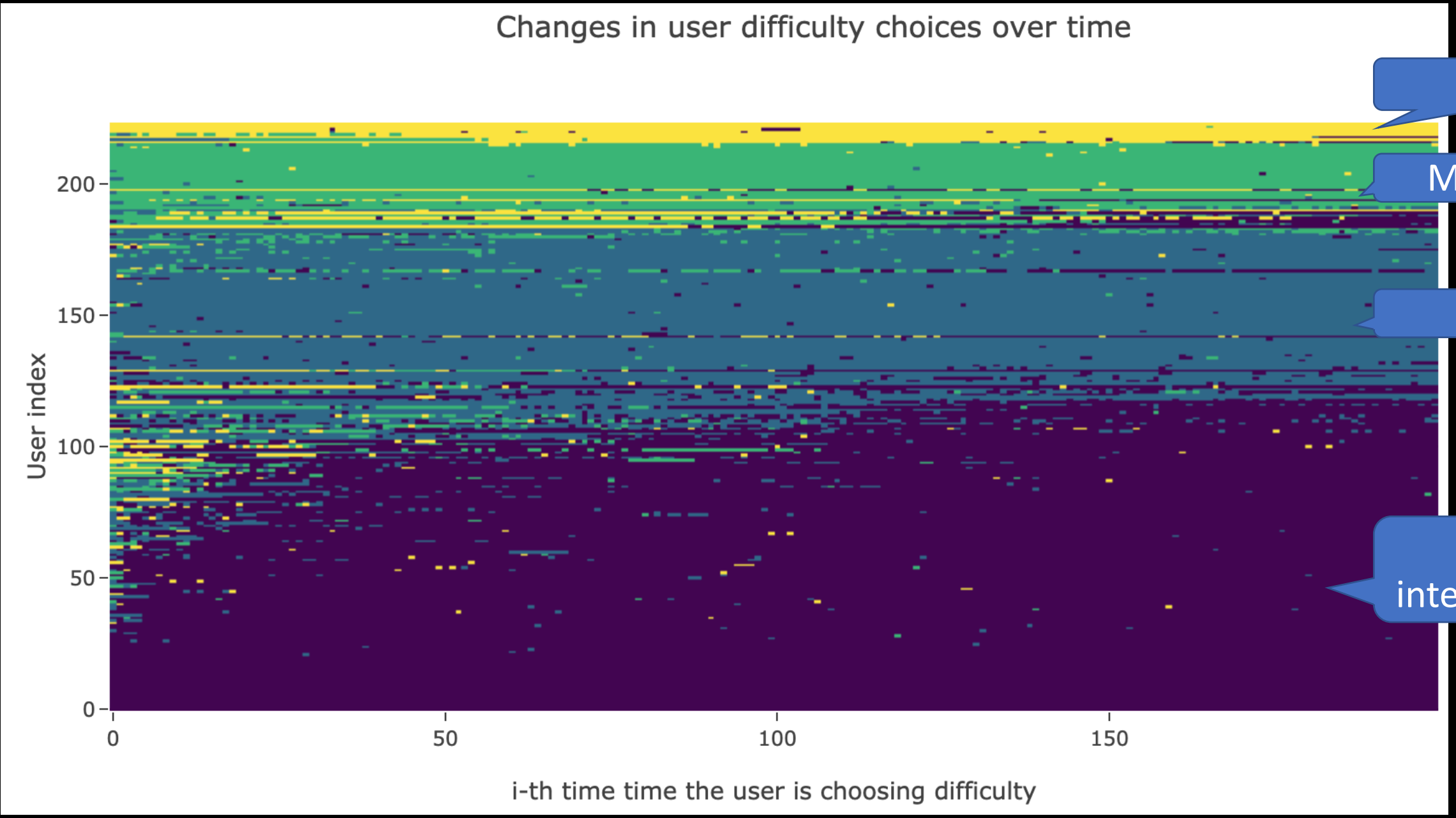


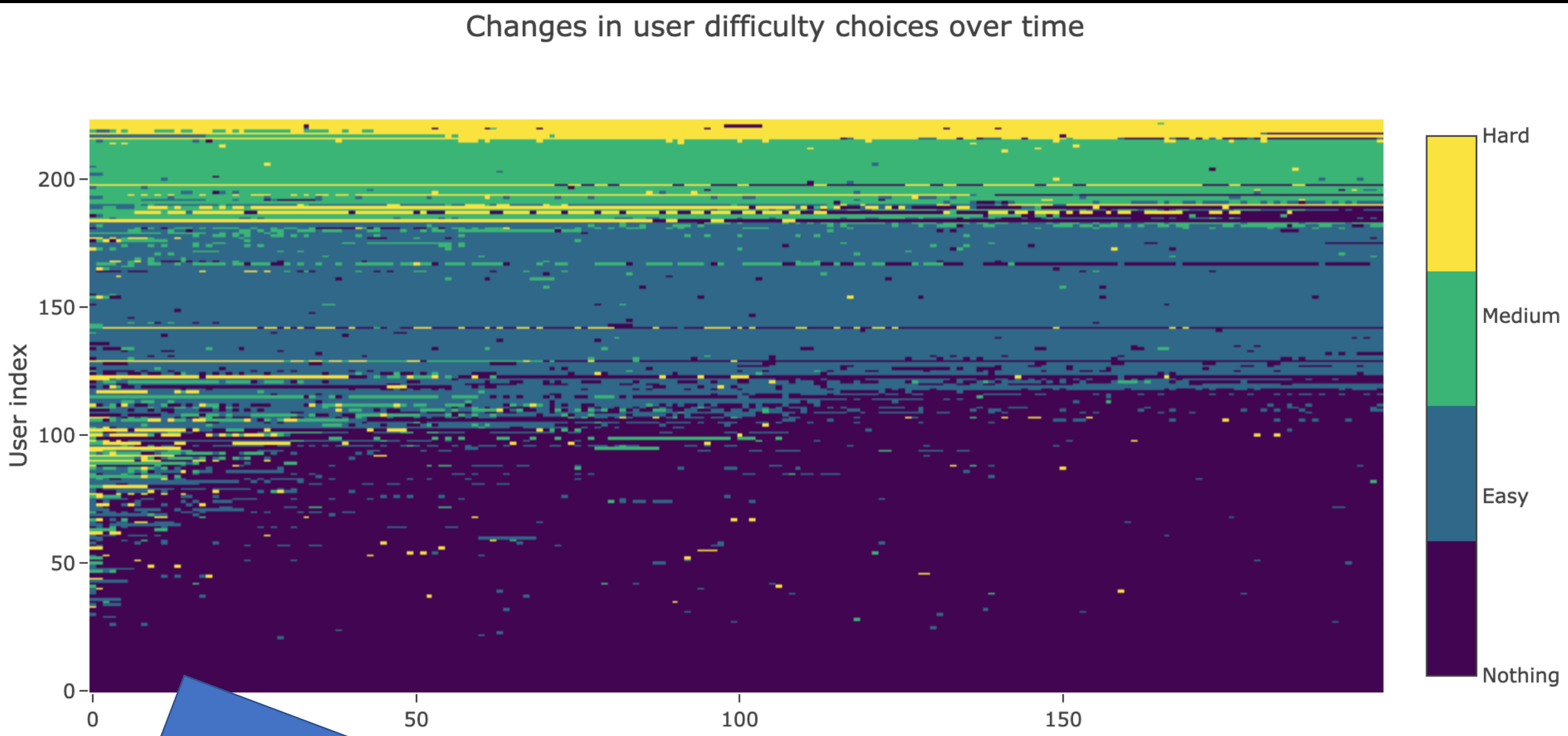

Changes in user difficulty choices over time
223 users
User index
200
150
100
50
0
0
50
100
150
i-th time time the user is choosing difficulty
Hard
Medium
Easy
Nothing
Answered at least 200 times


Changes in user difficulty choices over time
User index
200
150
100
50
0
0
50
100
150
i-th time time the user is choosing difficulty
No intervention

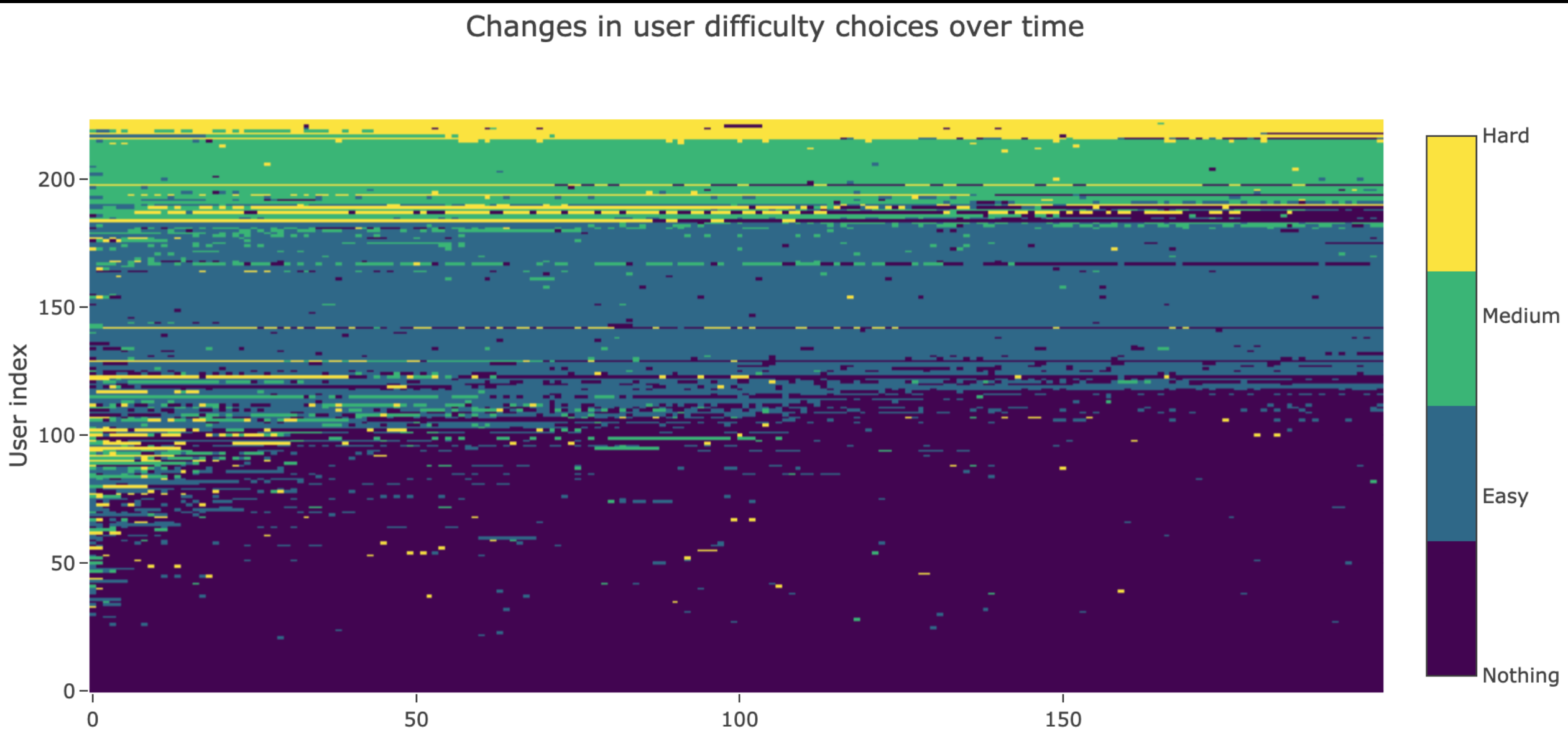

Changes in user difficulty choices over time
User index
200
150
100
50
0
0
50
100
150
i-th time time the user is choosing difficulty
Easy
No intervention


Changes in user difficulty choices over time
User index
200
150
100
50
0
0
50
100
150
i-th time time the user is choosing difficulty
Medium
Easy
No intervention

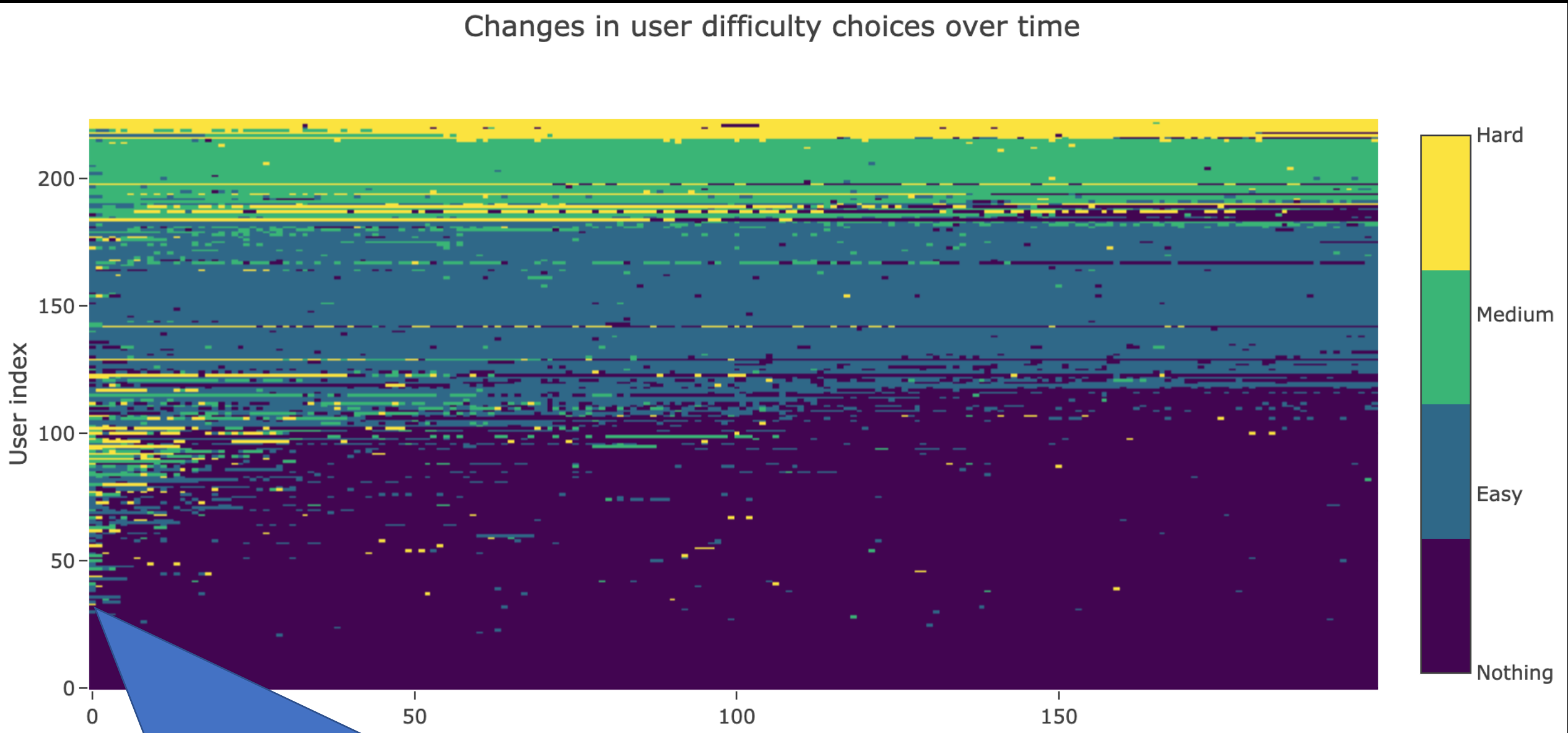

Changes in user difficulty choices over time
User index
0
50
100
150
200
0
50
100
150
i-th time time the user is choosing difficulty
Hard
Medium
Easy
No intervention



