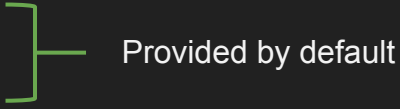


A/B Testing

Human in the Loop

Archit YADAV
Zakary BOUCHEMA

Agenda

- Objective
 - What is it?
 - Explanation
 - Case Study 1
 - Case Study 2
 - Other Case Studies
 - Personal Case Study
 - Benefits
 - Drawbacks
- Provided by default
- 

Objective

A/B Testing

What is it?

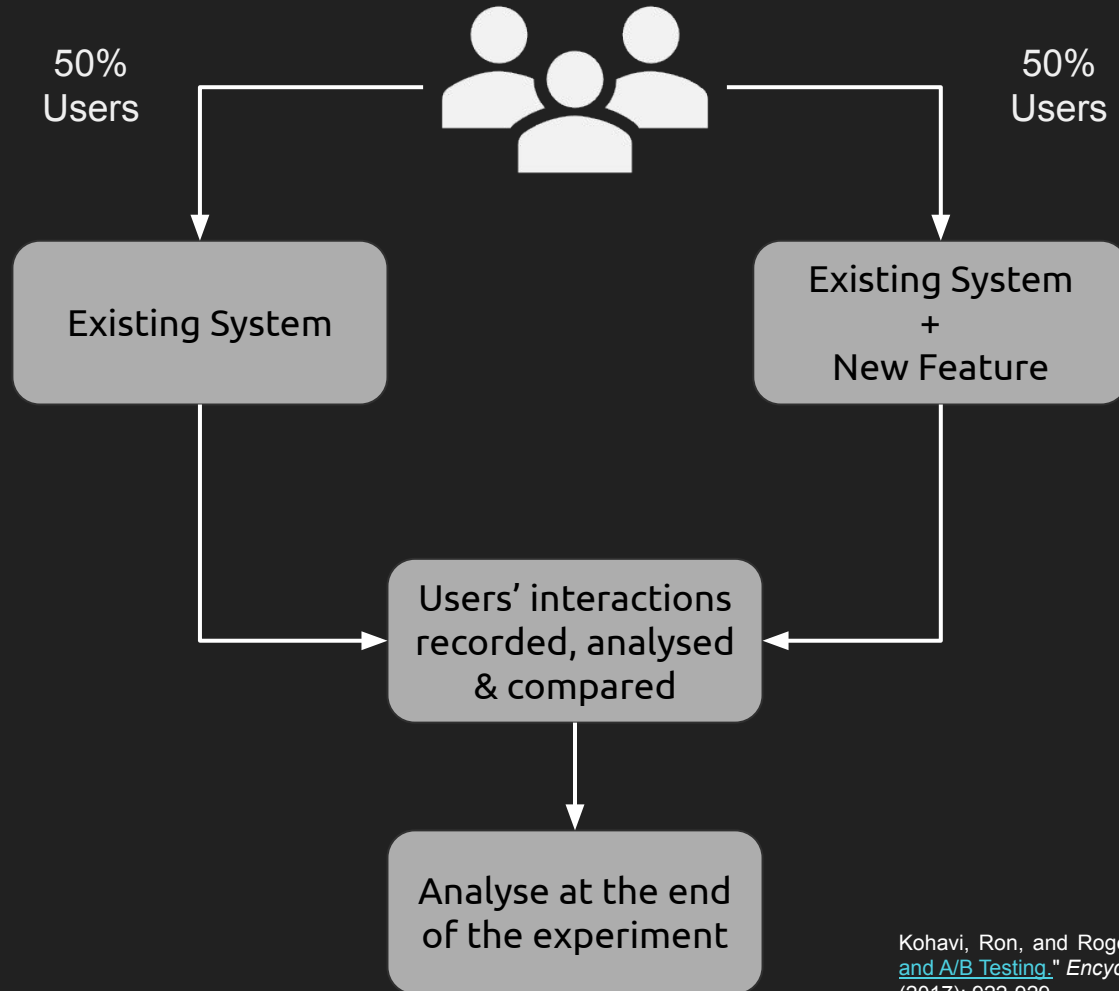
A/B Testing: What is it?

- Same experiment, different variants

Separate experiment!

