

# Lecture 5:

# Designing efficient systems

Measuring and optimising human performance through quantitative experimental methods.

# Overview of the course

- Theory driven approaches to HCI
- Design of visual displays
- Goal-oriented interaction
- Designing smart systems (guest lecturer)
- **Designing efficient systems**
- Designing meaningful systems (guest lecturer)
- Evaluating interactive system designs
- Designing complex systems

# Lessons from text entry - recap

- It's possible to model human action
- It's possible (in part) to predict human action
- Efficiency can be predicted, and also measured
- A really fundamental trade-off:
  - Speed versus accuracy

# Fitts' Law

# User actions are information-constrained

How many bits of information to select one of these choices?



How many bits of information to select one of these choices?



The human neuromotor system is limited by information rate - size of target relative to movement

# Demonstration of Fitts' Law

# Fitts' Law – the only equation in HCI!

- How long does it take to point at something?
- Proportional to the **D**istance to target
- Inversely proportional to **W**idth of target
- Like most human performance (and most things in information theory), it's a log function:
- $\text{Time} = k \log (2D/W)$

# Speed-accuracy tradeoff

- Users are capable of doing things faster
- But making more mistakes as a result
- Did your application need speed, or accuracy?



## 1. State EOC

### 1. TEST Message

DRILL-PACOM (DEMO) STATE ONLY

False Alarm BMD (CEM) - STATE ONLY

Monthly Test (RMT) - STATE ONLY

PACOM (CDW) - STATE ONLY



# Hacking Fitt's Law: “semantic pointing”



# Small changes can have a big effect (1972)

## Psychological Evaluation of Two Conditional Constructions Used in Computer Languages

M. E. SIME, T. R. G. GREEN AND D. J. GUEST

---

NEST solution:

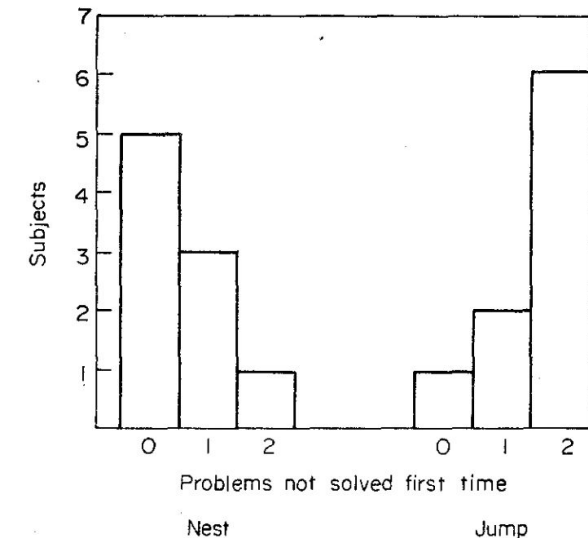
```
IF JUICY THEN
  IF LEAFY THEN
    IF GREEN THEN GRILL
  OTHERWISE BOIL
OTHERWISE FRY
OTHERWISE
  IF HARD THEN ROAST
  OTHERWISE REJECT
```

vs

JUMP solution:

```
IF JUICY GOTO L1
IF HARD GOTO L2
REJECT
L2 ROAST
L1 IF LEAFY GOTO L3
  FRY
L3 IF GREEN GOTO L4
  BOIL
L4 GRILL
```

=>



KLM/GOMS: Predicting time

# Keystroke Level Model (KLM)

Model an interaction as series of operators, to predict the time an expert takes to do something

Operator	Time/s	Description
K	0.2	Key or button press
P	1.1	Pointing
H	0.4	Homing, switching hand between keyboard/mouse
M	1.35	Mental preparation
R	?	System response time

# Keystroke Level Model (KLM)

Rules for when you should insert operators (NOT EXAMINABLE)

1. Insert Ms in front of Ks and Ps that select commands
2. Remove any Ms that are fully anticipated
3. Remove all but the first M from runs of MK that are a single cognitive unit
4. Remove any Ms where the K is a redundant terminator
5. Remove Ms from terminate constant strings

# Example

M (before command)  
H (hand -> mouse)  
P (point at "Tweet")  
K (Click)  
R (wait for response)

The screenshot displays the Twitter mobile app interface. At the top, navigation tabs include Home, Moments, Notifications, Messages, and a profile icon. A search bar and a 'Tweet' button are also visible. The main content area shows a tweet from 'The Onion' (@TheOnion) with the text 'National News Highlights' and a video titled 'Columbia, MO' showing a map of Missouri with a location pin. Below the tweet, there's a retweet and a video player. To the right, a 'Who to follow' section lists accounts like Whitney Museum, J. Paul Getty Museum, and Hillary Clinton. A red arrow points from the top right towards the 'Tweet' button.

Home Moments Notifications Messages

Search Twitter Tweet

What's happening?