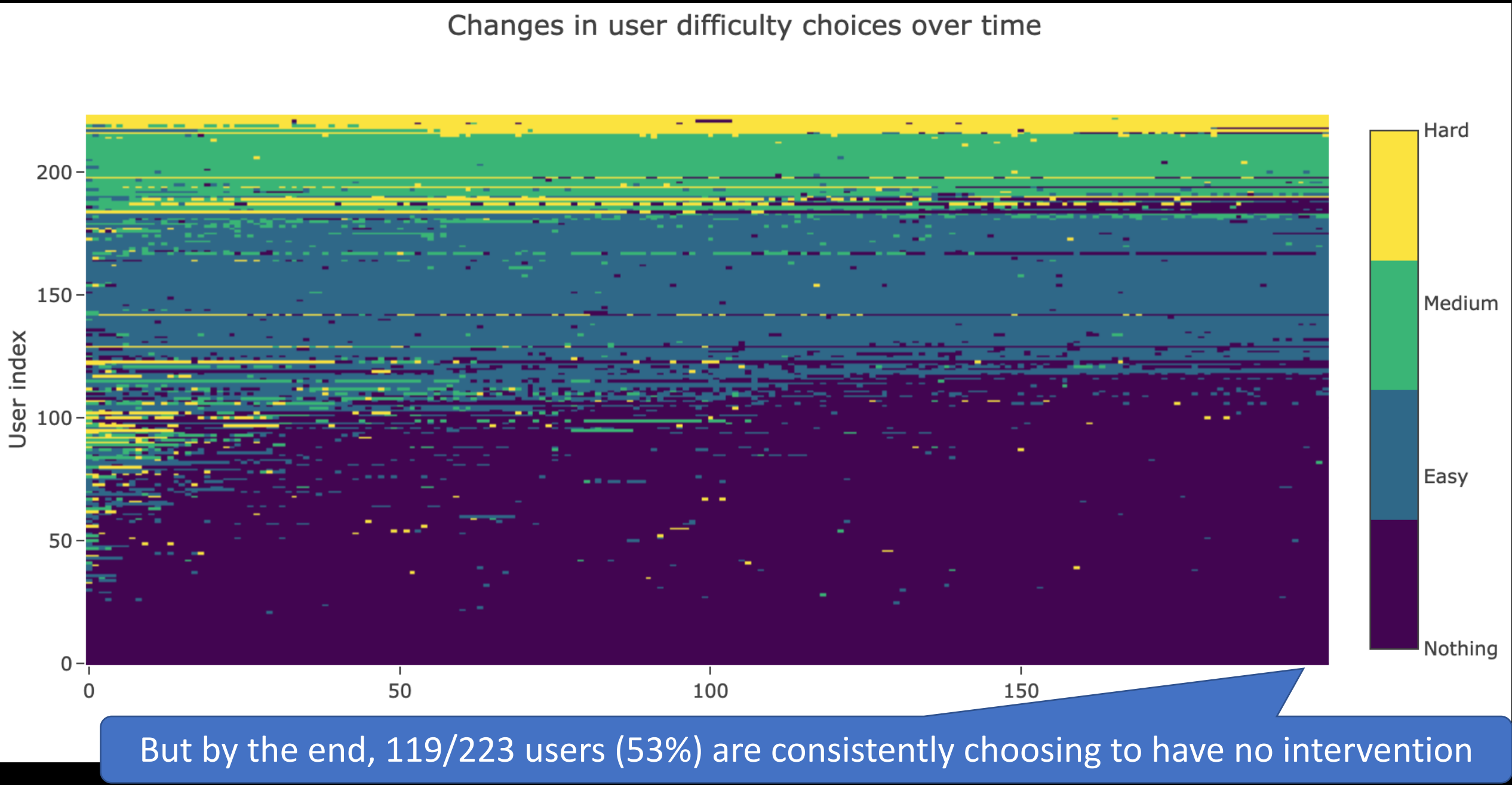
Lots of initial exploration between intervention difficulties



After answering 100 times, preferences become mostly stable



Initially, only 29 are consistently choosing “no intervention”



But by the end, 119/223 users (53%) are consistently choosing to have no intervention

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

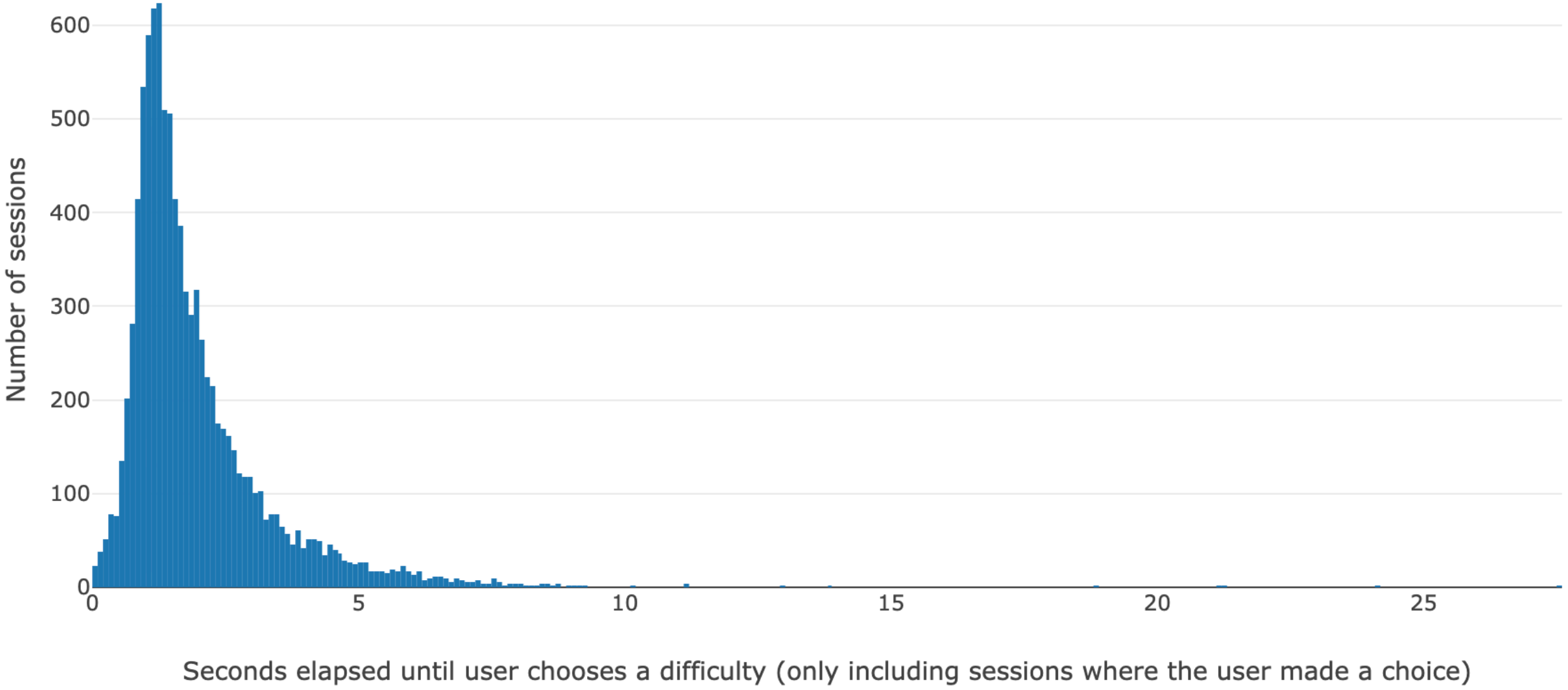
- Users initially are optimistic about how difficult they want their interventions
  - "No intervention" is least frequently chosen during onboarding
- Intervention difficulty choices decline over time
  - **Half of users eventually choose "no intervention" nearly always**

How can we determine when preferred intervention difficulty changes?

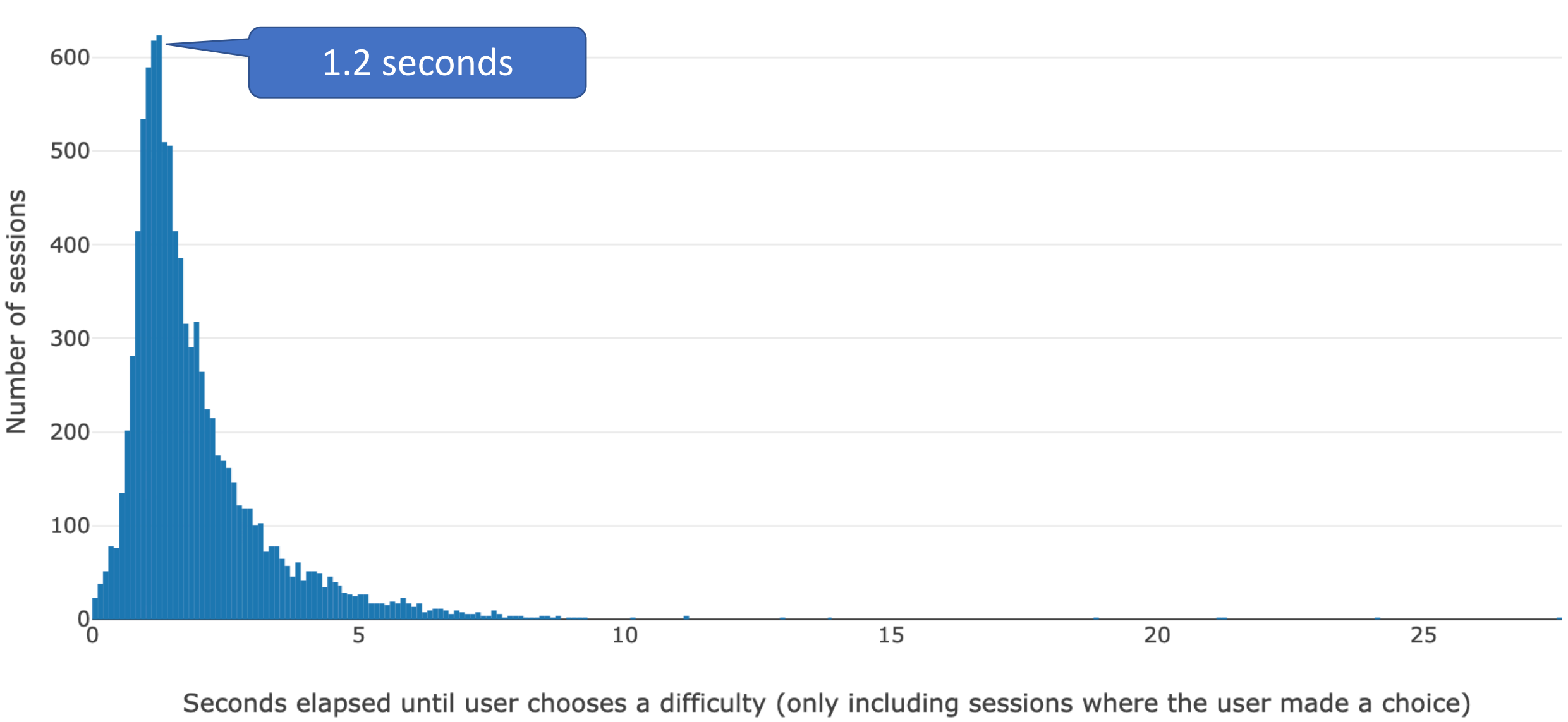
If we periodically ask users their preferred intervention difficulty:

- How frequently do we need to ask to get accurate results?
- What are the costs of asking? (Time? Attrition? Response rate?)