A B





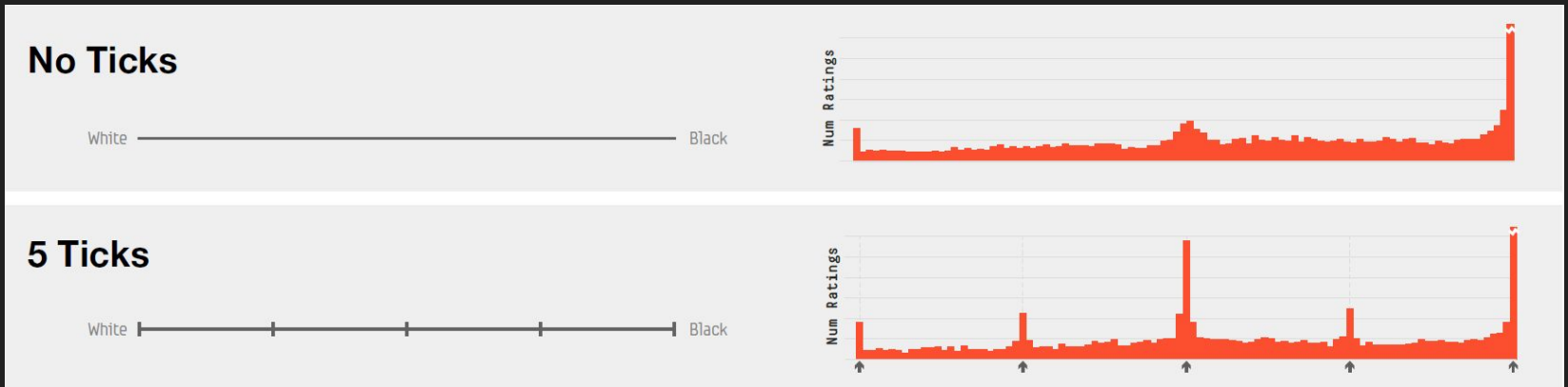
Explanation

Case Study 1:

Visual Analogue Scales and Continuous Sliders

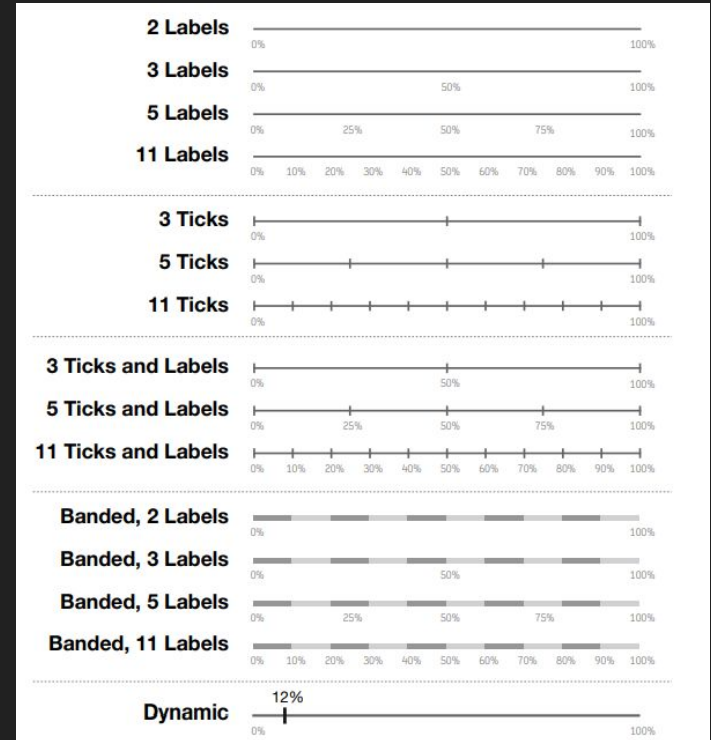
Case Study 1: VAS and Sliders

Question : Do decorations on scales have any effect on the values chosen ?



