



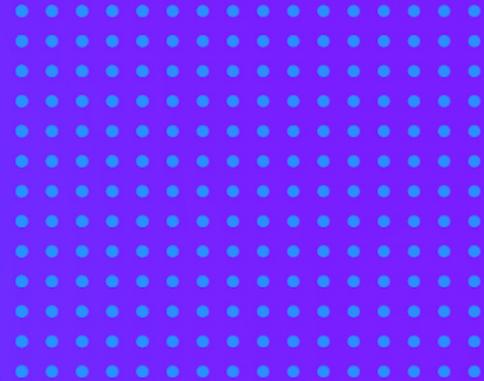
Reduce risk and fail faster:
experimenting across the product
development lifecycle

Gregory Sherwin

Sr Principal Engineer,
Farfetch

Isabel Meijaard

Sr Manager Strategy Consulting,
Optimizely





Gregory Sherwin

Sr Principal Engineer

FARFETCH



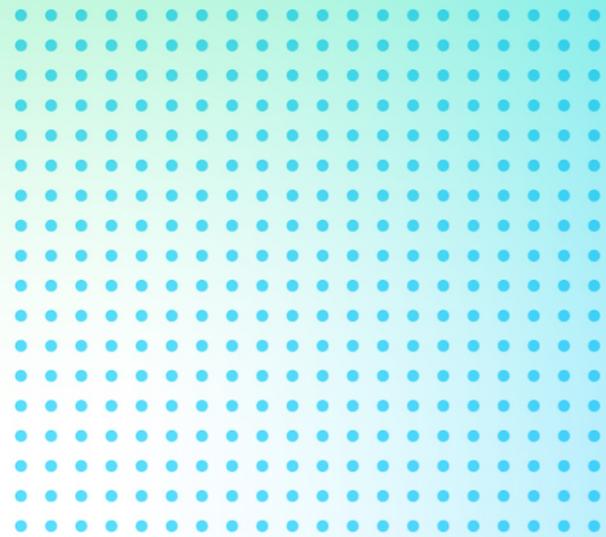
Isabel Meijaard

Sr Manager Strategy Consulting

 **Optimizely**

Agenda

- 1. Product Dev & Experimentation**
- 2. Determine the right direction**
- 3. Drive Continuous Optimization**