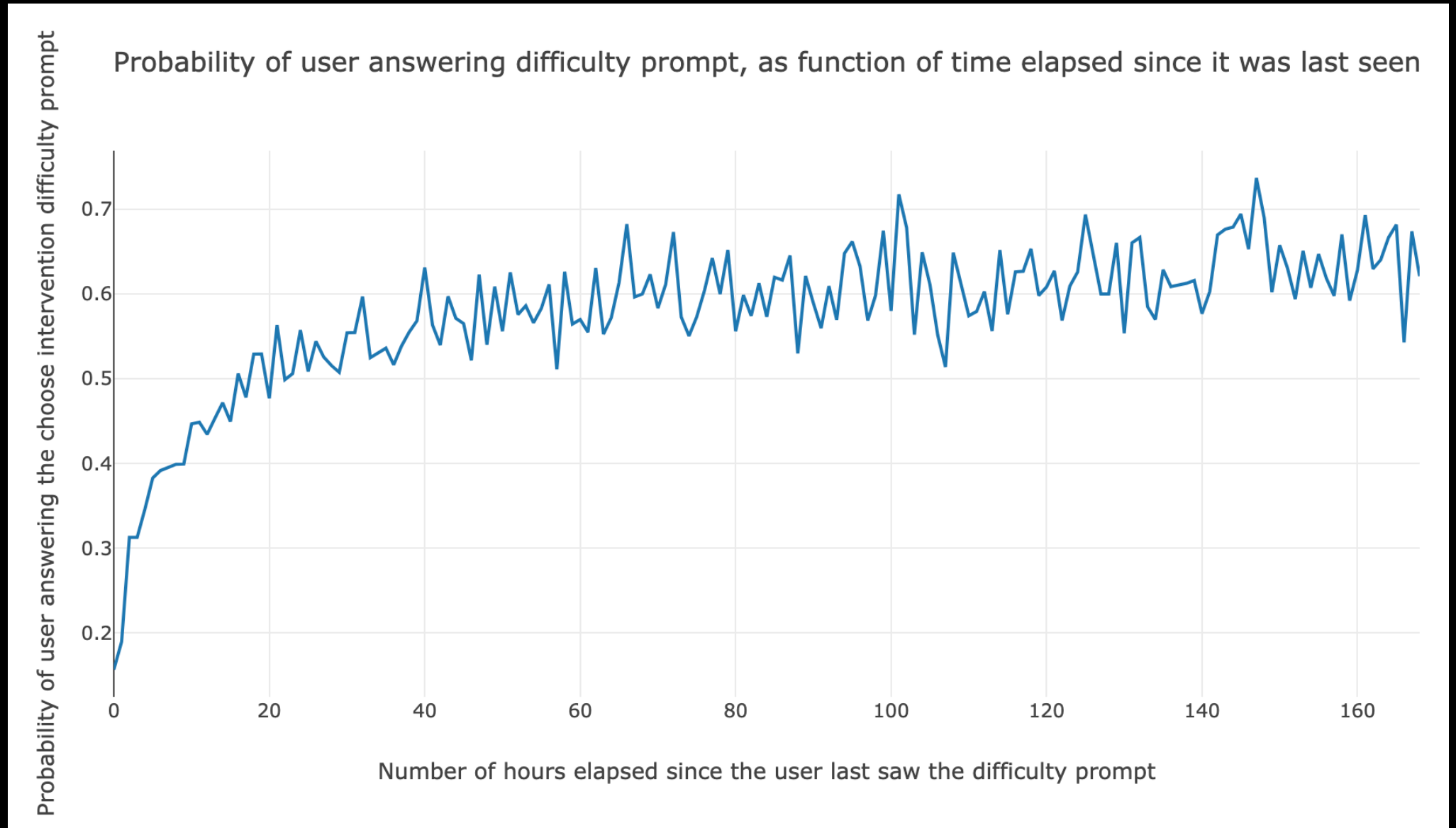
Time it requires users to choose difficulty



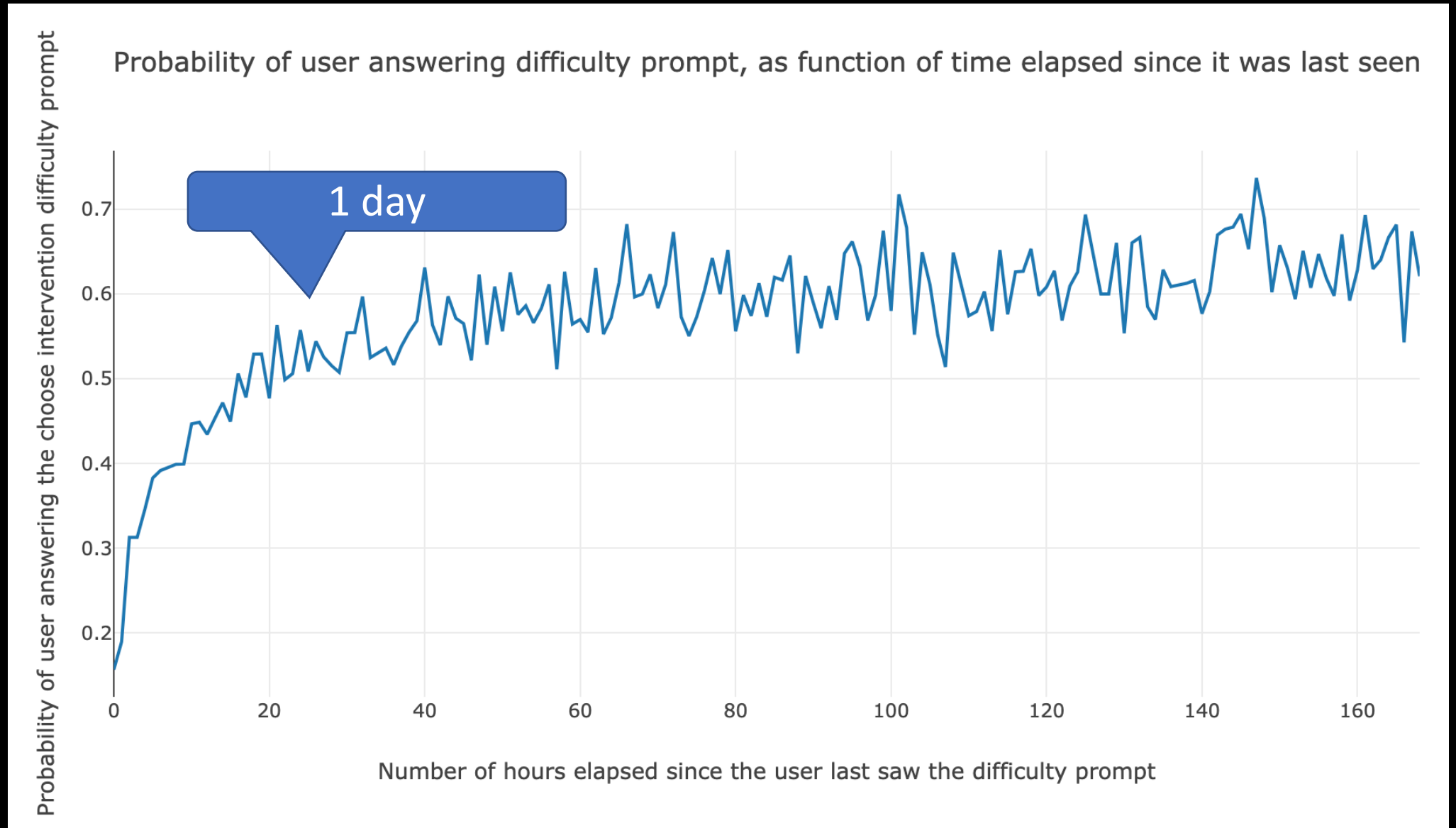
Time costs are low: 1.2 seconds to choose difficulty



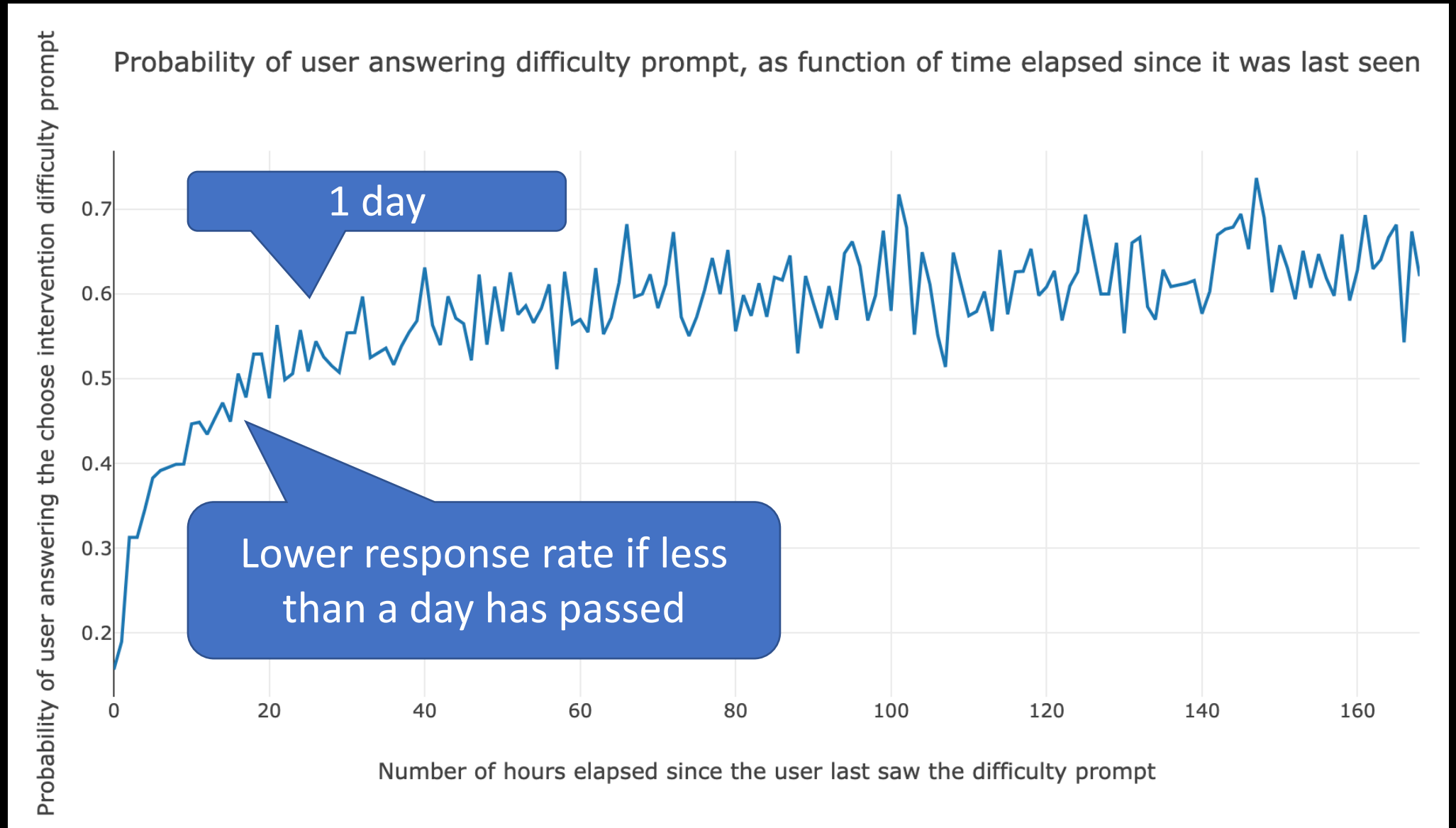
If less than a day has passed since they were last asked, users are less likely to answer



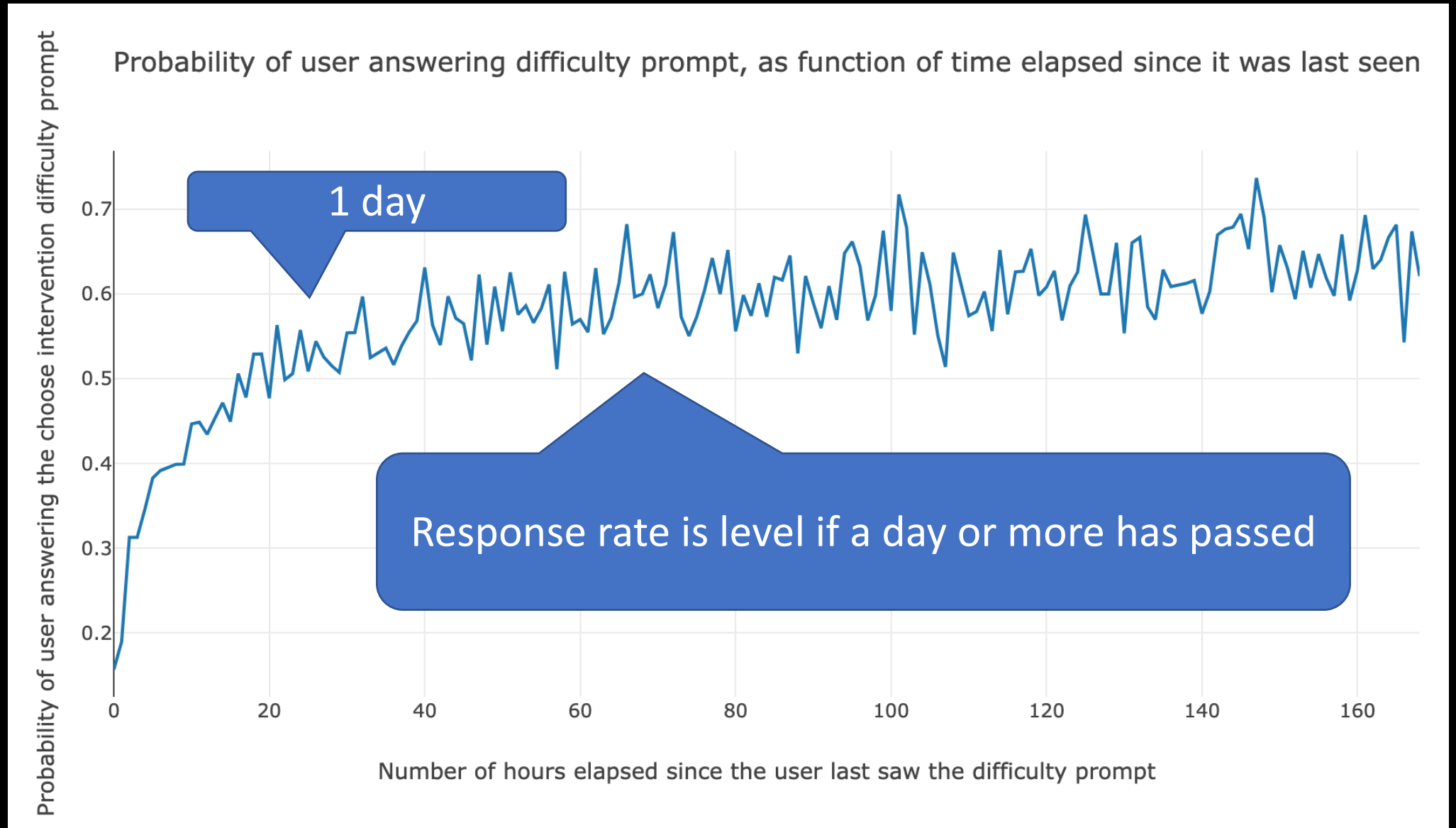
If less than a day has passed since they were last asked, users are less likely to answer



