Click ‘Rate Session’
Rate 10 sessions to get the supercool GOTO reward
Follow us on Twitter @GOTOber
www.gotober.com
building a culture of experimentation at Spotify

@bendressler - experimentation lead
user retention
product satisfaction
# of track plays
predicting things is hard
hi
why and how we experiment - in a consumer oriented live product where we aim to drive a defined set of KPIs.
expectations

building and evaluating
a/b testing basics
graduating towards experimentation
showcase
prologue
“and God said, ‘Let there be light,’ and there was light”
“God saw all that he had made. and it was very good”
a successful idea does what it’s supposed to do
ALMIGHTY BUSINESS = KPI
I use it on the subway

All my favorites are there

I don’t get it

IT'S AMAZING

It doesn’t support Linux, I’m out
I use it on the subway

All my favorites are there

I don’t get it

ITS AMAZING

It doesn’t support Linux, I’m out

ALMIGHTY BUSINESS KPI = X 100,000,000
I don’t get it
used to be an iTunes user

“what’s the difference between saving and adding to a playlist?”

struggles with finding his friend’s profile

doesn’t know there is a Premium version
91% feeling ‘super pumped’
so what else do you want?
problem #1
“omg, people who build tons of playlists stay with us forever!”

“let’s call the CEO!!”

“mandatory playlist creation for everyone!!!”

#hypotheticalscenario
playlisting → retention
playlisting ← retention
problem #2
enter: science
“I wonder why some people end up being super mean to others”
personality

experiences

etnicity

socio-economic status

genes

education

ideology

situational factors
situational factors

- personality
- experiences
- ethnicity
- socio-economic status
- genetics
- education
- ideology
“now punch those guys over there”
situational factors

personality

experiences

ethnicity

socio-economic status

genetics

education

ideology
population

a

(control)

observation a

difference = effect

b

(control + change)

observation b
enough participants

population

observation a

(control) a

observation b

(control + change) b

difference = effect

random assignment

enough participants

population

observation a

(control) a

observation b

(control + change) b

difference = effect

random assignment
Population

Observation a

(make changing things cheap & easy, QA)

Observation b

(control + change)

difference = effect
observation a

(population)

observation b

(reliable data, good logging, meaningful metrics)

difference = effect

(control + change)
population

observation a (control)

observation b (control + change)

a

observation a

b

observation b

difference = effect

statistical test
guarantee statistical assumptions
minimise delay from start to finish
handle many concurrent tests
minimise human error
“God looked at his conversion rate. and it was very good”
large buttons are swee-heet

blue links, no, underlined, no, red links

flashing text gets more attention

kittens. just kittens.
“A/B testing is inevitably reductive, darwinistically evolving the ‘fittest’ design ... it forces you to follow your audience, not lead them ... but it is best suited to niche-testing elements, not layouts.”

Martin Gittins
“an experiment is a means of gathering information to compare an idea against reality”.

from *Experiment!* by Colin McFarland
“God said ‘let our value proposition be clearer to new users’”
+ strong hypothesis
+ meaningful metrics
+ community
strong hypothesis
what. why. how. who.
meaningful metrics

who. when. where. what.
relevant and useful.
community
critique. share. repeat.
in practice
example
“God said ‘let our value proposition be clearer to new users’”
direction
epilogue
building

“and God said, ‘Let there be light,’”
“and God said, ‘Let there be light,’ - so the angels asked what his hypothesis was.”
thanks :)
further reading

microsoft (ronny kohavi)
booking.com (lukas vermeer, erin weigel)
skyscanner (colin mcfarland)
etsy (dan mckinley)
linkedin (ya xu)
riot games (jeffrey lin)
Please
Remember to rate session
Thank you!

Follow us on Twitter @GOTOber
www.gotober.com