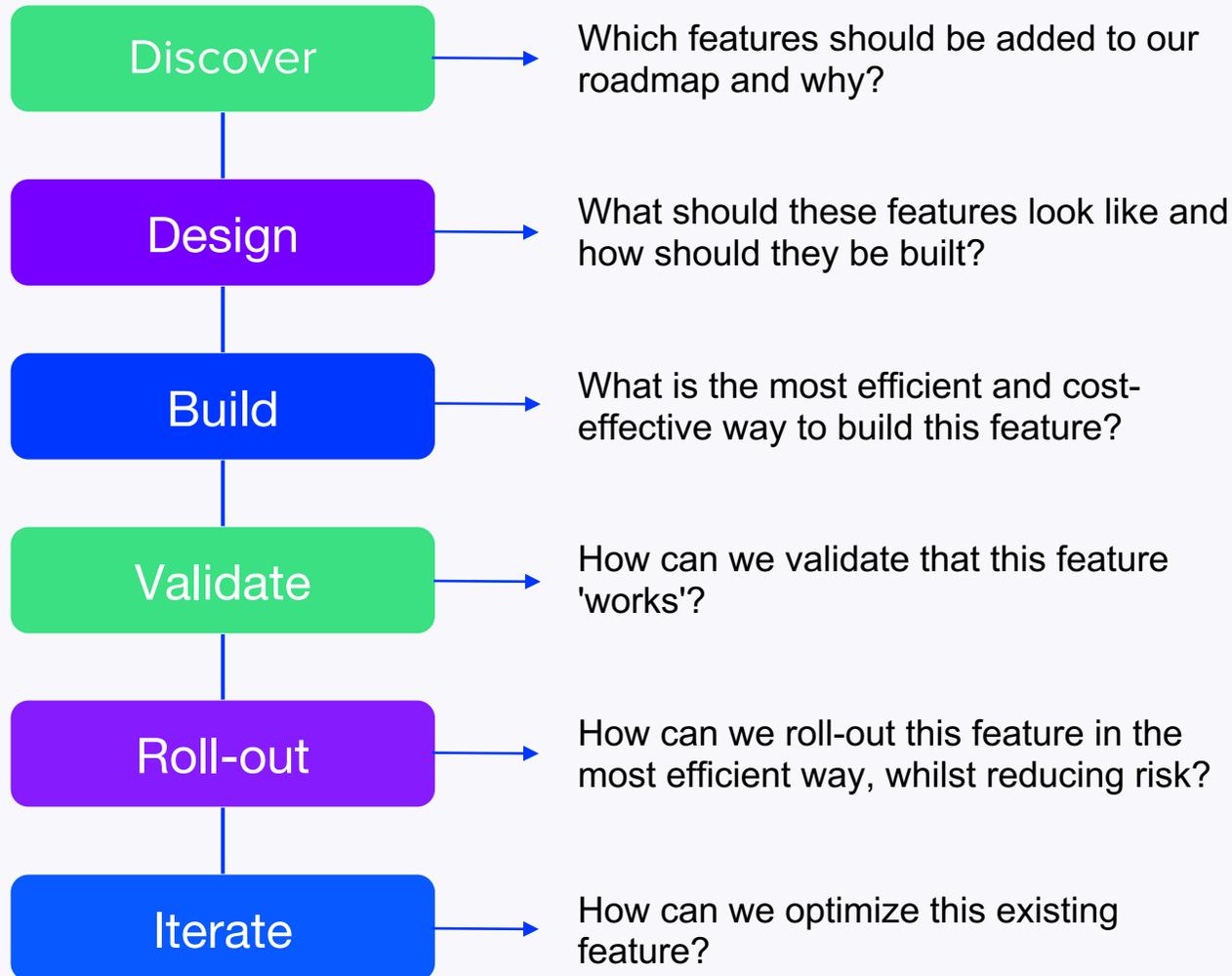


01

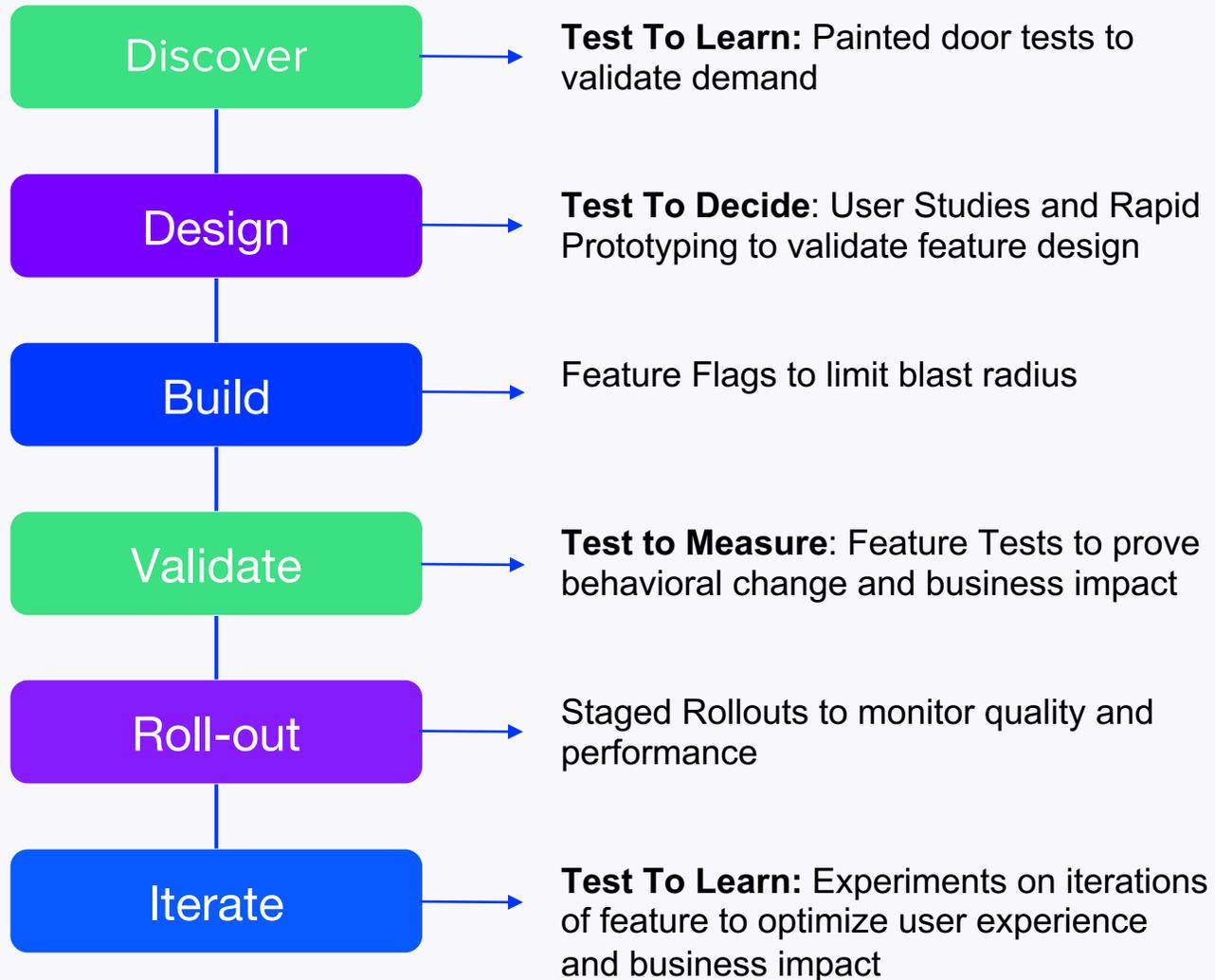
Introduction

**Product
Development &
Experimentation**

The Product Development Process



And Experimentation



Isabel



Three Product Test Types



Test To Learn

Learn about user demand and interest for new features. Learn *how* users respond to your feature and why they respond the way they do.



Test To Decide

Decide on feature design, build and roll-out strategy



Test To Measure

Validate impact of feature on user experience and/or key business metrics

Isabel

Greg

Common Challenges

Ship-first Mentality

Not asking enough questions of whether we're shipping the *right* thing.

Productivity Threat

Goodhart's Law challenge: when pushing changes becomes the goal vs. pushing only the *right* changes.

Org-wide Buy-in

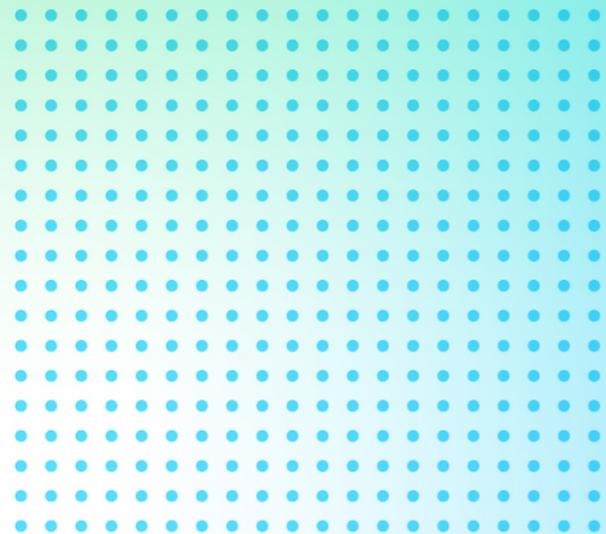
HIPPO decrees, customer & compliance demands, and inertia conflict with distributed decision-making.

Validation Only

Experimentation perceived as a necessary burden to "rubber stamp" a product or feature design decision.

Agenda

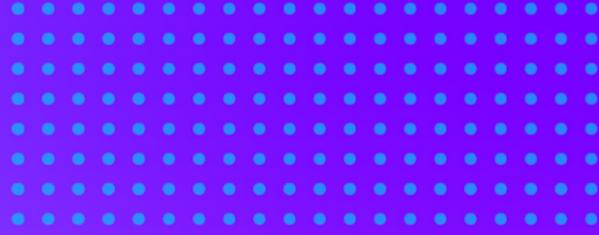
- 1. Product Dev & Experimentation**
- 2. Determine the right direction**
- 3. Drive Continuous Optimization**



02

Determine the right direction

Build the right 'thing'

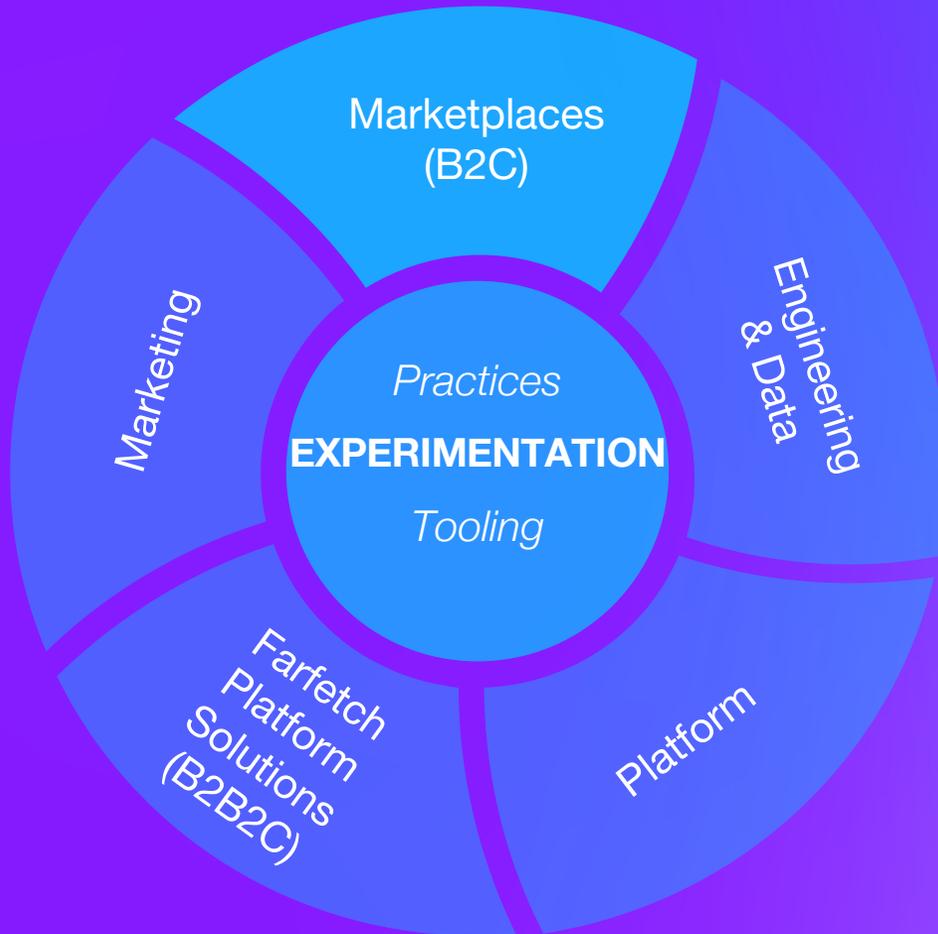


Build the right 'thing'
Build the right 'thing'
Build the right 'thing'
Build the right 'thing'
Build the right 'thing'