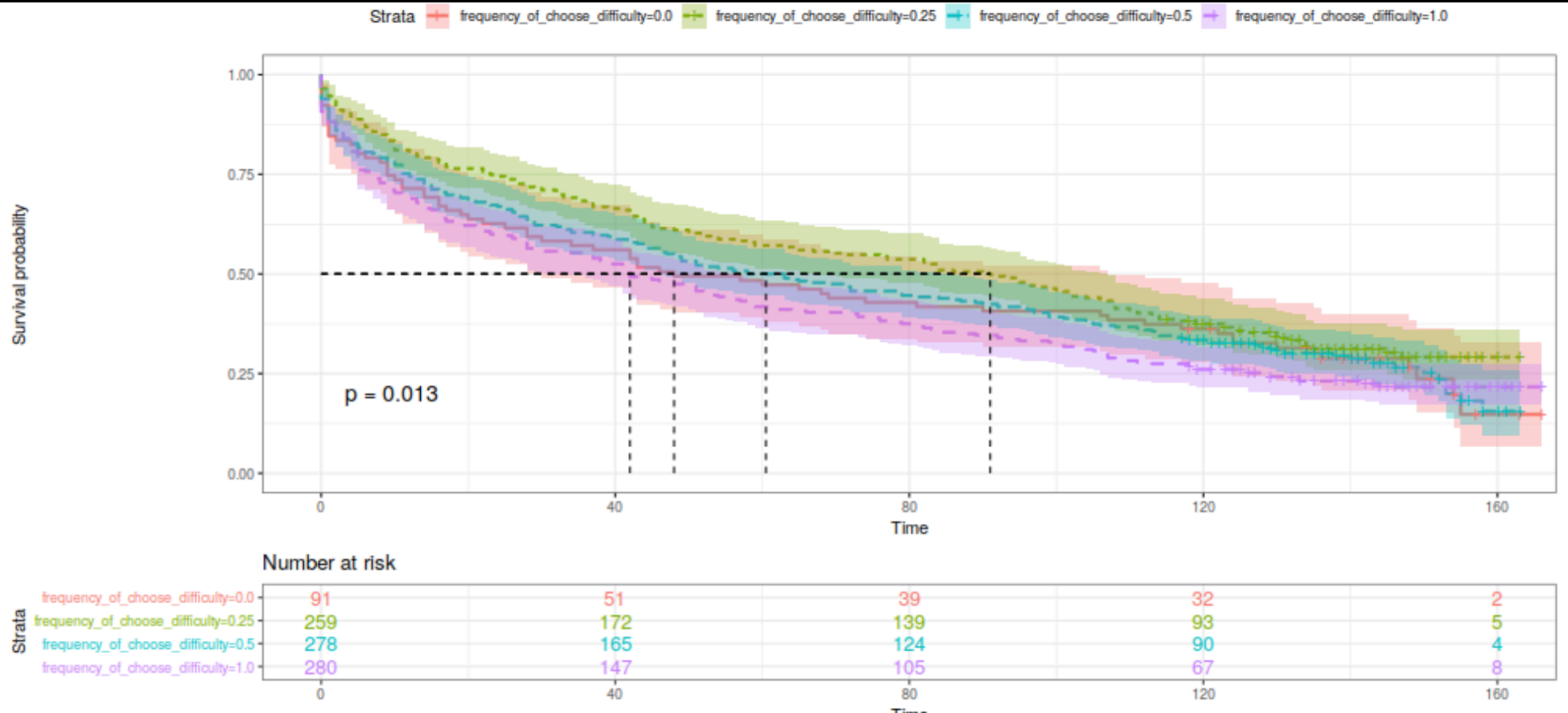
If less than a day has passed since they were last asked, users are less likely to answer



If less than a day has passed since they were last asked, users are less likely to answer



Excessive experience sampling increases attrition

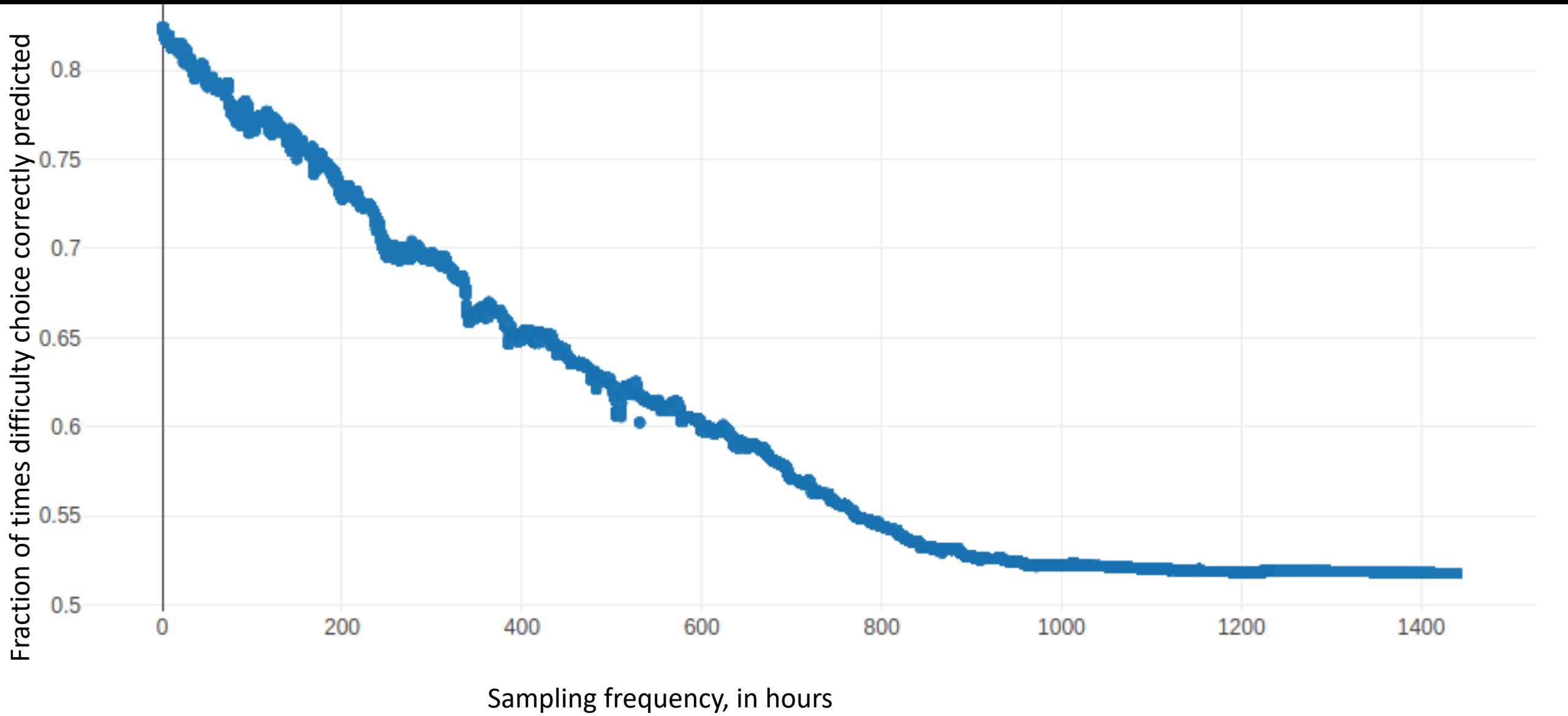


How can we determine when preferred intervention difficulty changes?

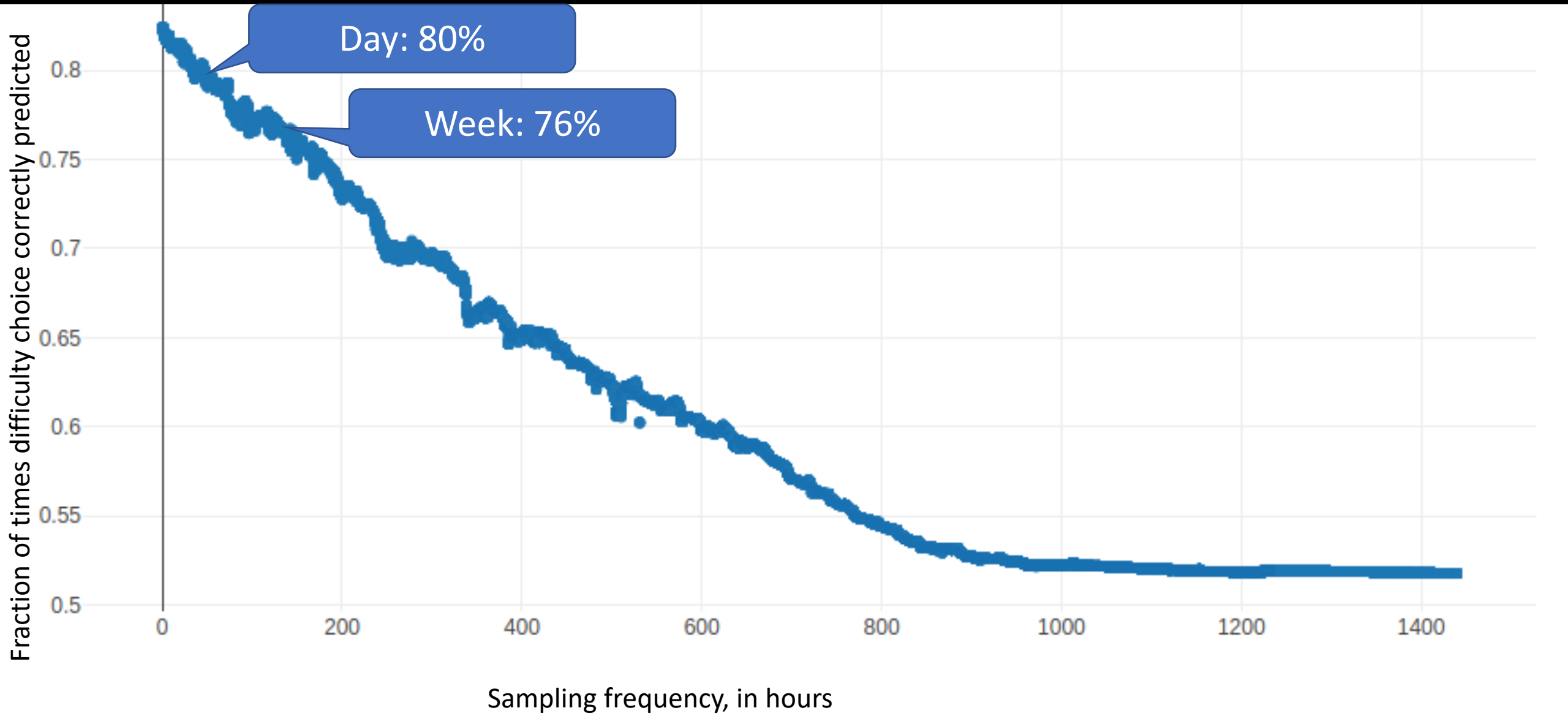
If we periodically ask users their preferred intervention difficulty:

- What are the costs of asking? (Time? Attrition? Response rate?)
  - Low time cost (1.2 seconds)
  - Response rate declines if more frequent than daily
- How frequently do we need to ask to get accurate results?

Prediction accuracy declines as sampling frequency decreases



Prediction accuracy declines as sampling frequency decreases



How can we determine when preferred intervention difficulty changes?

If we periodically ask users their preferred intervention difficulty:

- What are the costs of asking? (Time? Attrition? Response rate?)
  - Low time cost (1.2 seconds)
  - Response rate declines if more frequent than daily
- How frequently do we need to ask to get accurate results?
  - Daily still gets reasonably high (80%) accuracy

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

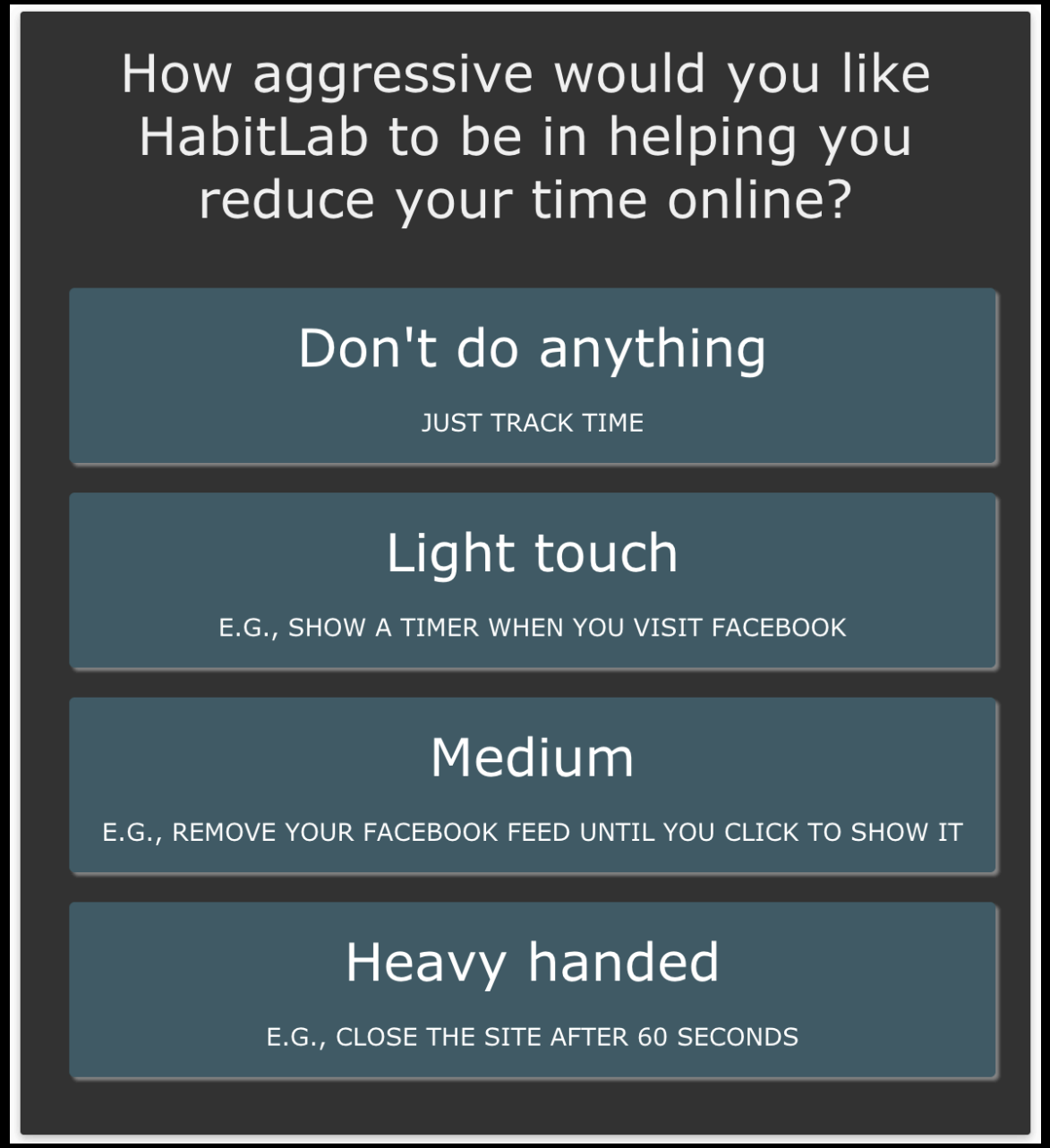
What are the side effects of interventions?