Case Study 1: VAS and Sliders

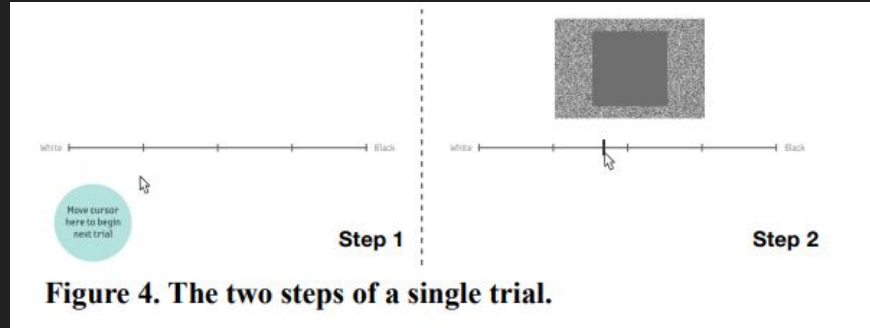
- Sliders
 - Shape of thumb
- Labels
 - Numbers of labels
- Ticks
 - Numbers of ticks
 - Visual weight
 - Alternating major/minor tick
- Labels x Ticks
- Labels x Banded



Case Study 1: VAS and Sliders

Experimental design:

- Put cursor on 'start position'
- Display stimulus and wait user input
- Rinse and repeat

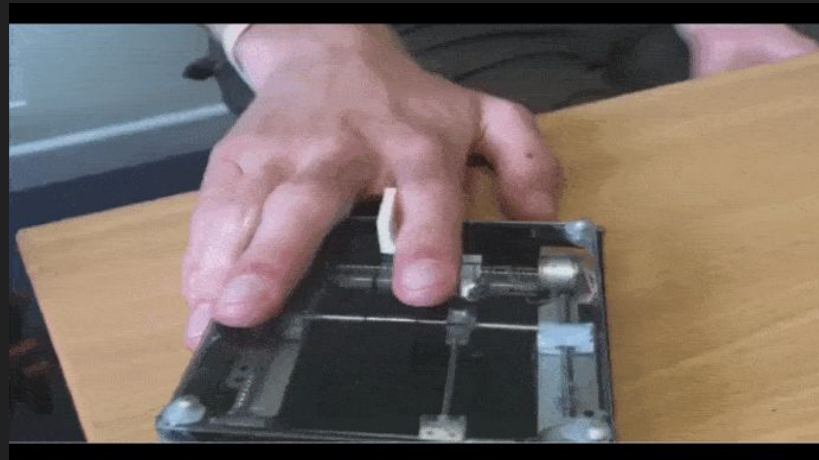
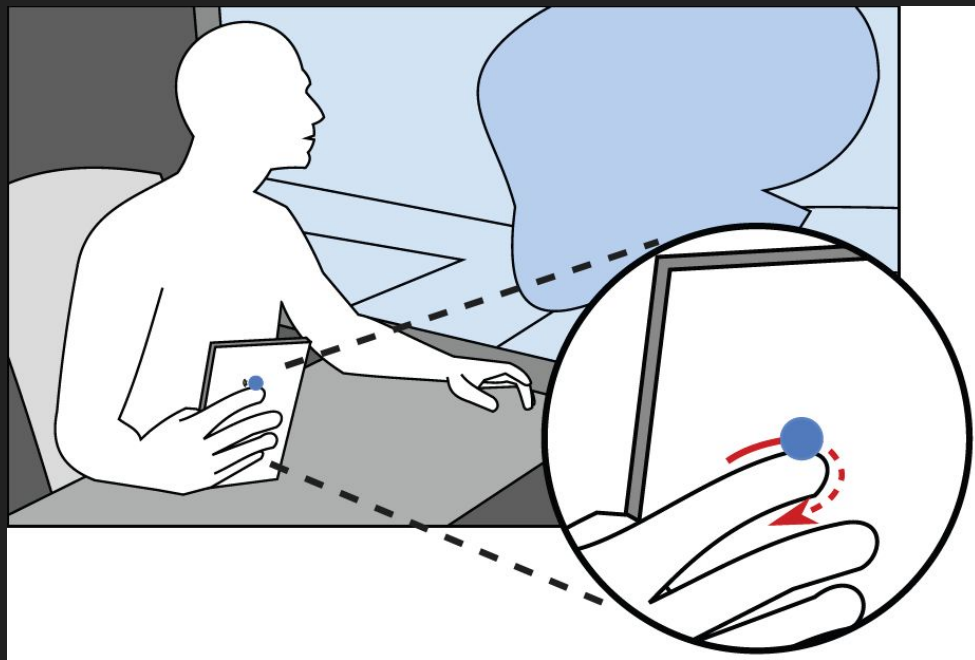