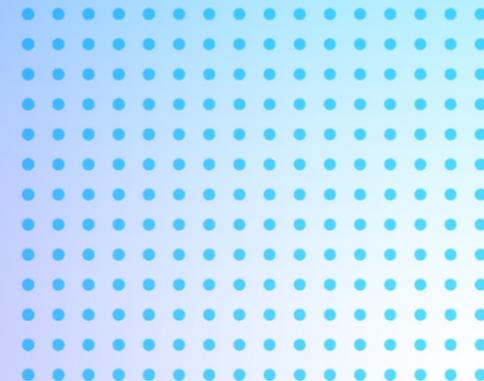


Experimentation mission

Greg



Provide the right tools and elevate standard practices to support experimentation across the company

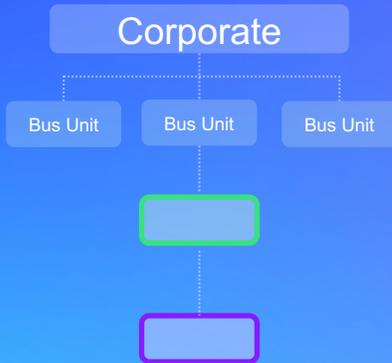


Models for program structures

Greg



Individual

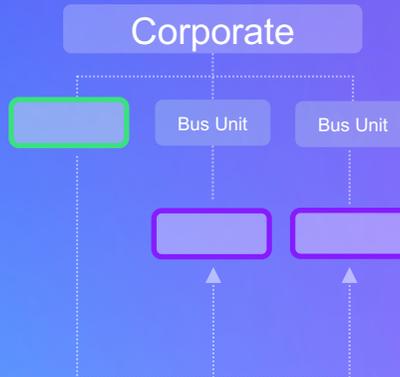


Single team and/or single site ownership

Small focus, simple tests, build momentum

Program Manager with access to Dev, Design, etc.

Centralized

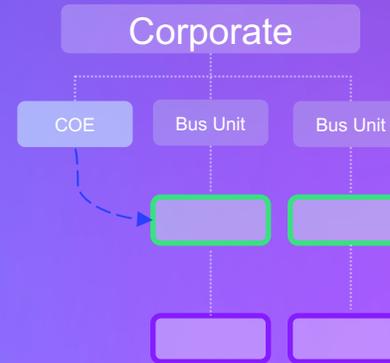


Internal consultants for other teams

Single team owns all experiment strategy/tools and launches all experiments

Program Manager with access to dev, design etc.

Center of excellence

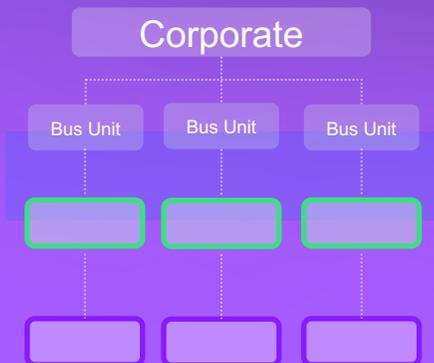


Central support team for business units who launch their own tests

CoE enables individual teams of separate sites/domains

Central Expert team support other business functions who want to start testing

Experimentation council



Committee planning across business units

Driven by a shared ownership of pages, kpis, traffic and prioritization

Testing council for product, marketing and UX/Creative

Definition

Characteristics

Team

Experimentation rituals and artifacts

Greg

**Experiment
definition
documents**

Hypothesis
Peer-Review
Sessions

Regular
Training
sessions &
clinics

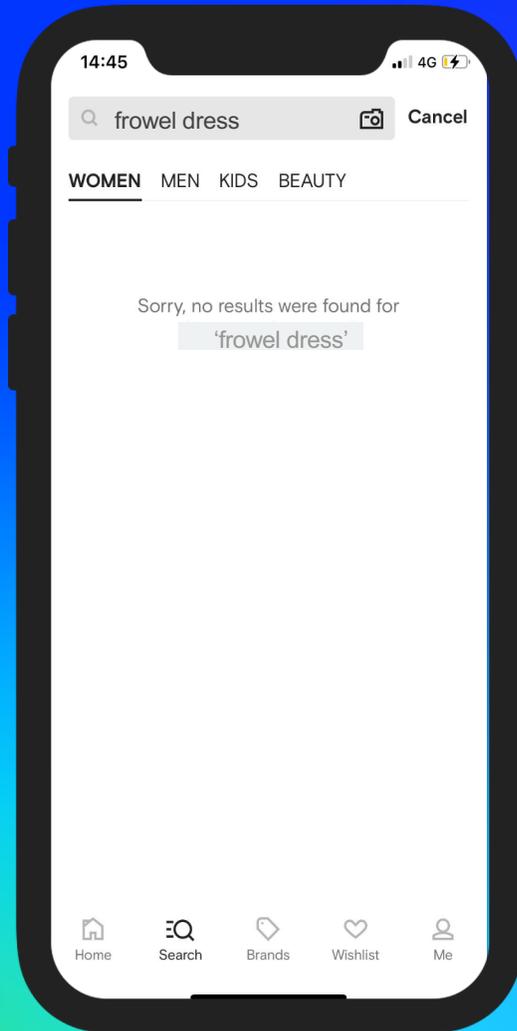
Test & Learn
shared
learning
sessions

User
Research
pain points
library

Self-service
deep dives
and business
impact

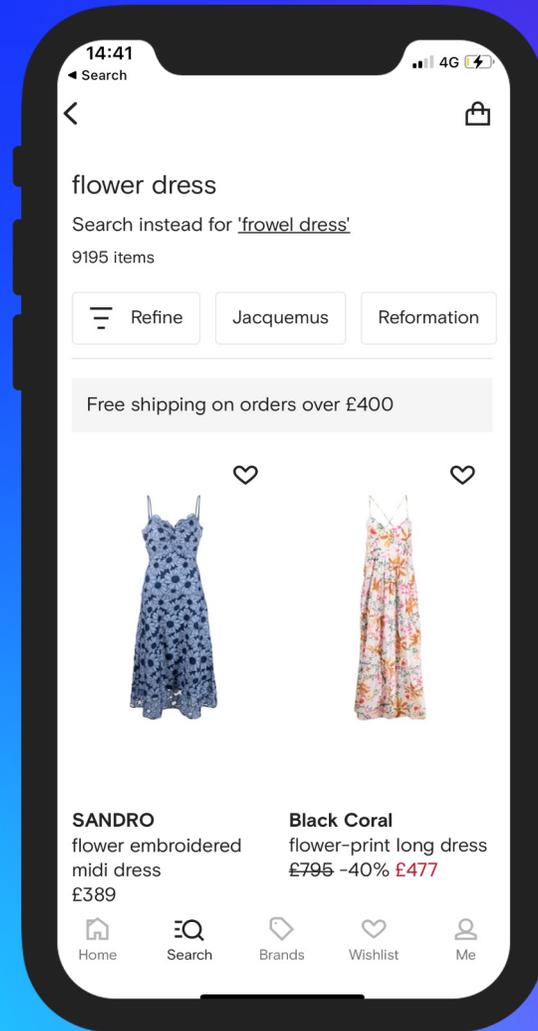
Example “Test & Learn” sharing slide

Greg



CONTROL

Searching without spellchecker



ALTERNATIVE

Searching with spellchecker

Spellchecker

Customers on IOS that misspell or mistype their queries, struggle to find the product they like and often end up in no-results pages, without knowing what to do next. Spellchecker can help with customer journey, by correcting more queries, allowing the customers to get what they need faster and easier.