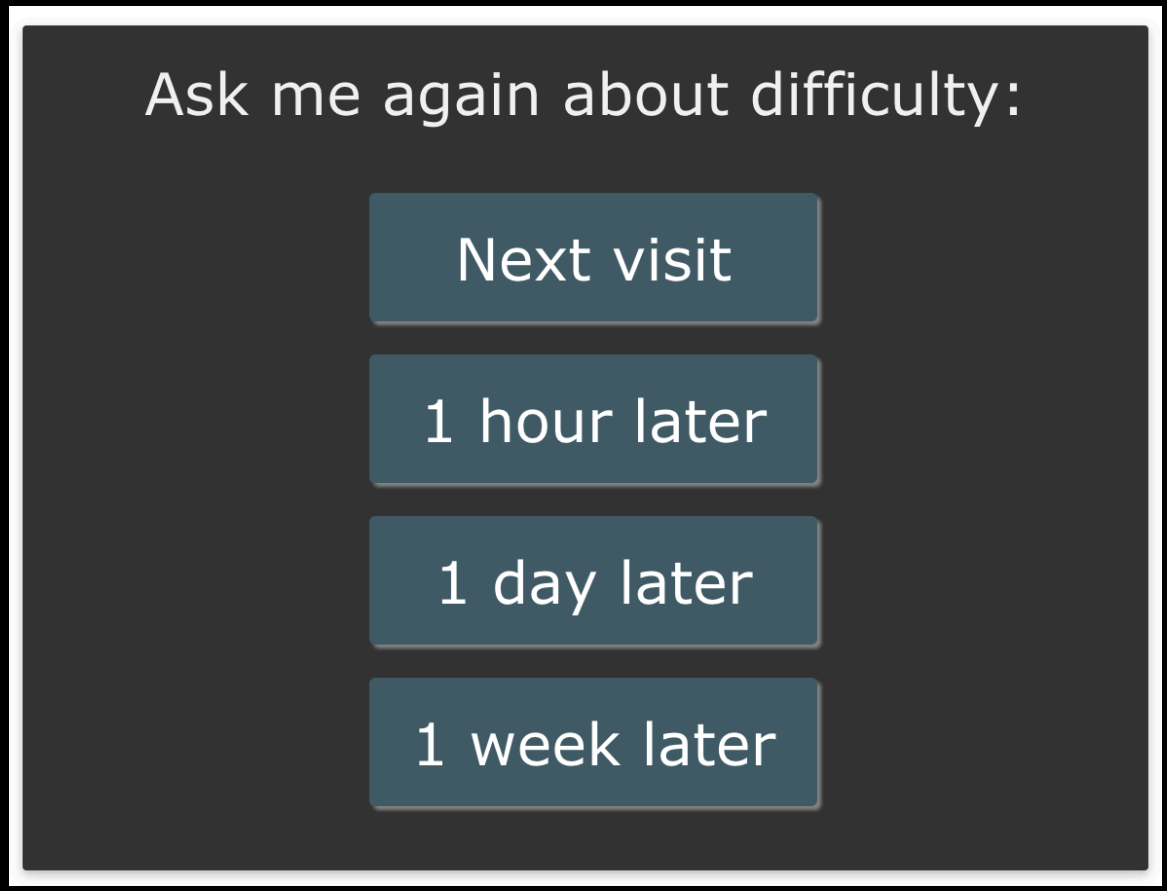
How do users' preferences change over time?

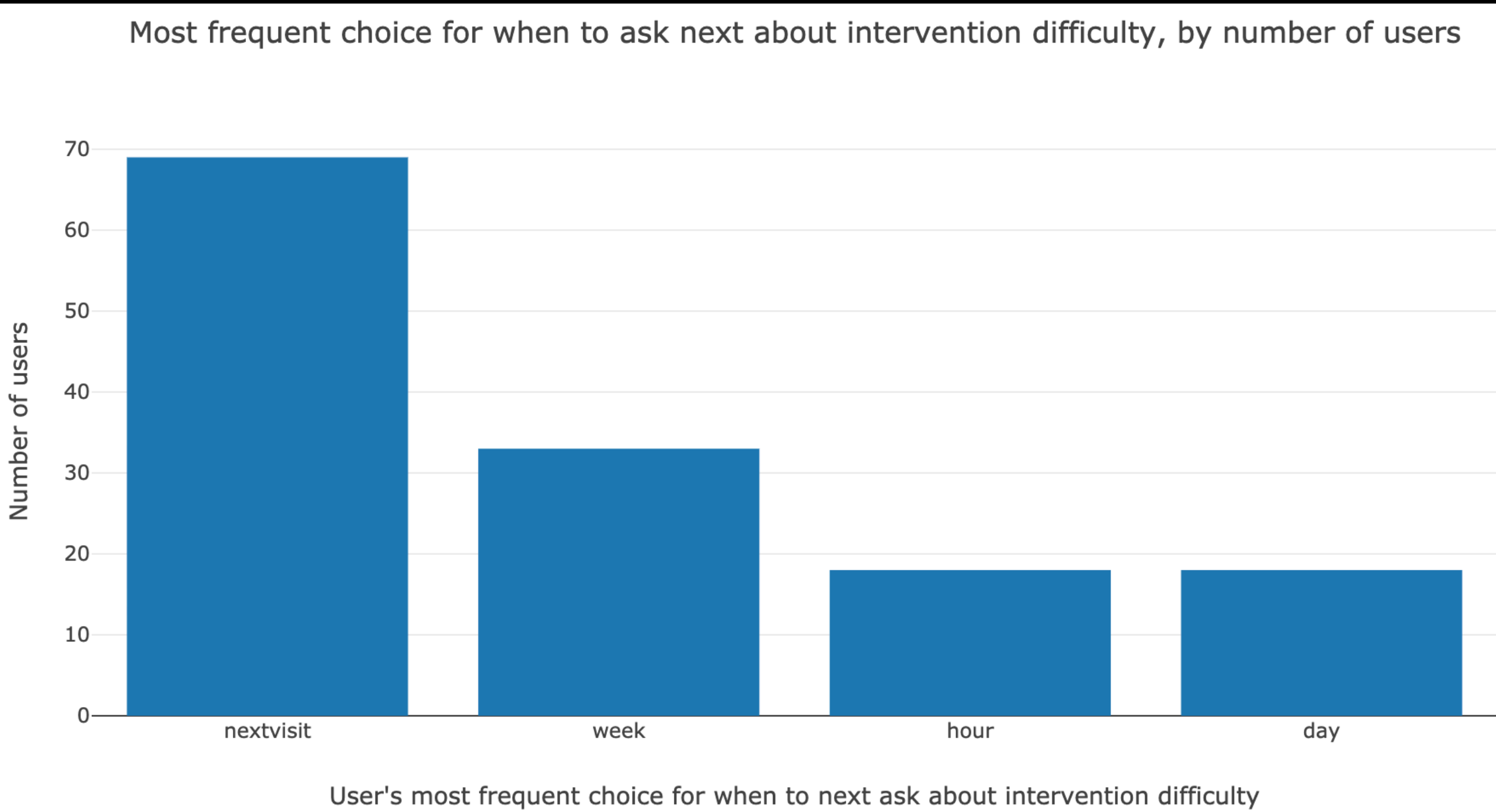
- Users initially are optimistic about how difficult they want their interventions
  - "No intervention" is least frequently chosen during onboarding
- Intervention difficulty choices decline over time
  - Half of users eventually choose "no intervention" nearly always
- **Desired difficulty can be predicted with periodic experience sampling**

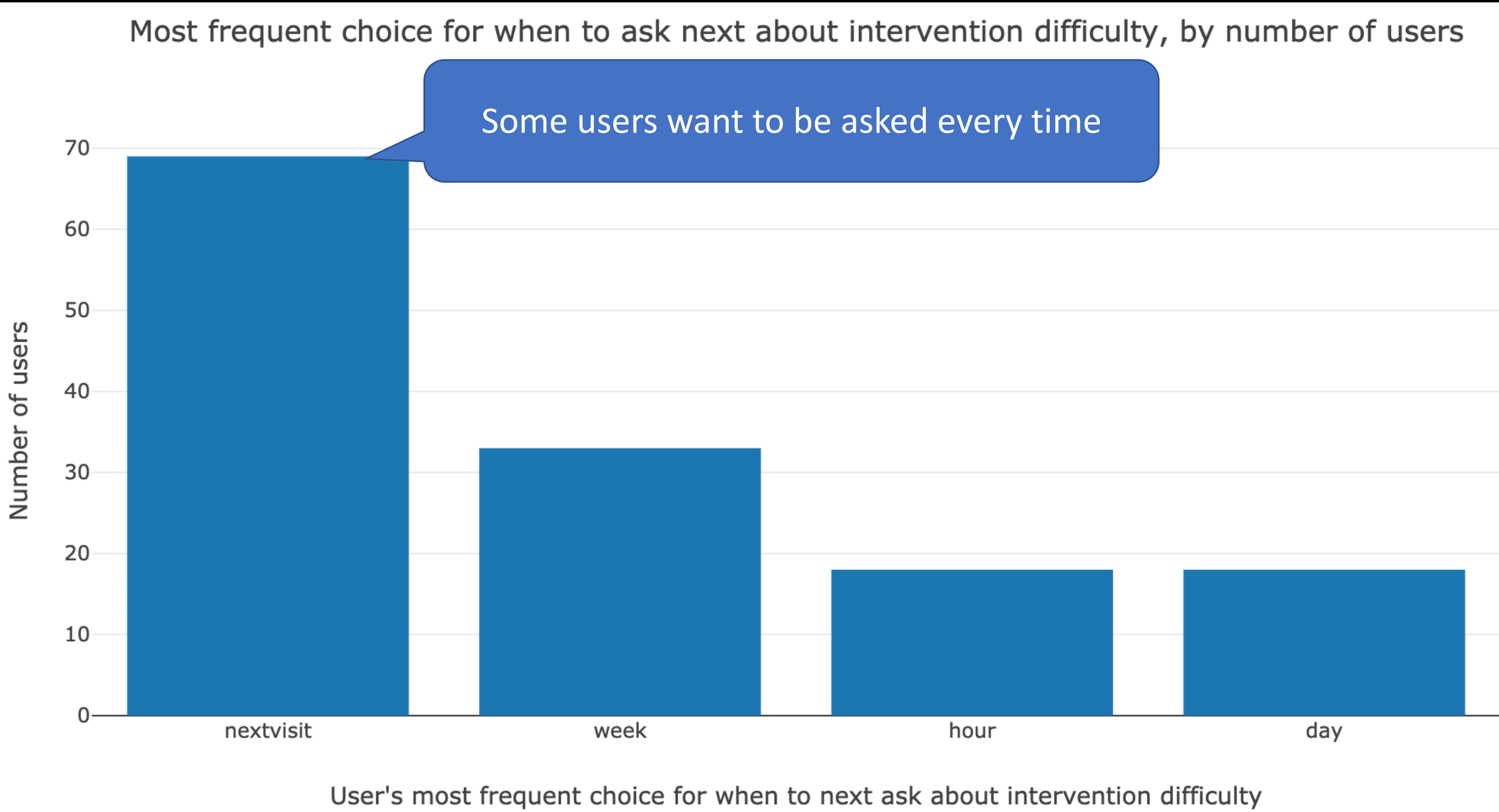
User preferences change over time, and our results suggest asking once a day should get a good balance between accuracy vs sampling costs