Results: In short, yes, the decorations do have an effect on the judgemental task

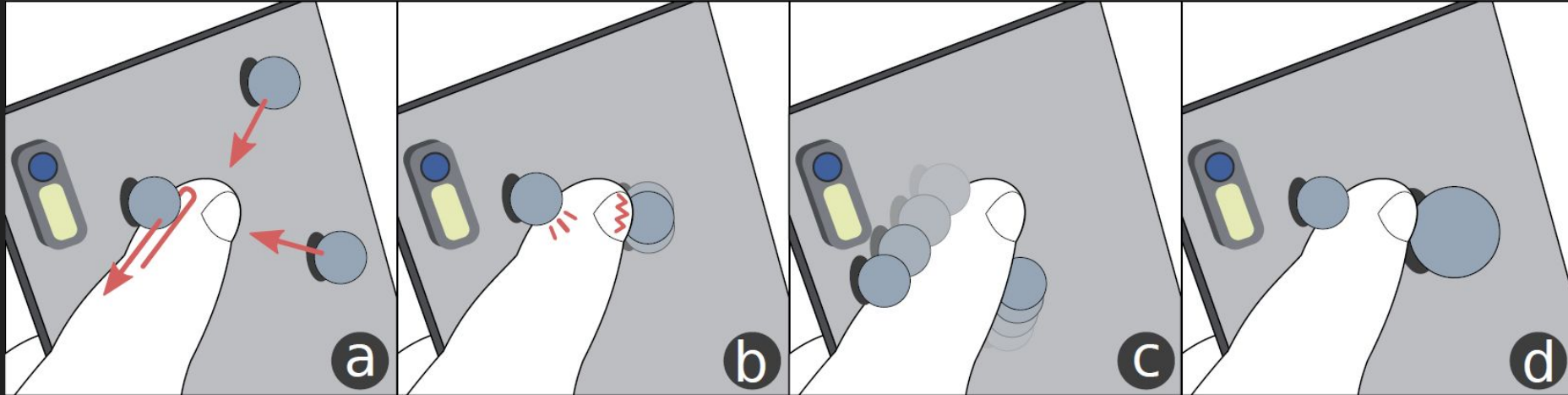
Case Study 2:

Pulp Friction

Case Study 2: Pulp Friction



Case Study 2: Pulp Friction



- Angle
- Pattern

- Poke
- Vibration

- Speed

- Size

Case Study 2: Pulp Friction

- Angle
- Type
 - Poke
 - Vibration
- Pattern
 - CW
 - ACW
- Speed
- Size
- Surface
- Strength
- Comfort

Factors

(Not considered)

Ability to recognise:	
Localized Contact Point	Patterns
Angle	Pattern
Type	Speed
	Size

A

50%
Users

B

50%
Users

Case Study 3:

A/B Testing Challenges in Large Scale Social Networks

Case Study 3: A/B Testing Challenges in Large Scale Social Networks

- Platform into consideration - LinkedIn
- > 400 concurrent A/B testings / day
 - Affects sales and interaction of users



What cause(s) do you care about? ×

Add more color to your professional identity by showing what you care about.

<input type="checkbox"/> Animal Welfare	<input type="checkbox"/> Environment
<input type="checkbox"/> Arts and Culture	<input type="checkbox"/> Health
<input type="checkbox"/> Children	<input type="checkbox"/> Human Rights
<input type="checkbox"/> Civil Rights and Social Action	<input type="checkbox"/> Politics



14 % increase
in edits

1 Confirm your billing cycle

Monthly
30-day free trial
then US\$59.99 per month

Annual
Save 20%
US\$575.88 per year

Monthly	US\$59.99
30-day free trial	(US\$59.99)
<hr/>	
Subtotal	US\$0.00
Estimated tax (0.00%)	US\$0.00
<hr/>	
Today's total	US\$0.00

2 Select your payment method

[Why do we need this for a free trial?](#)