iOS, Search (English only). Jun 06th - Jul 4th.
Francesca A. [LEAR](#), [Results dashboard](#)

+0.14 to 1.21% Search Engagement Rate
+0.00 to 4.24% AOVt
+0.3 to 6.87% GTV per Visit



Rollout

Experimentation at every step

Greg

**Product
Discovery**

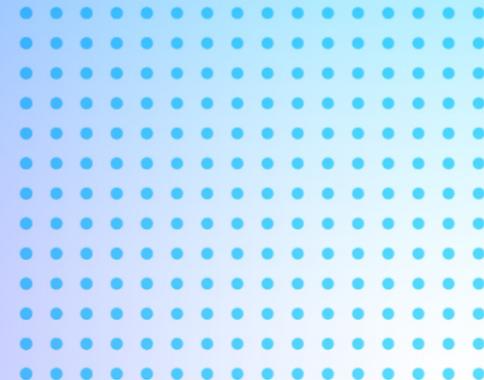
“test to learn”

Build up

“test to
decide”

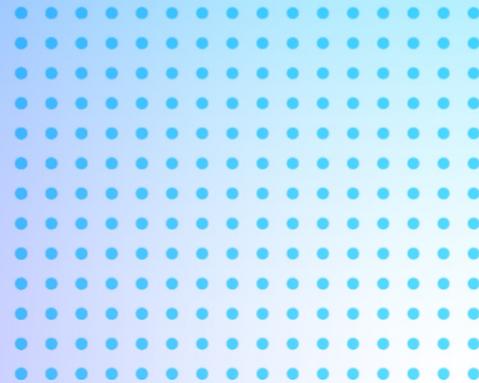
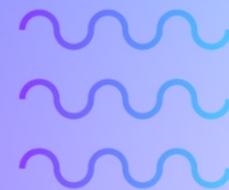
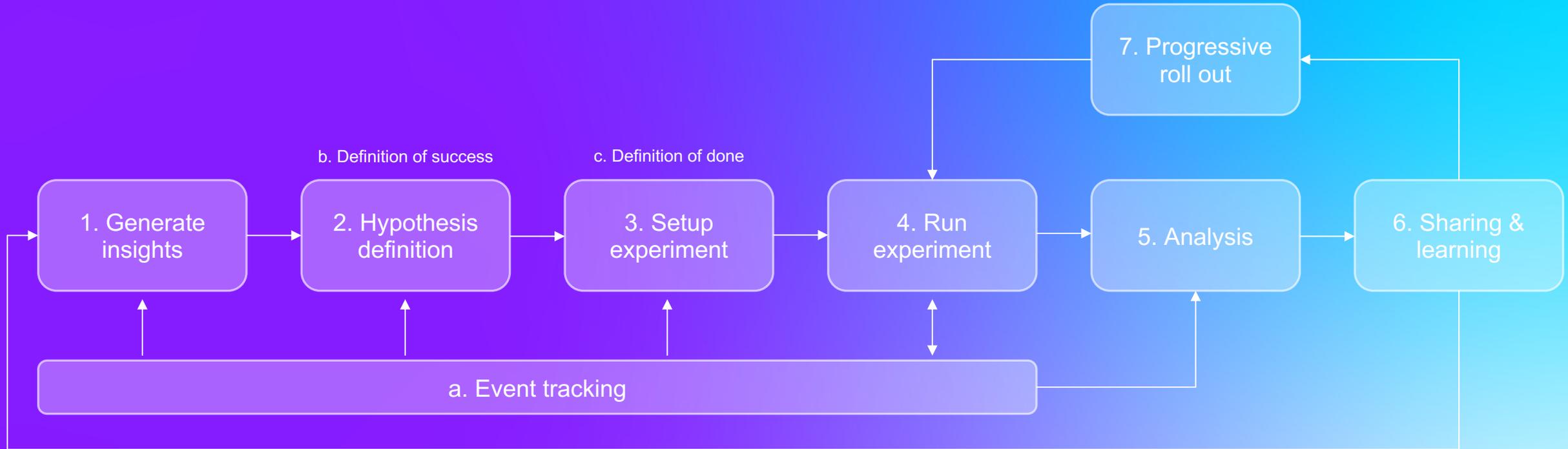
Delivery

“test to
measure”



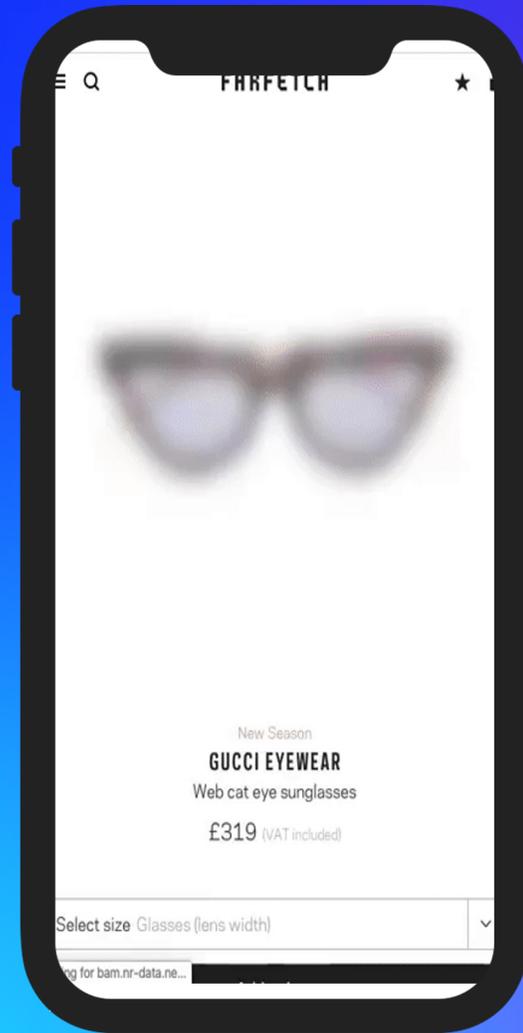
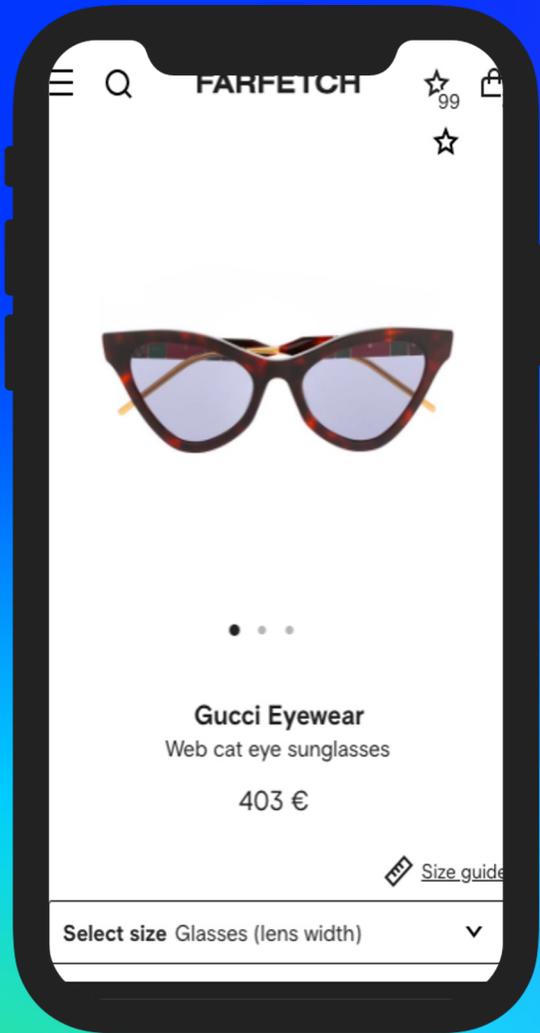
Farfetch Experimentation process flow