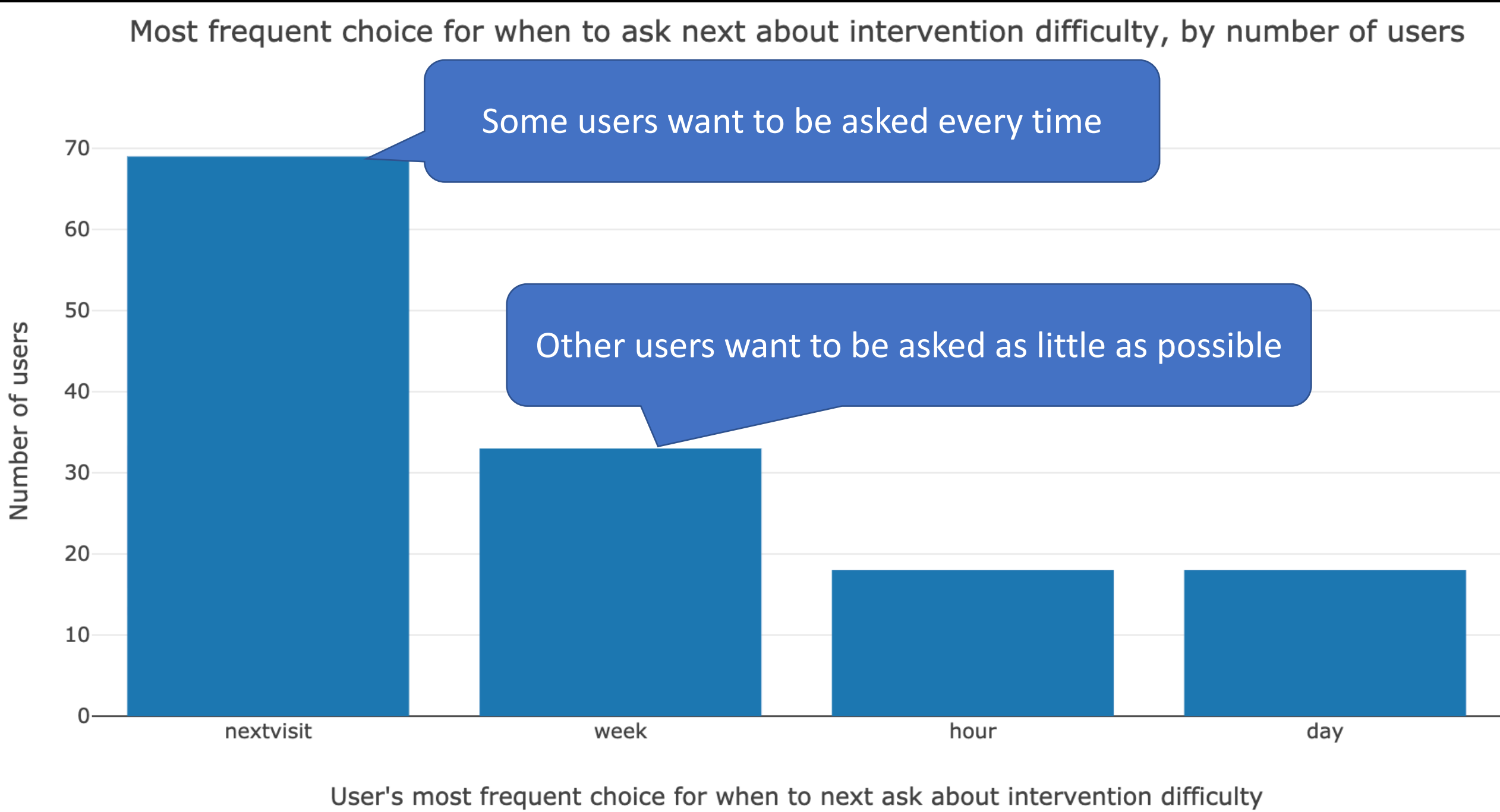
What do users actually want?

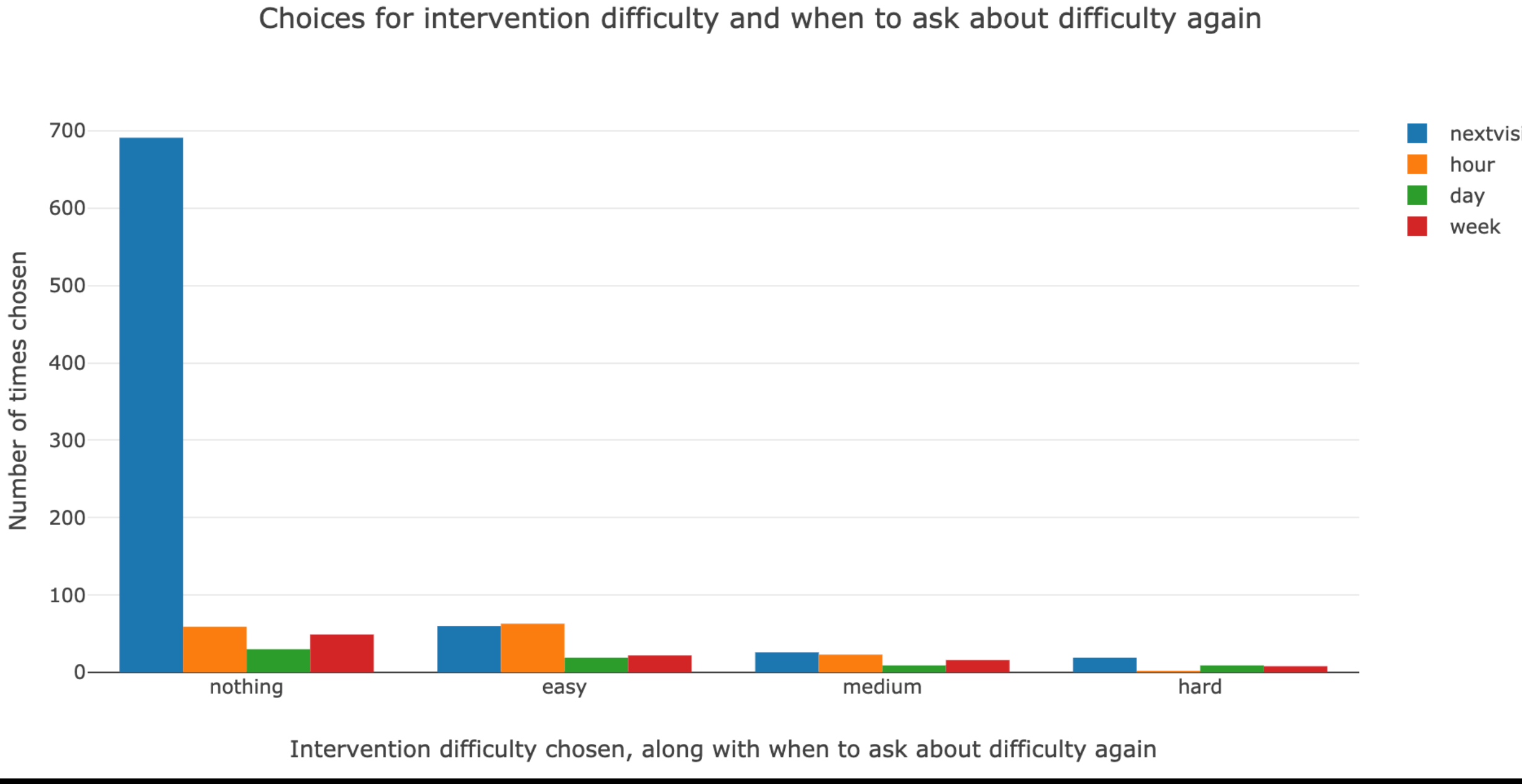

How aggressive would you like HabitLab to be in helping you reduce your time online?
Don't do anything
JUST TRACK TIME
Light touch
E.G., SHOW A TIMER WHEN YOU VISIT FACEBOOK
Medium
E.G., REMOVE YOUR FACEBOOK FEED UNTIL YOU CLICK TO SHOW IT
Heavy handed
E.G., CLOSE THE SITE AFTER 60 SECONDS

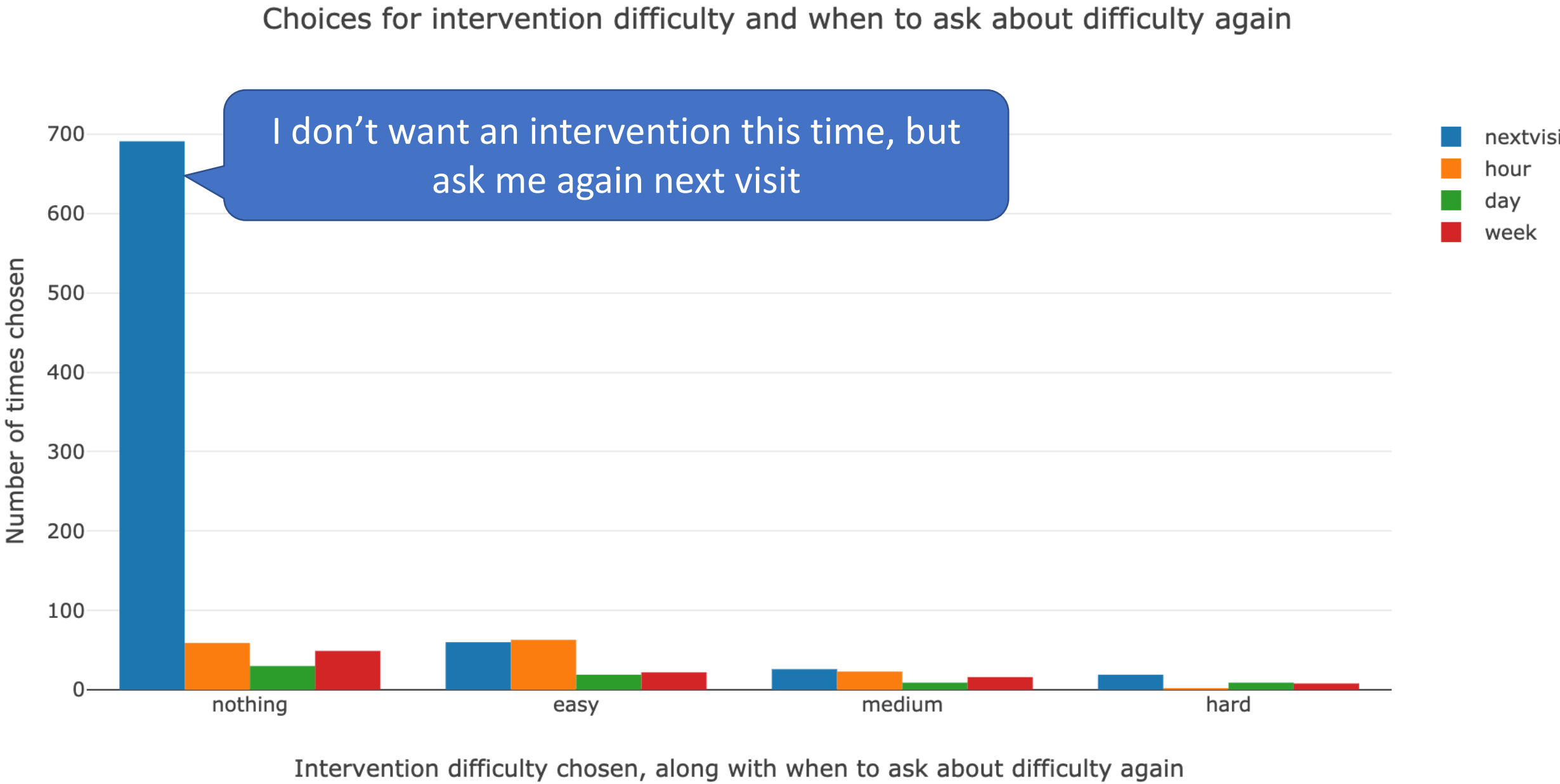

Ask me again about difficulty:
Next visit
1 hour later
1 day later
1 week later


Most frequent choice for when to ask next about intervention difficulty, by number of users
Number of users
70
60
50
40
30
20
10
0
nextvisit
week
hour
day
User's most frequent choice for when to next ask about intervention difficulty


Most frequent choice for when to ask next about intervention difficulty, by number of users
Some users want to be asked every time
Number of users
70
60
50
40
30
20
10
0
nextvisit
week
hour
day
User's most frequent choice for when to next ask about intervention difficulty


Most frequent choice for when to ask next about intervention difficulty, by number of users
Some users want to be asked every time
Other users want to be asked as little as possible
Number of users
70
60
50
40
30
20
10
0
nextvisit
week
hour
day
User's most frequent choice for when to next ask about intervention difficulty


Choices for intervention difficulty and when to ask about difficulty again
Number of times chosen
700
600
500
400
300
200
100
0
nothing
easy
medium
hard
Intervention difficulty chosen, along with when to ask about difficulty again
nextvisit
hour
day
week


Choices for intervention difficulty and when to ask about difficulty again
I don’t want an intervention this time, but ask me again next visit
nextvisit
hour
day
week
Number of times chosen
700
600
500
400
300
200
100
0
nothing
easy
medium
hard
Intervention difficulty chosen, along with when to ask about difficulty again

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

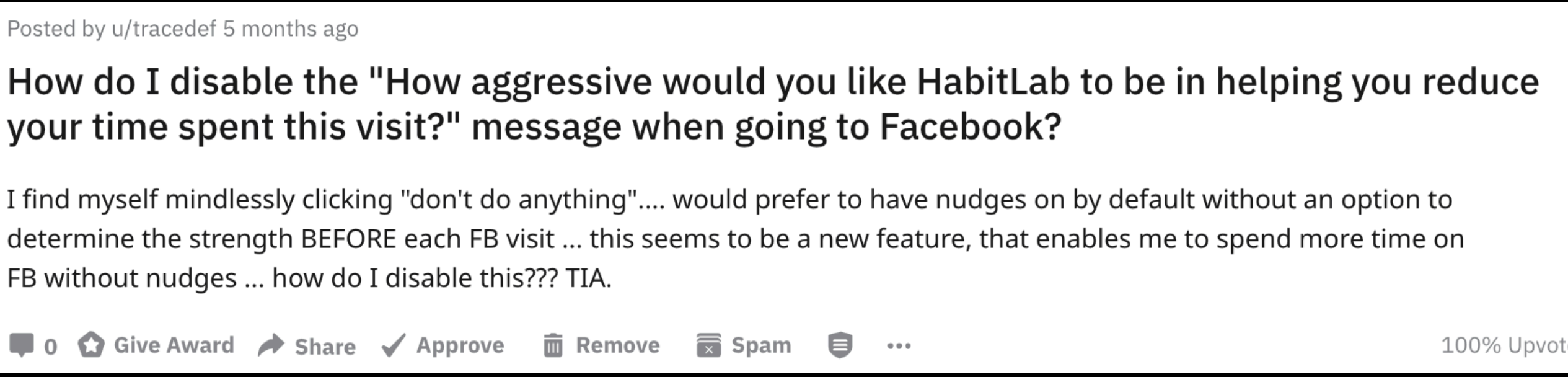
What are the side effects of interventions?

How do users' preferences change over time?

- Users initially are overly optimistic about how difficult they want their interventions
  - "No intervention" is least frequently chosen during onboarding
- Intervention difficulty choices decline over time
  - Half of users eventually choose "no intervention" nearly always
- Desired difficulty can be predicted with periodic experience sampling
- **Hope springs eternal: Don't want an intervention this visit, but ask next time**

People initially overestimate their motivation

Give people choices and many will eventually gravitate to the easy path



Posted by u/tracedef 5 months ago

How do I disable the "How aggressive would you like HabitLab to be in helping you reduce your time spent this visit?" message when going to Facebook?