Frequently asked questions ▾

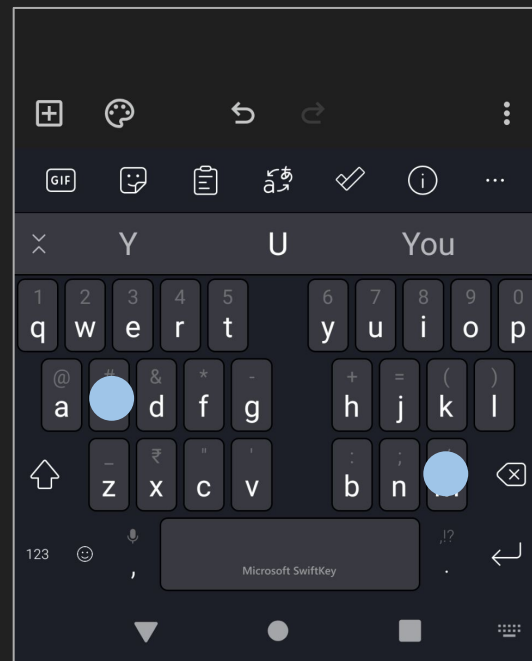
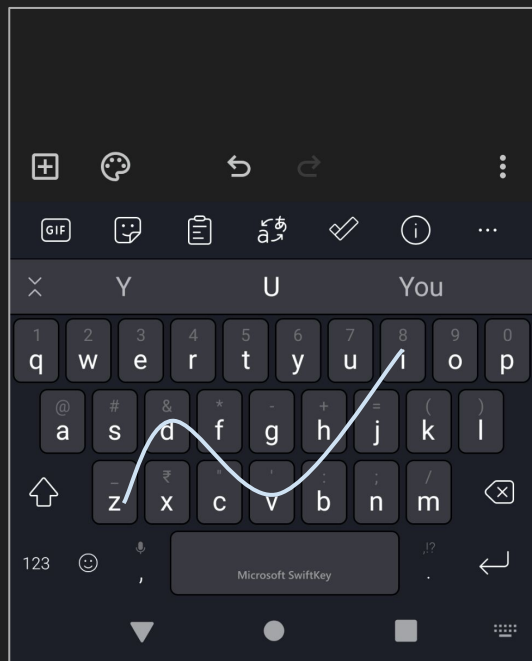
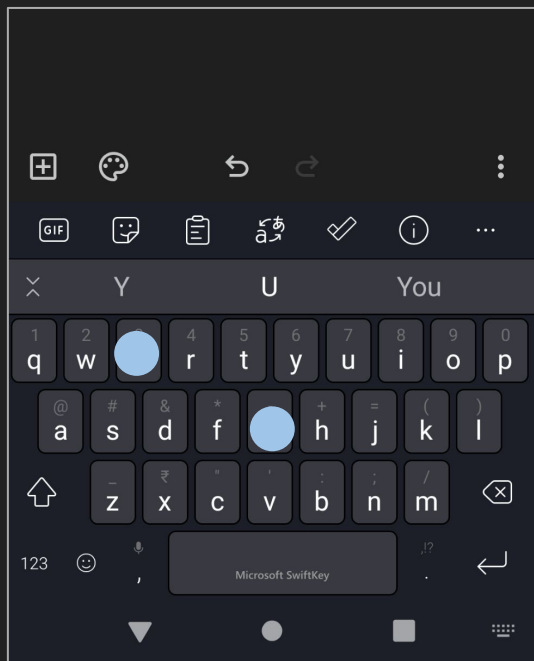
- Redesign of the payment flow
- > 10% ↑ in free trial orders
- 30% ↓ on refund orders
 - Users better understood the premium plans
- Resulted in increasing earnings as well as UX