Greg



A “Test & Learn” learning

Greg



Virtual try on - Eyewear

Despite the appetite for AR Virtual Try-On for the eyewear category (13.4% CTR), we noticed high PDP bounce rate for users that interacted with VTO for less than 10s (30%), leading us to conclude that the long asset loading time harmed the overall user experience

Web, PDP (300 products). April 17th 2020 - June 1st. Racquel G. + Mary B. [LEAR](#), [Results dashboard](#)

-11% to -4% Add to Bag (primary metric)

-26% to -2% Add to wishlist (secondary metric)



Stop – exploring next steps for the category with the eyewear commercial team

Some Key Learnings

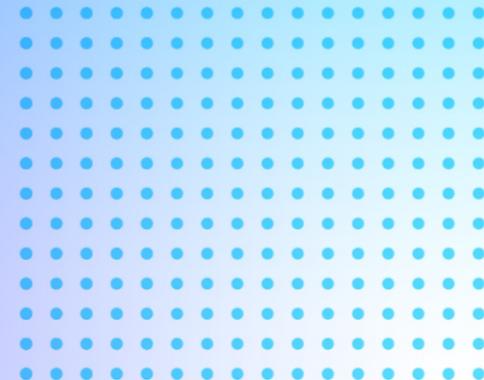
Isabel

Greg

**Go Beyond
The Numbers**

**Be more
sensitive to
Technical
Fidelity of
AR**

**Context
Matters.
Don't
generalize
results.**



Key Take-aways



Experiment at every step

Test to learn, Test to decide and Test to measure.



CoE, Rituals & Artifacts

Build a support structure for your Product Teams to help them embed experimentation in their product dev process.



You win some, you lose some

There's as much value in knowing what not to build, as knowing what to build. Celebrate losing experiments as much as winners.

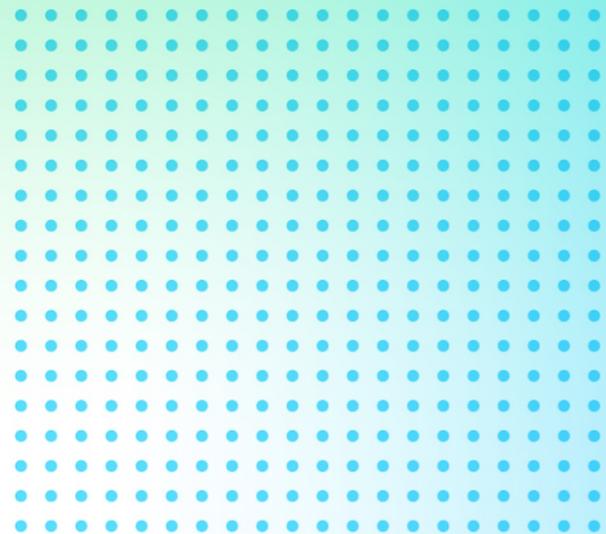


Build the right thing

Building the right thing trumps building 'it' right. Be more diligent in deciding what to build.

Agenda

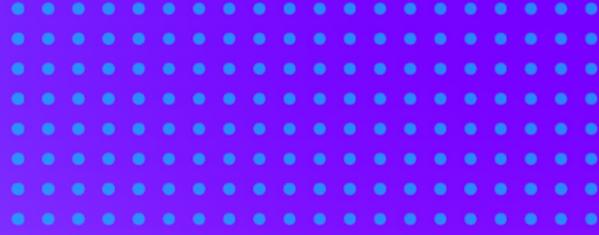
- 1. Product Dev & Experimentation**
- 2. Determine the right direction**
- 3. Drive Continuous Optimization**



03

Drive Continuous Optimization

**Validate, iterate..
repeat**



Validate, iterate.. repeat



Experimentation at every step

**Product
Discovery**

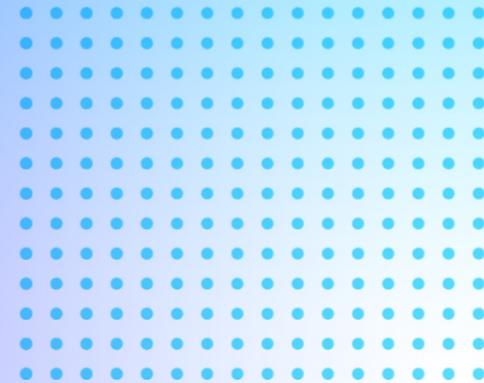
“test to learn”

Build up

“test to
decide”

Delivery

“test to
measure”



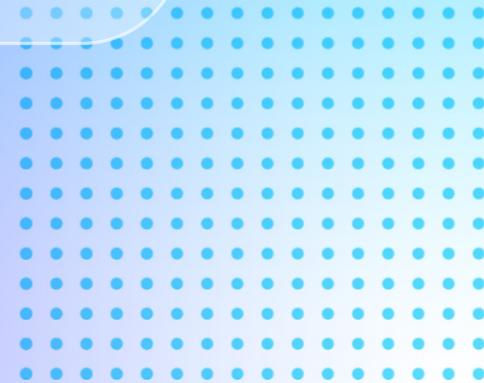
Continuous Optimization Mindset

Adoption Challenges

**Longer Term
Strategic
Planning**

**False
Confidence
in product
definition**

**Drop in
Test
Velocity**



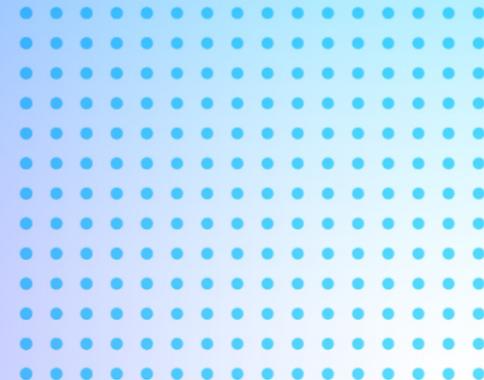
Validating Program Performance

Program Metrics

- Test Velocity
- Conclusive Rate
- Win Rate / Lose Rate
- Roll-out and Roll-back Rate
- Test set-up time
- Tests per initiative
- % Product Discovery Tests
- Test Duration

Value Metrics

- Positive Net GTV Impact
- Negative Net GTV Impact
- Annual GTV Estimated Impact
- Annual OC Estimated Impact