I find myself mindlessly clicking "don't do anything".... would prefer to have nudges on by default without an option to determine the strength BEFORE each FB visit ... this seems to be a new feature, that enables me to spend more time on FB without nudges ... how do I disable this??? TIA.

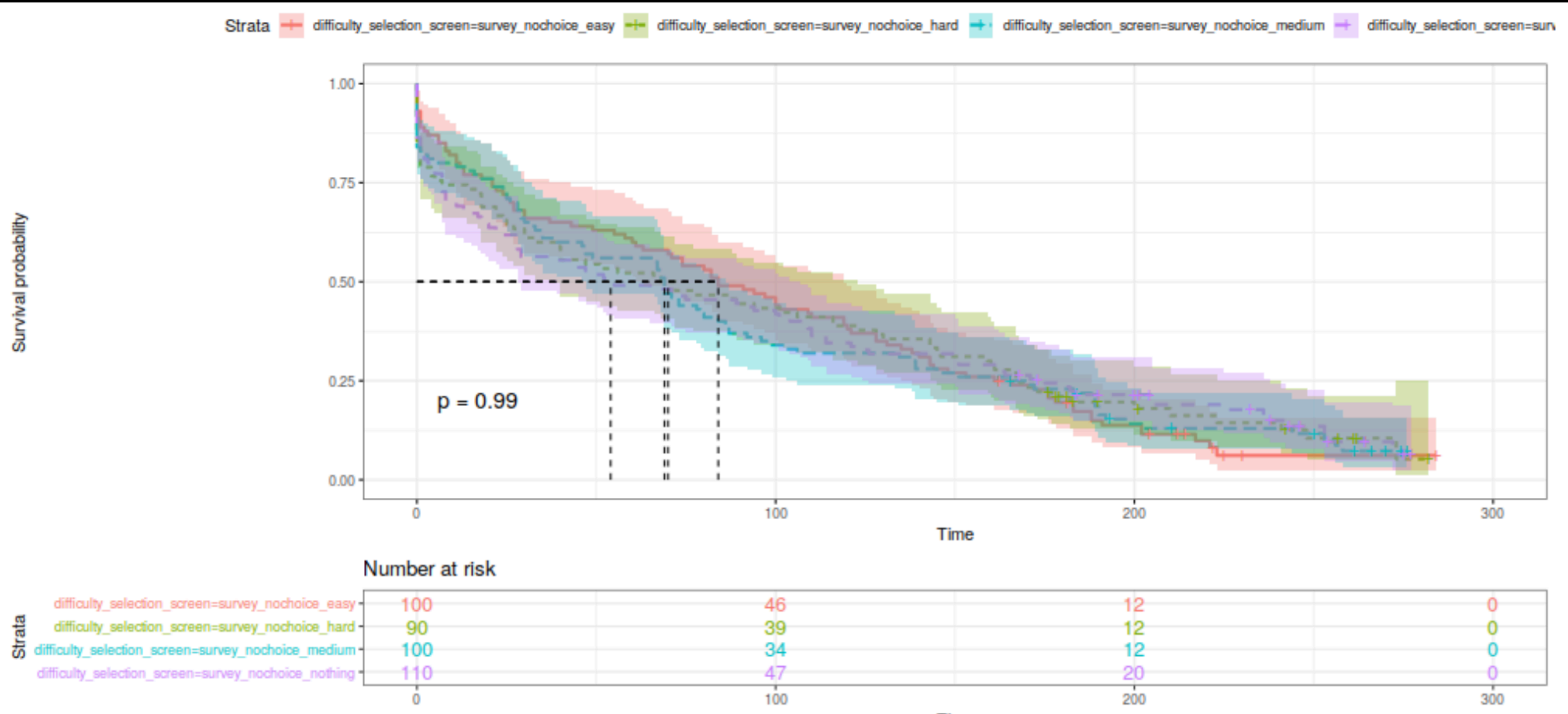
0 Give Award Share Approve Remove Spam ••• 100% Upvote

What happens when we remove choices?

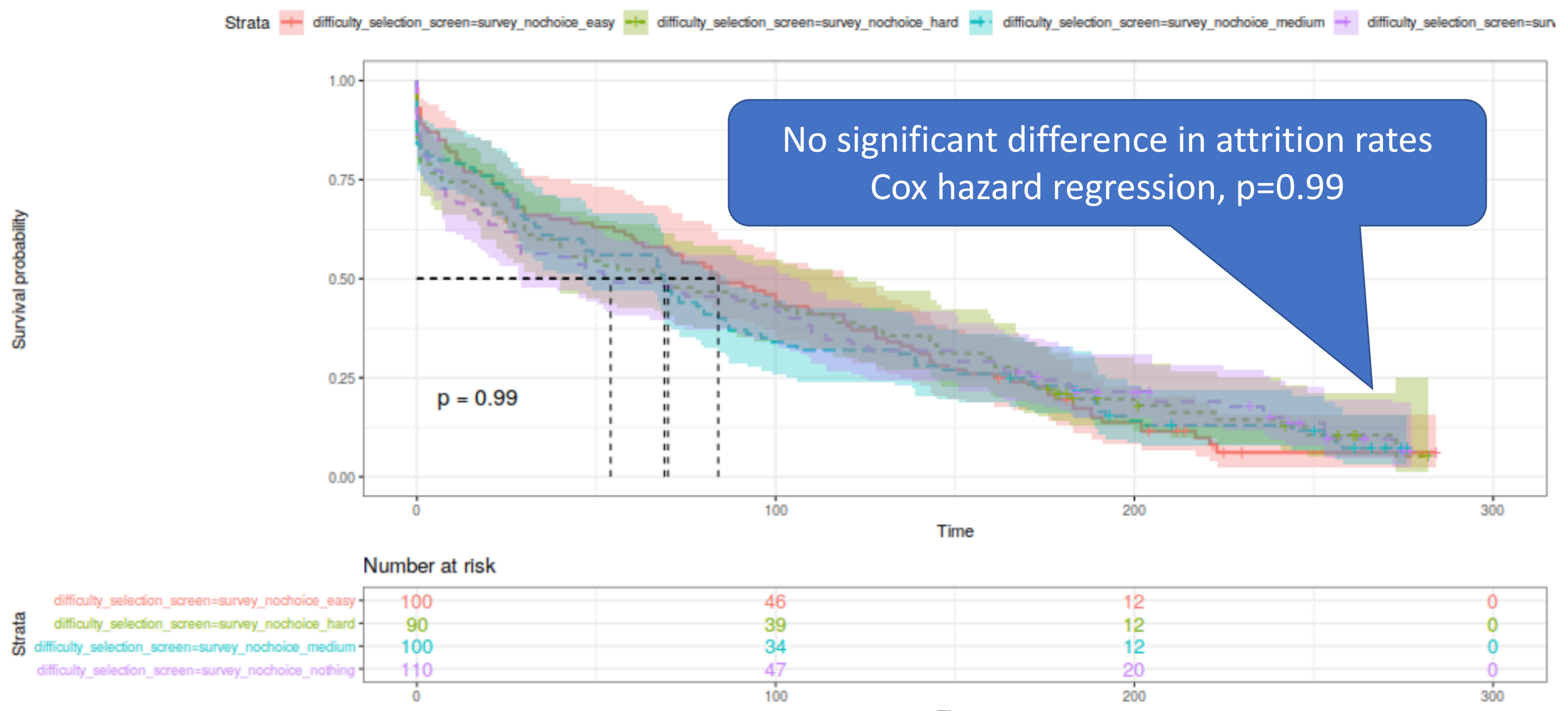
Asked users for their initial preferences during onboarding

Ignored their preferences and assigned them randomly to various intervention difficulty levels

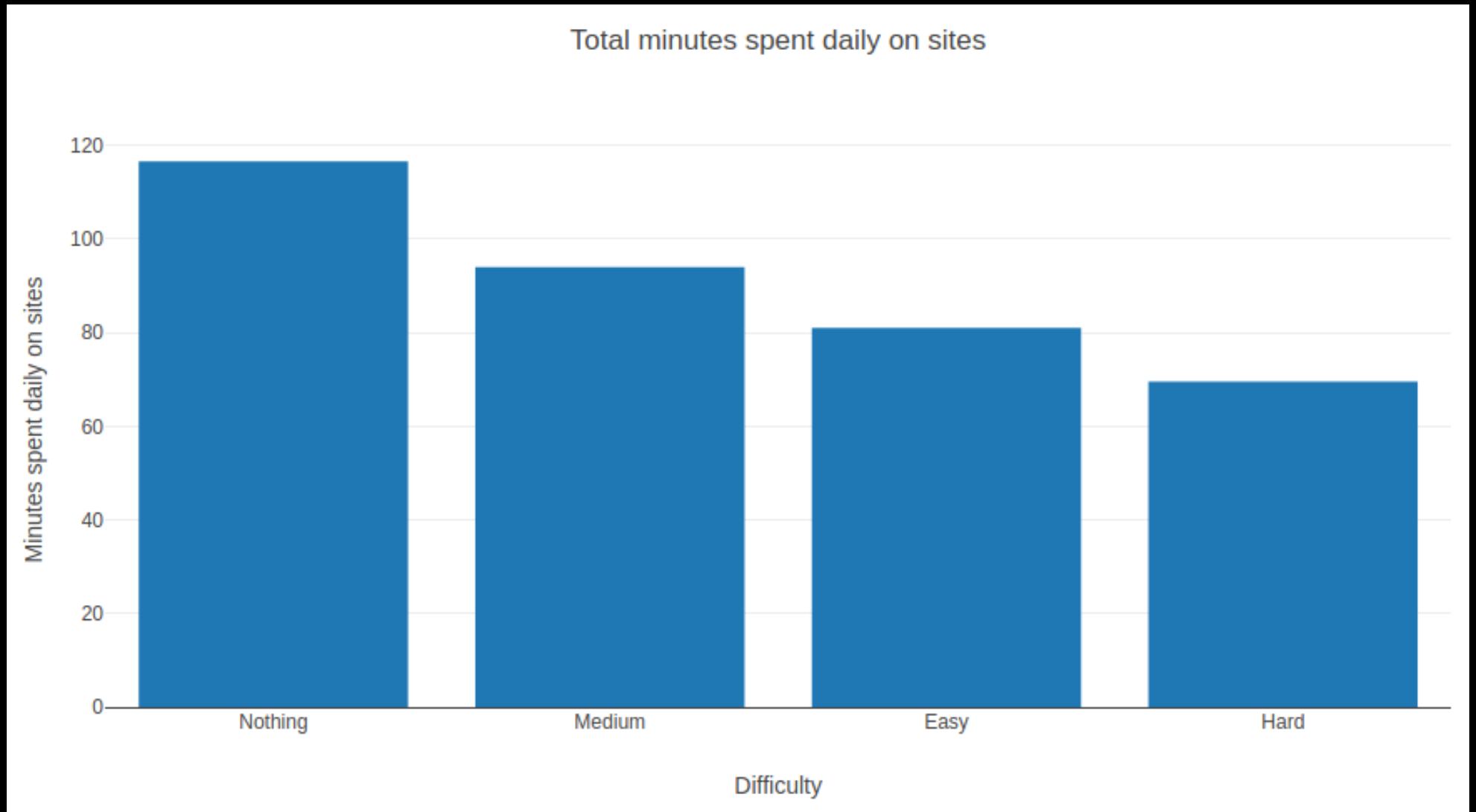
Randomly assigning difficulty levels has no effect on attrition



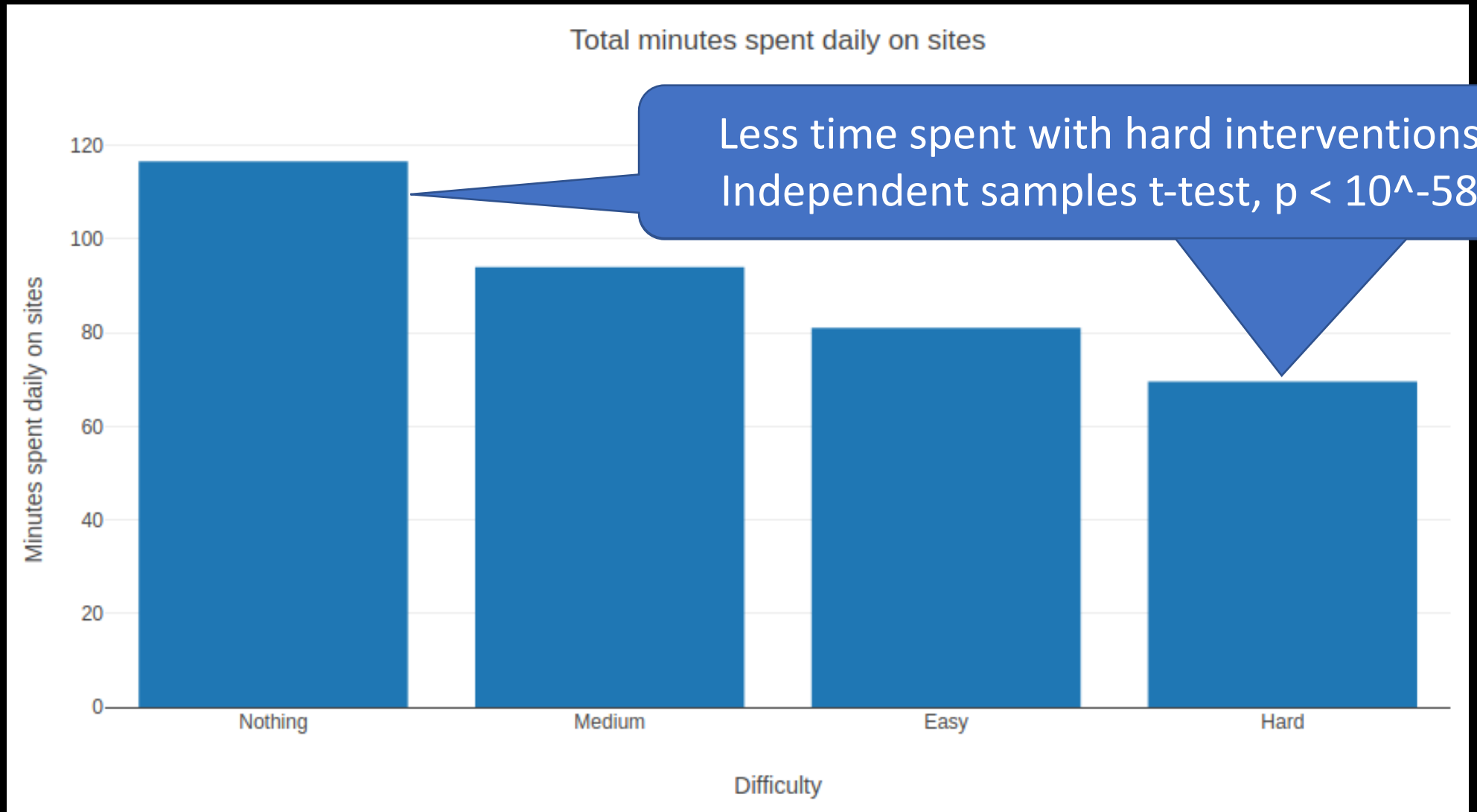
Randomly assigning difficulty levels has no effect on attrition



Randomly assigning difficulty levels improves the efficacy of interventions



Randomly assigning difficulty levels improves the efficacy of interventions



HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

- Users initially are overly optimistic about how difficult they want their interventions
  - "No intervention" is least frequently chosen during onboarding
- Intervention difficulty choices decline over time
  - Half of users eventually choose "no intervention" nearly always
- Desired difficulty can be predicted with periodic experience sampling
- Hope springs eternal: Don't want an intervention this visit, but ask next time
- **Assigning users harder interventions works to combat these issues**

Discussion

Randomly giving users harder interventions seems to perform better than giving users a choice

We would still prefer to respect user preferences

Discussion

Randomly giving users harder interventions seems to perform better than giving users a choice

We would still prefer to respect user preferences

Perhaps we can solve this issue by changing the set of choices?