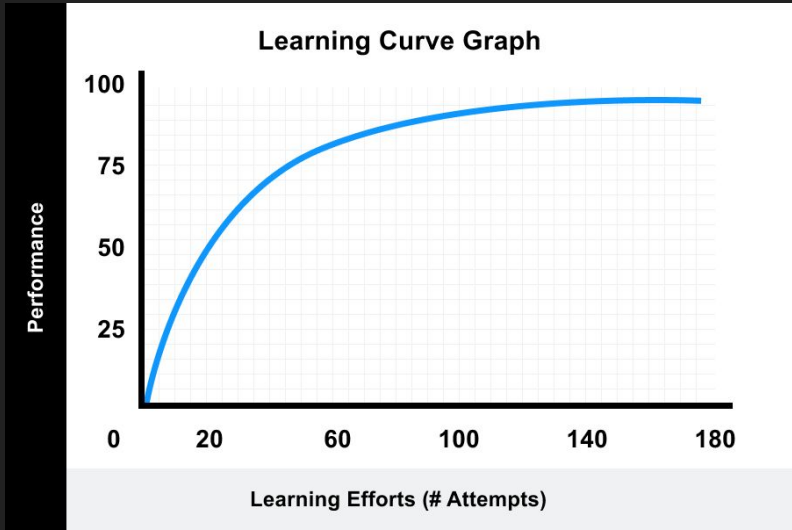
Personal (Ongoing) Case Study

Effect of Various Keyboards on Text Input

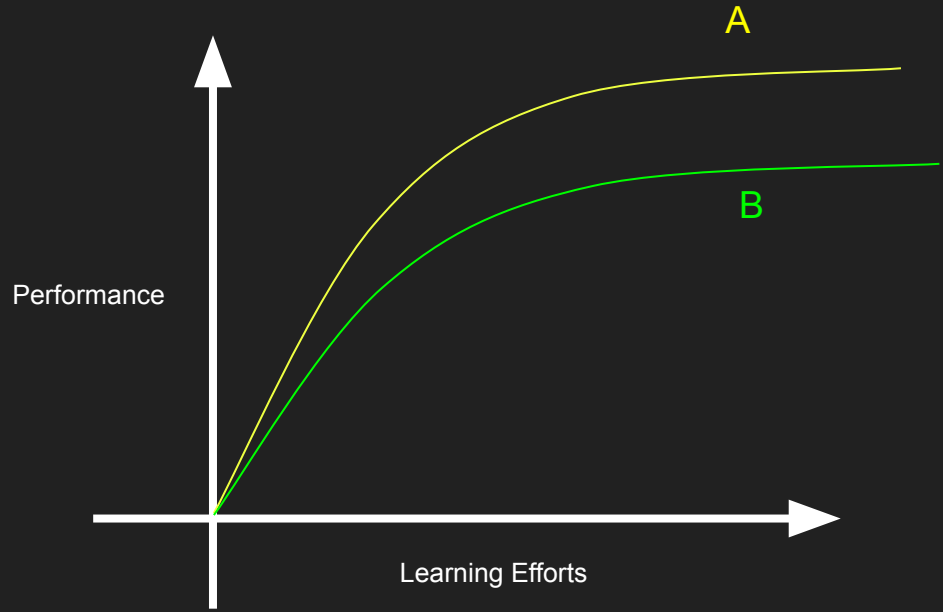
Personal Case Study: Effect of Various Keyboards on Text Input