2023 Q2 EXPERIMENTATION METRICS

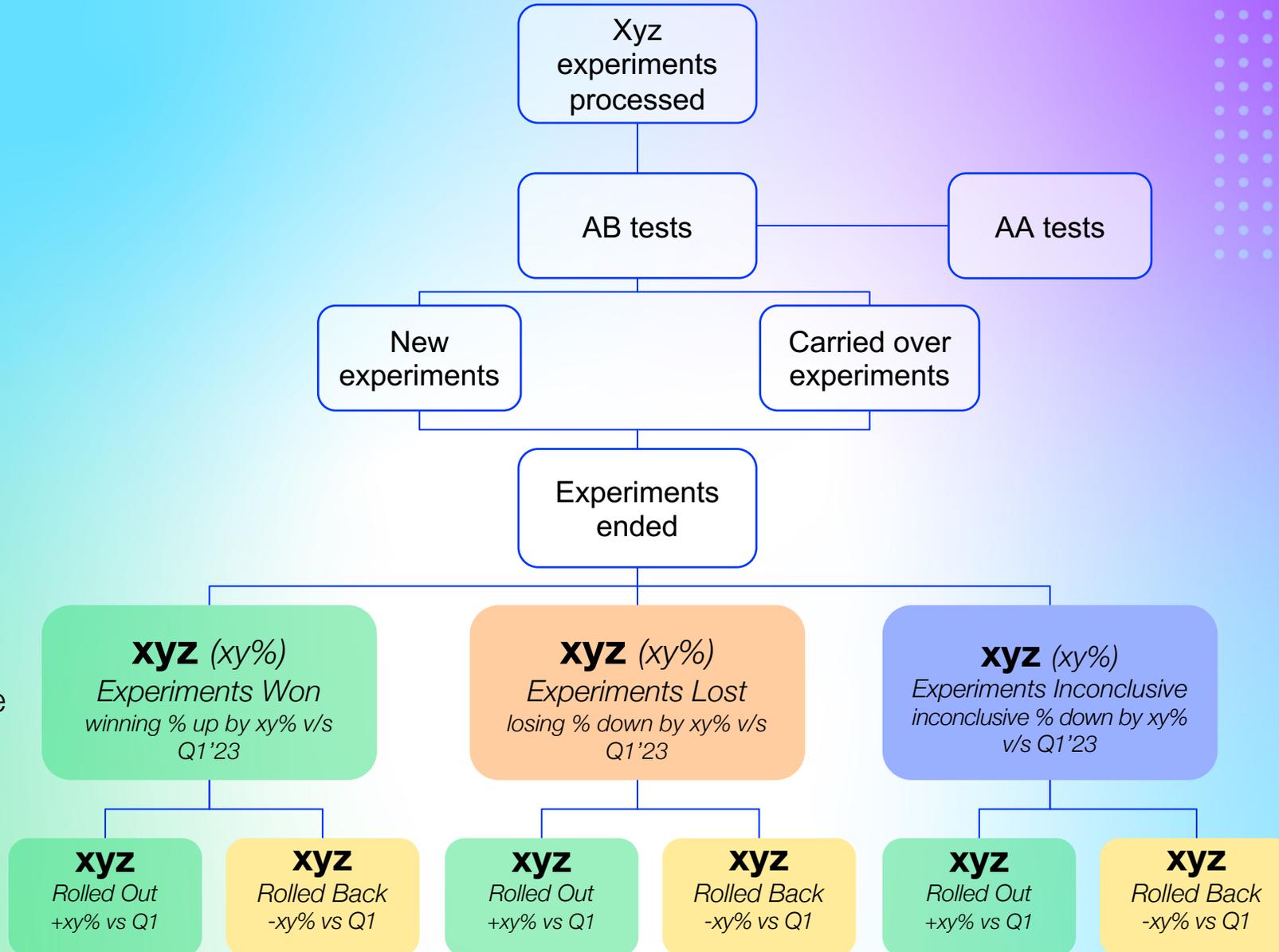
Greg

Total Experiments

Ended Experiments

Win/Lose/Inconclusive

Rollout/Rollback





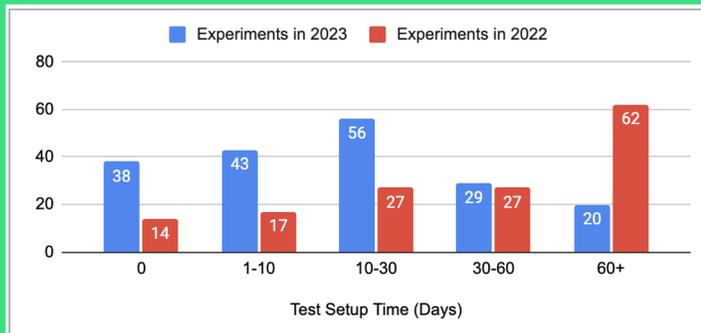
Self-service Platform

Autonomy & Easy to Use Experimentation Platform

BASELINE

Median setup time **2023** (until July) = xy days

Median setup time **2022** = xy days



Time values are for indicative purposes only

KPI: Test Setup Time

Measurement: Median of duration from draft to running state on AB Test Platform

2023 KEY TACTICS TO IMPROVE KPI

- Educate PMs and Analysts on our KPI
- Track & alert PMs of tests where setup time > 15 days
- Streamline creation of trackers & metrics for experimentation
- Optimize metrics catalog
- Improvements on AB Test Platform UI for Analytical setup (Metrics, Scope, Dimensions)

Greg

H2 2023:
median x days

2024:
median y days



Adoption & Expansion

Advance the experimentation mindset
across the company



Tests per Initiative

H2 2023: no change

2024 Target: > x.y

% of Product Discovery Tests

H2 2023: keep target (xy%) but with
better categorization

2024: yz%

KPIs: Tests Per Initiative & Percentage Product Discovery Tests

Measurement: # Tests per initiative

**Measurement: Ratio of Product Discovery tests vs. all
other tests (Risk Mitigation + Iterative/Incremental)**

2023 KEY TACTICS TO IMPROVE KPIs

- Educate Product leaders and PMs and create alignment for a 6-12 months
- Encourage Product Discovery as a key way of working for all PMs
- Ensure capture of all Initiative information related to AB tests
- Promote earlier & cheaper testing methods, alongside product discovery, during Clinics, Drop-in sessions and with PM leadership.
- Improve catalog of non-AB tests in a standardized format

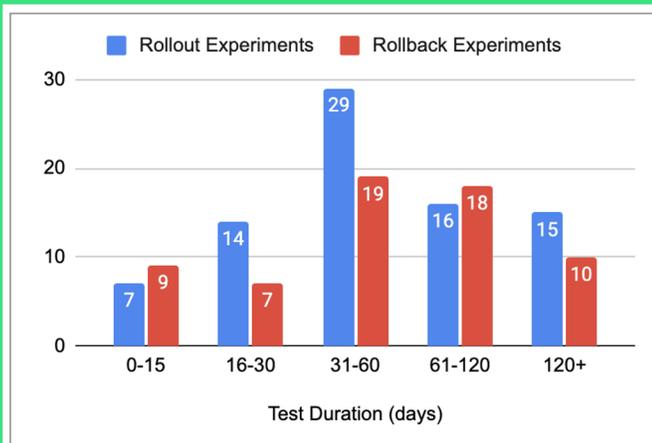


Insights
& Knowledge

Improve trust & speed of decision making

BASELINE

Median test duration **2023** (until July) = xy days



Time values are for indicative purposes only

KPI: Test Duration

Greg

H2 2023:
median xy days

2024:
median yz days

Measurement: Median of duration from running state to ended state on AB Test Platform for only rolled out or rolled back tests

2023 KEY TACTICS TO IMPROVE KPI

- Create reports and alerts/checks for PMs:
 - after 3 months; after significance reached; low significance (<50%) after 2 months; GTV per visit or “end of customer journey” as primary metrics; setup issues
- Guidelines about key metrics selection per touchpoint.
- Greater investments in self-service experiment results dashboards.
- Guidelines and education on Product-change-to-measurement-distance
 - Supported with alerts when distance is too high.