Discussion

Randomly giving users harder interventions seems to perform better than giving users a choice

We would still prefer to respect user preferences

Perhaps we can solve this issue by changing the set of choices?

- Instead of giving a choice between fries or salad, make the choice between broccoli or spinach?

Discussion

Randomly giving users harder interventions seems to perform better than giving users a choice

We would still prefer to respect user preferences

Perhaps we can solve this issue by changing the set of choices?

- Instead of giving a choice between fries or salad, make the choice between broccoli or spinach?
- Instead of giving a choice to not see an intervention, make the choice between intervention A or B?

Project summary

Users differ in preferences for intervention difficulty

Users initially overestimate their motivation

Give users choices and many will eventually gravitate to the easy path

Giving them harder interventions without asking improves outcomes

Future work: Give users choices without adverse effects

HabitLab: Our in-the-wild behavior change platform

Do interventions remain effective as time passes?

What are the side effects of interventions?

How do users' preferences change over time?

Conclusion

Discussion and Implications

Behavior change systems' effects on users change constantly

- Users' preferences during onboarding may not be true a week later
- Initial observations of effectiveness are subject to novelty effects

There are secondary effects in addition to the targeted outcomes

- By reducing time on aggregator sites, time elsewhere is also reduced

Recommendations

Behavior change systems' effects on users change constantly

- Periodically do non-intrusive experience sampling and changes if possible. Don't assume everything during onboarding will be true forever

There are secondary effects in addition to the targeted outcomes

- When measuring how well your behavior change system works, measure outcomes holistically in addition to the target behavior

Future work

There are many behavior change taxonomies organizing theories for how behavior change interventions can work

Which of these theories actually work? How much does theory matter, as opposed to the implementation?

Implement interventions covering a taxonomy (90 total), and measure intervention effectiveness and attrition for each

Future work (more distant future)

We have focused on online behavior change

With the increasing ubiquity of sensors and wearables, could we build an in-the-wild behavior experimentation platform in the physical world?

Thesis Statement

In-the-wild experimentation is a powerful tool to gain insights about behavior change systems at scale – specifically, allowing us to conduct a wide range of studies about interventions and their outcomes.

Acknowledgements

Advisor: Michael Bernstein

Many contributors: Zhengxuan Wu, Jianghezi Zheng, Lisa Liao, Helen Qiu, Drew Gregory, Zilin Ma, Golrokh Emami, Jacob Ray, Matthieu Rolfo, Sarah Sukardi, Matthew Mistele, Julie Ju Young Kim, Wenqin Chen, Radha Jain, James Carroll, Sara Valderrama, Catherine Xu, Esteban Rey, Lewin Cary, Carmelle Millar, Colin Gaffney, Swathi Iyer, Sarah Tieu, Danna Xue, Britni Olina Chau, Na He Jeon, Armando Banuelos, Kaylie Zhu, Brahm Capoor

Labmates in the HCI group

HabitLab: In-the-wild Behavior Change Experiments at Scale

Outcomes

Does effectiveness remain constant as time passes?

Effectiveness of static interventions falls over time, rotating them helps

What externalities exist?

On browsers, reducing time on one site leads to reductions elsewhere

Interventions

How do users' preferences change over time?

Users initially choose harder interventions, but choice of difficulty falls over time

Backup slides

Examples of research questions we can potentially study using HabitLab

Question type	Questions related to goals of users	Questions related to choice of intervention	Questions related to outcomes of interventions
Observed state	What goals to users have?	What interventions do users choose?	What interventions are effective?
Changes	How and why do user goals change over time?	How and why do intervention choices change over time?	How and why does intervention effectiveness change over time?
Measures	How do users' stated goals differ from their actual goals?	How do users' stated intervention preferences differ from their actual choices?	How does effectiveness measured for an intervention differ from overall effect?

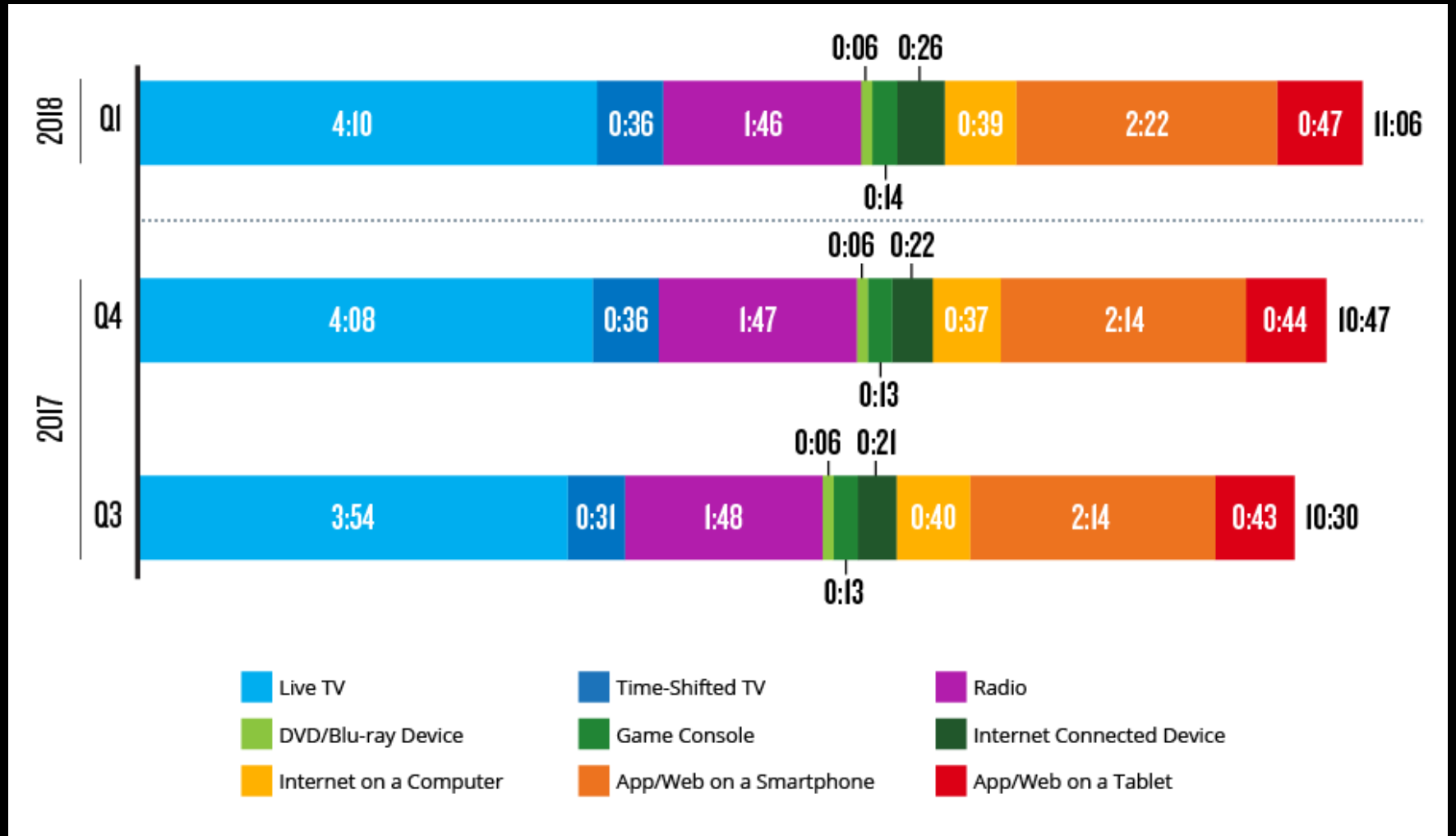
Why study online behavior change?

Can measure effectiveness (time spent per visit)

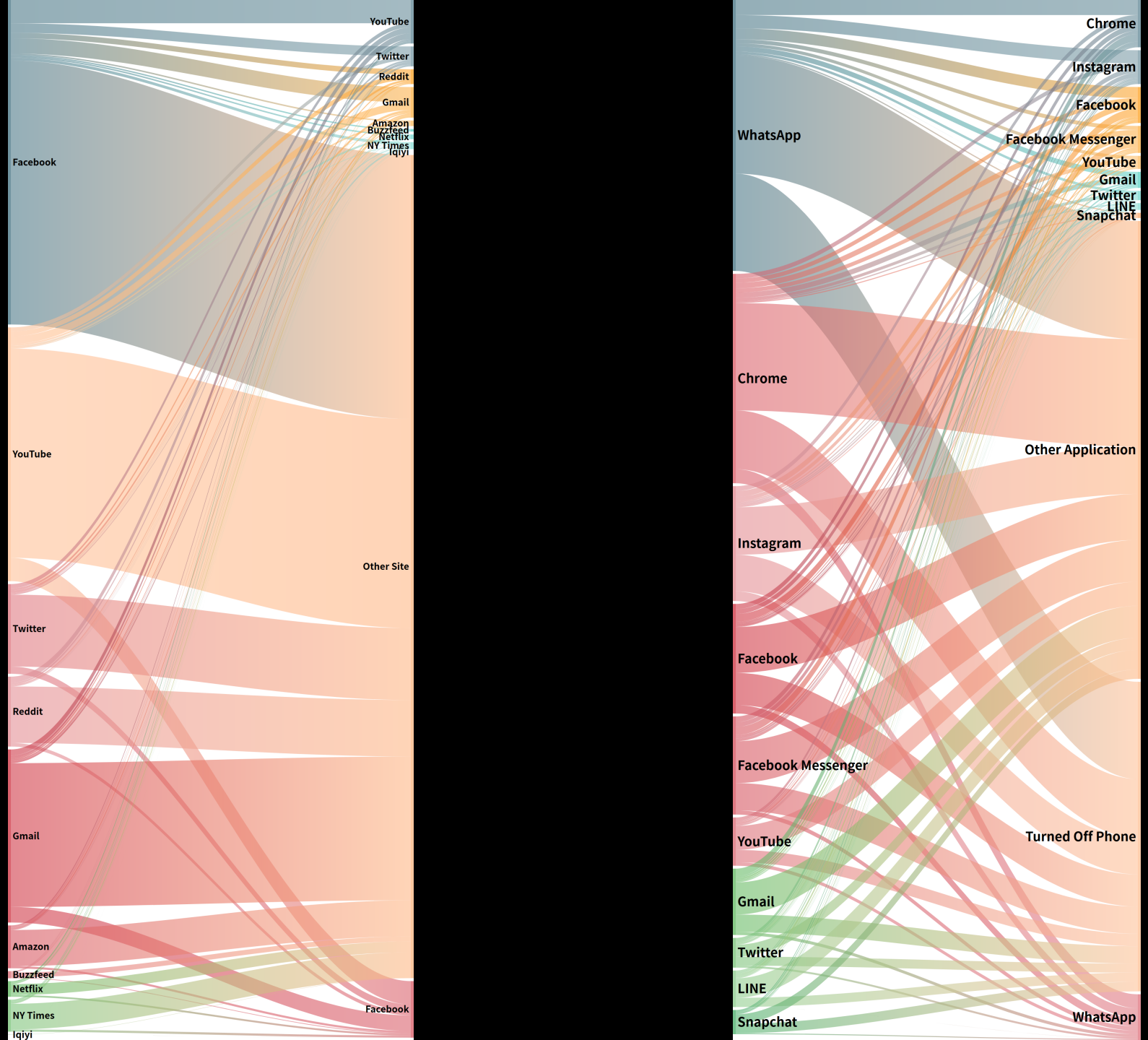
Many interventions possible (can modify sites)

Can frequently alter interventions (can change intervention every visit)

US adults spend 4.2 hours each day browsing the web and using phone apps



Q1 2018 Nielsen Total Audience Report


Facebook
YouTube
Twitter
Reddit
Gmail
Amazon
Buzzfeed
Netflix
NY Times
Iqiyi
YouTube
Twitter
Reddit
Gmail
Amazon
Buzzfeed
Netflix
NY Times
Iqiyi
Other Site
Facebook
WhatsApp
Chrome
Instagram
Facebook
Facebook Messenger
YouTube
Gmail
Twitter
LINE
Snapchat
Chrome
Instagram
Facebook
Facebook Messenger
YouTube
Gmail
Twitter
LINE
Snapchat
Other Application
Turned Off Phone
WhatsApp