- User Bandwidth / Performance Evaluation



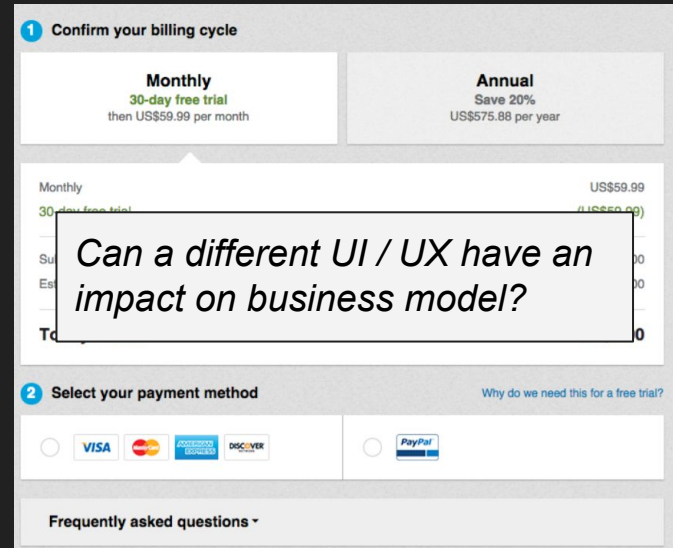
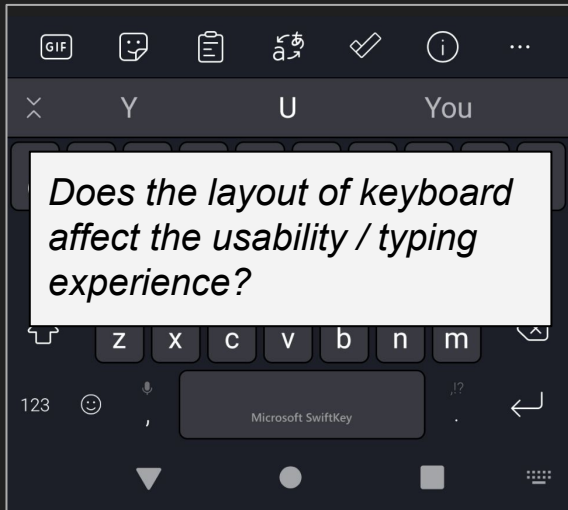
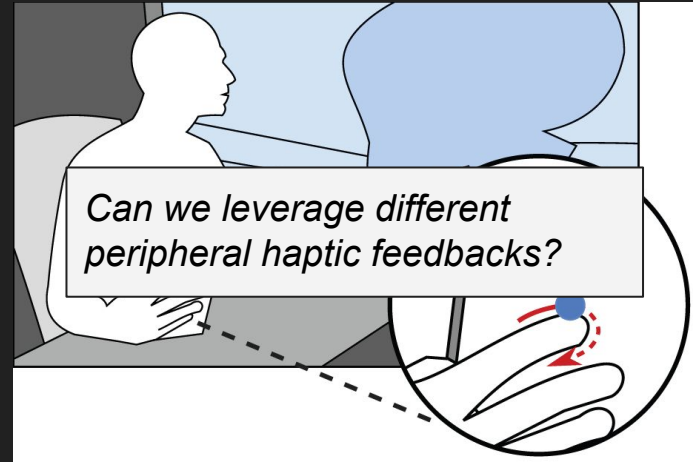
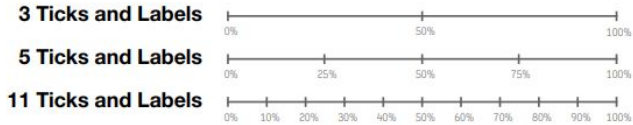


<https://www.valamis.com/documents/10197/520324/learning-curve.png>





Do Different sliders have an impact on the output?



Benefits

Benefits

- Test new ideas

- Get feedback



- Incremental improvement

- Addition of small new feature



- Distributed improvement

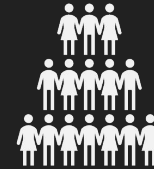
- Doesn't affect all in case of failure



Drawbacks

Drawbacks

- Only answers the '*what*', not '*why*'
- Problematic in case of wrong group targeted
 - Bias in results
 - May lead to disgruntled users
- Time and resource intensive



References

- Kohavi R., Longbotham R. (2017) *Online Controlled Experiments and A/B Testing*. In: Sammut C., Webb G.I. (eds) *Encyclopedia of Machine Learning and Data Mining*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7687-1_891
- Justin Matejka, Michael Glueck, Tovi Grossman, and George Fitzmaurice. 2016. *The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales*. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5421–5432. DOI:<https://doi.org/10.1145/2858036.2858063>
- Alix Goguey, Deepak Ranjan Sahoo, Simon Robinson, Jennifer Pearson, and Matt Jones. 2019. *Pulp Friction: Exploring the Finger Pad Periphery for Subtle Haptic Feedback*. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Paper 641, 1–13. DOI:<https://doi.org/10.1145/3290605.3300871>
- Ya Xu, Nanyu Chen, Addrian Fernandez, Omar Sinno, and Anmol Bhasin. 2015. *From Infrastructure to Culture: A/B Testing Challenges in Large Scale Social Networks*. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*. Association for Computing Machinery, New York, NY, USA, 2227–2236. DOI:<https://doi.org/10.1145/2783258.2788602>

A/B Testing

Human in the Loop