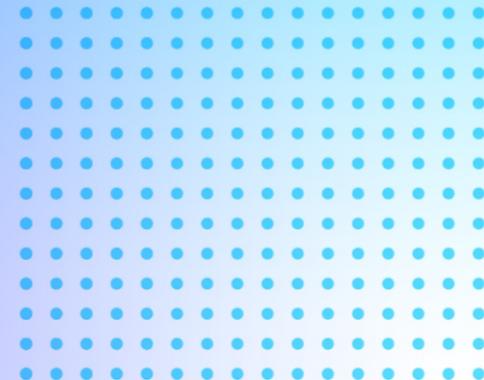
Validating Program Performance

Program Metrics

- Test Velocity
- Conclusive Rate
- Win Rate / Lose Rate
- Roll-out and Roll-back Rate
- Test set-up time
- Tests per initiative
- % Product Discovery Tests
- Test Duration

Value Metrics

- Positive Net GTV Impact
- Negative Net GTV Impact
- Annual GTV Estimated Impact
- Annual OC Estimated Impact



Annual GTV Estimated
Impact Realized

Not considering overlapping between experiments

\$xym

- \$wzyzM to + \$wxyzM

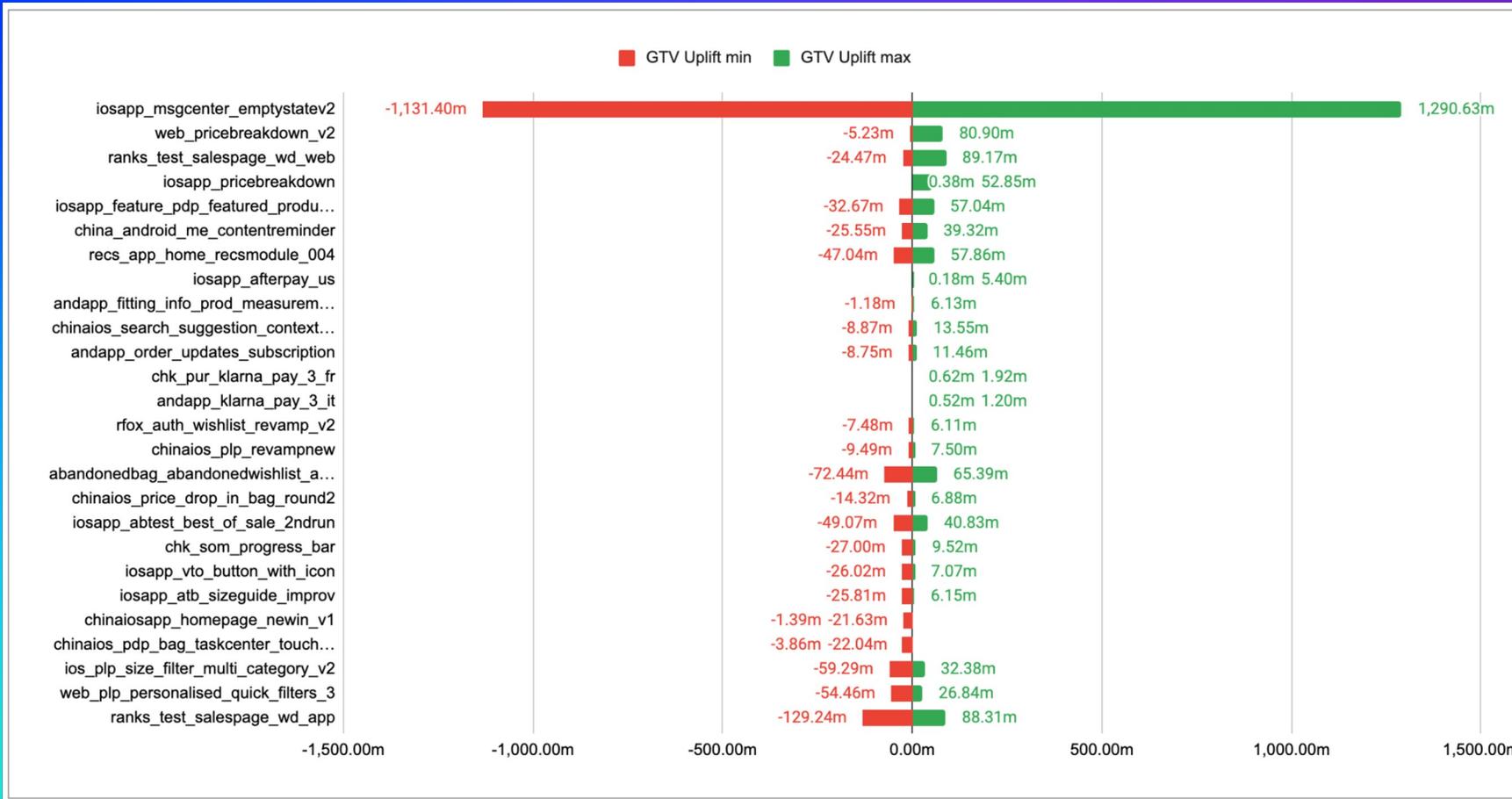
Annual OC Estimated
Impact Realized

Not considering overlapping between experiments

\$X.yzm

- \$wzyzM to + \$wxyzM

GMV values
are for
indicative
purposes only



positive net
GTV impact
realized

negative net
GTV impact
realized

**Annual GTV Estimated
Impact Realized**

Not considering overlapping between experiments

\$xyM

- \$wzyzM to + \$wxyzM

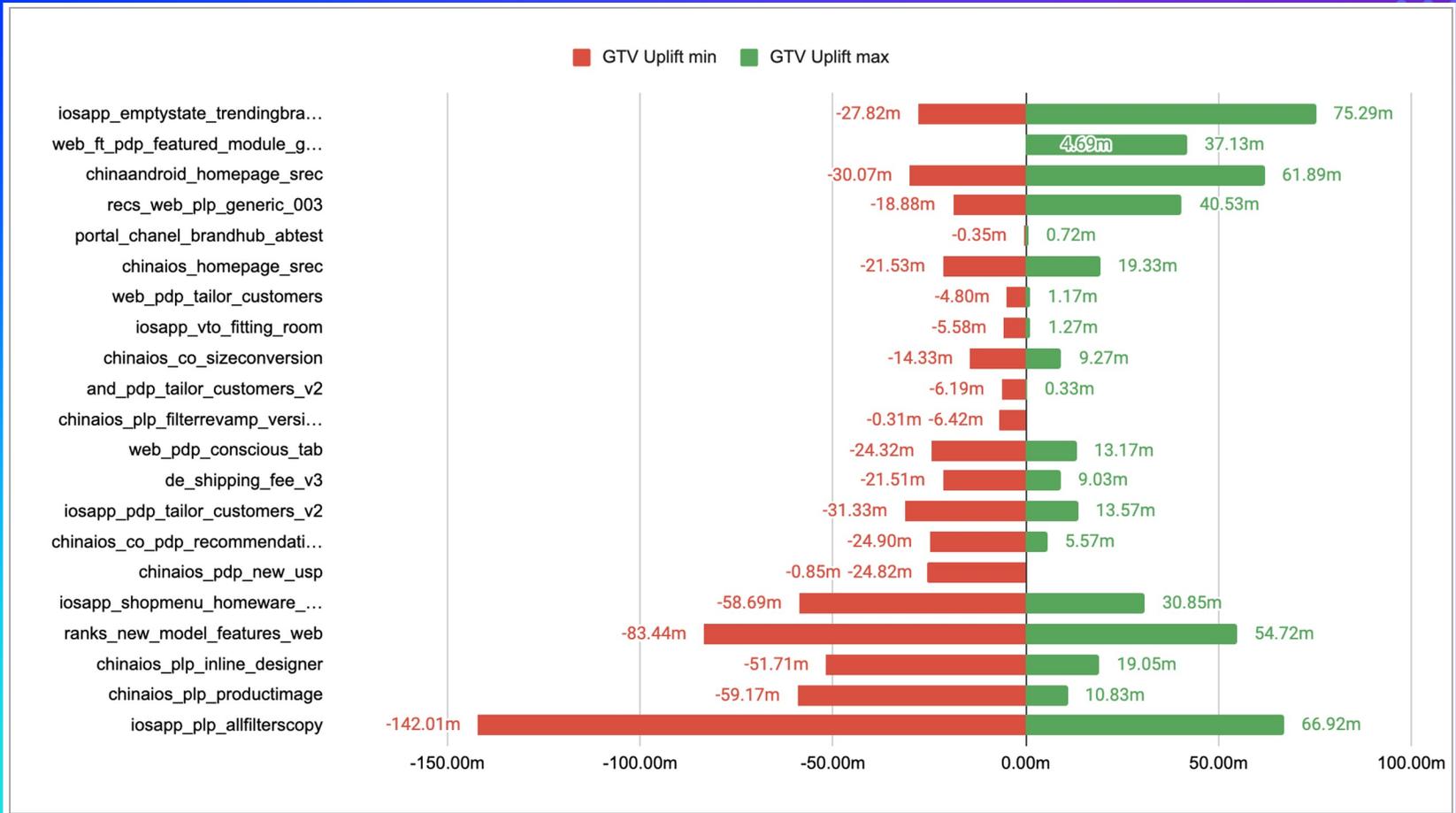
**Annual OC Estimated
Impact Realized**

Not considering overlapping between experiments

\$x.yzM

- \$xyzM to + \$xyzM

*GMV values
are for
indicative
purposes only*



**positive net
GTV impact
prevented
(Potential GTV
uplift lost)**

**negative net
GTV impact
prevented**

[Key Take-aways]

Isabel

Greg



**First main point
goes right here:
Arial bold 18pt**

And optional
supporting text: Arial
reg 16pt



**Second main point
goes right here:
Arial bold 18pt**

And optional
supporting text: Arial
reg 16pt



**Third main point
goes right here:
Arial bold 18pt**

And optional
supporting text: Arial
reg 16pt



**Fourth main point
goes right here:
Arial bold 18pt**

And optional
supporting text: Arial
reg 16pt

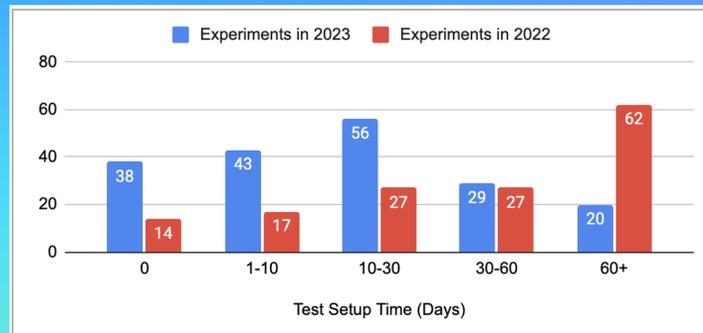
KPI: Test Setup Time

Measurement: Median of duration from draft to running state on AB Test Platform

BASELINE

Median setup time 2023 (until July) = xy days

Median setup time 2022 = xy days



2023 KEY TACTICS TO IMPROVE KPI

1. Impart knowledge to PMs and analysts on our aim to reduce test setup time
2. Track & alert PMs of tests having ongoing setup time > 15 days
3. Continue influencing a more efficient process to create trackers & metrics for experimentation
 - Include Analytics Team as part of the product dev lifecycle
 - Optimize metrics catalog
 - Tracking Brilliant Basics
 - Customer behavior telemetry (Omni-tracking) evolution
4. Improvements on AB Test Platform UI for Analytical setup (Metrics, Scope, Dimensions)