**Luke Church** @Ambiguitus  
Tweets 0 Following 6

Pick a profile photo  
Have a favorite selfie? Upload it now.  
30% complete  
Skip Add a photo

**Trends for you** · Change

- #TuesdayThoughts  
12.7K Tweets
- Britney Spears  
Britney Spears is coming to the UK to tour
- Kyle Edmund  
Kyle Edmund reaches Australian Open semi-finals
- #r4today  
2,618 Tweets
- #climatehack2018
- Hugh Masekela  
30.2K Tweets
- Charing Cross  
More than a thousand evacuated after gas leak in London's West End
- Neil Diamond  
Neil Diamond to end concert career due to Parkinson's disease
- #BIMASocial
- Germaine Greer  
Germaine Greer takes aim at the Me Too movement

**The Onion** @TheOnion · 4h  
National News Highlights

**NATIONAL NEWS HIGHLIGHTS**

**Columbia, MO**  
Contractor Lewis Petersen, 46, felt guilty about installing floors in a new house on Doyle Street, as he knows good and well that floors are unnecessary and a scam.

0:03 32.7K views

3 28 109

The Onion Retweeted

**ClickHole** @ClickHole · 18h  
Real Or Fake? Man Eats Entire Rice Cake Without Blinking

**Who to follow** · Refresh · View all

- Whitney Museum** @wh...  
Follow
- J. Paul Getty Museum** ...  
Follow
- Hillary Clinton** @Hillary...  
Follow

**Find people you know**  
Import your contacts from Gmail

Connect other address books

© 2018 Twitter About Help Center Terms Privacy policy Cookies Ads info Brand Blog Status Apps Jobs Marketing Businesses Developers

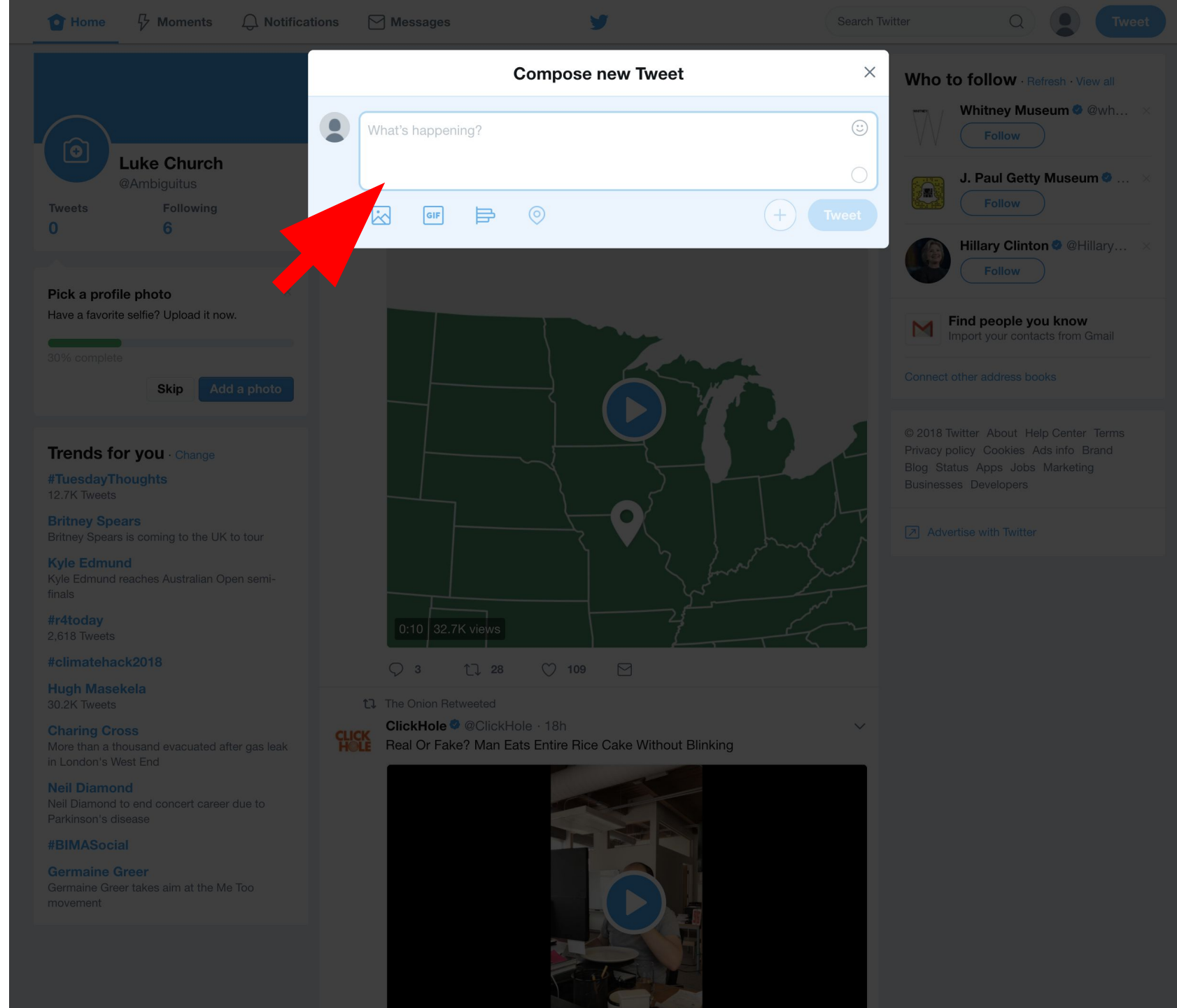
Advertise with Twitter

# Example

M (before command)