H2 2023:
median x days

2024:
median x days

KPI 1: # Tests per Initiative

Measurement: # Tests per initiative

KPI 2: % of Product Discovery Tests

Measurement: Ratio of Product Discovery tests vs. all other tests (Risk Mitigation + Iterative/Incremental)

BASELINE

- 2022 onwards: x.y Tests per initiative
- 2023 until July: xy% Product discovery, xy.z% Iterative Optimization, Risk Mitigation x.y%, Internal tests (AA tests) xy% (China-driven)

2023 KEY TACTICS TO IMPROVE KPIs

1. Create visibility of our objective with Product leaders and PMs and create alignment for a 6-12 month horizon
2. Influence that Product Discovery as a key aspect for the way of working for all PMs
3. Ensure capturing all Initiative info for AB tests
4. Keep promoting earlier & cheaper testing methods, alongside product discovery, during Clinics, Drop-in sessions and with PM leadership.
5. Improve catalog of non-AB tests in a standardized format



Tests per Initiative

H2 2023:
no change

2024 Target:
> x.y

% of Product Discovery Tests

H2 2023:
keep target (xy%) but better categorization

2024:
xy%

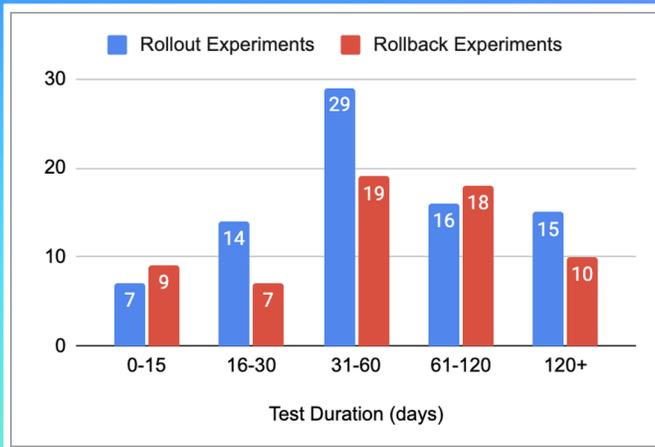


KPI: Test Duration

Measurement: Median of duration from running state to ended state on AB Test Platform for only rolled out or rolled back tests

BASELINE

Median test duration **2023** (until July) = xy days



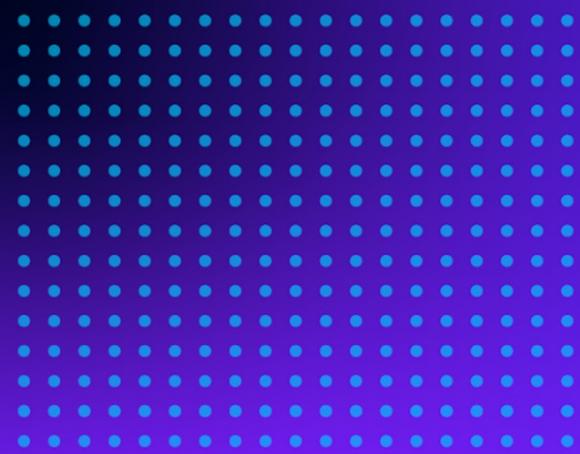
2023 KEY TACTICS TO IMPROVE KPI

1. Create reports and alerts/checks for PMs: running over 3 months; running after significance reached; low significance (<50%) after 2 months; GTV per visit or “end of customer journey” as primary metrics; setup issues
2. Guidelines about key metrics selection per touchpoint.
3. Improve our results dashboards with a more self-service approach, reducing dependency from Analytics Teams
4. Explore concept of Product-change-to-measurement-distance and provide guidelines and alerts when distance is too high.



H2 2023:
median x days

2024:
median x days



Build the right thing. Don't just build it *right*.



Embed experimentation along your product development process. Enable product teams to experiment and focus on building the right thing. Implement operational metrics to capture program success.





Reduce risk and fail faster: embedding experimentation across the entire product development lifecycle

Gregory Sherwin

Sr Principal Engineer,
Farfetch

Isabel Meijaard

Sr Manager Strategy
Consulting, Optimizely