P (point at “What’s happening?”)

K (Click)



# Example

M (Prepare to type)

K

K

K

K

K

K

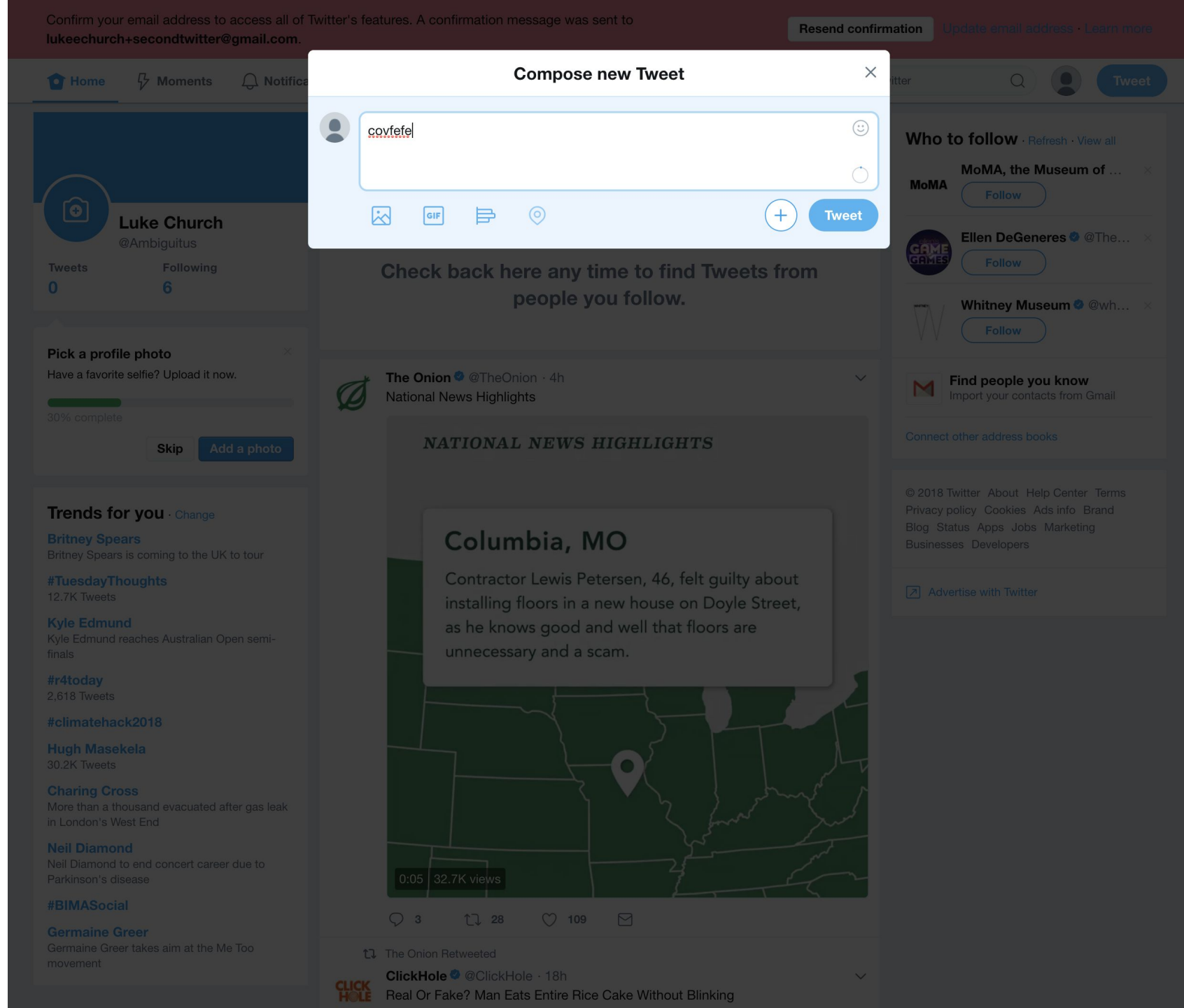
K

M (Prepare to click)

P (Point at “Tweet”)

K (Click)

R (Wait for response)





# Example

MHPKR

MPK

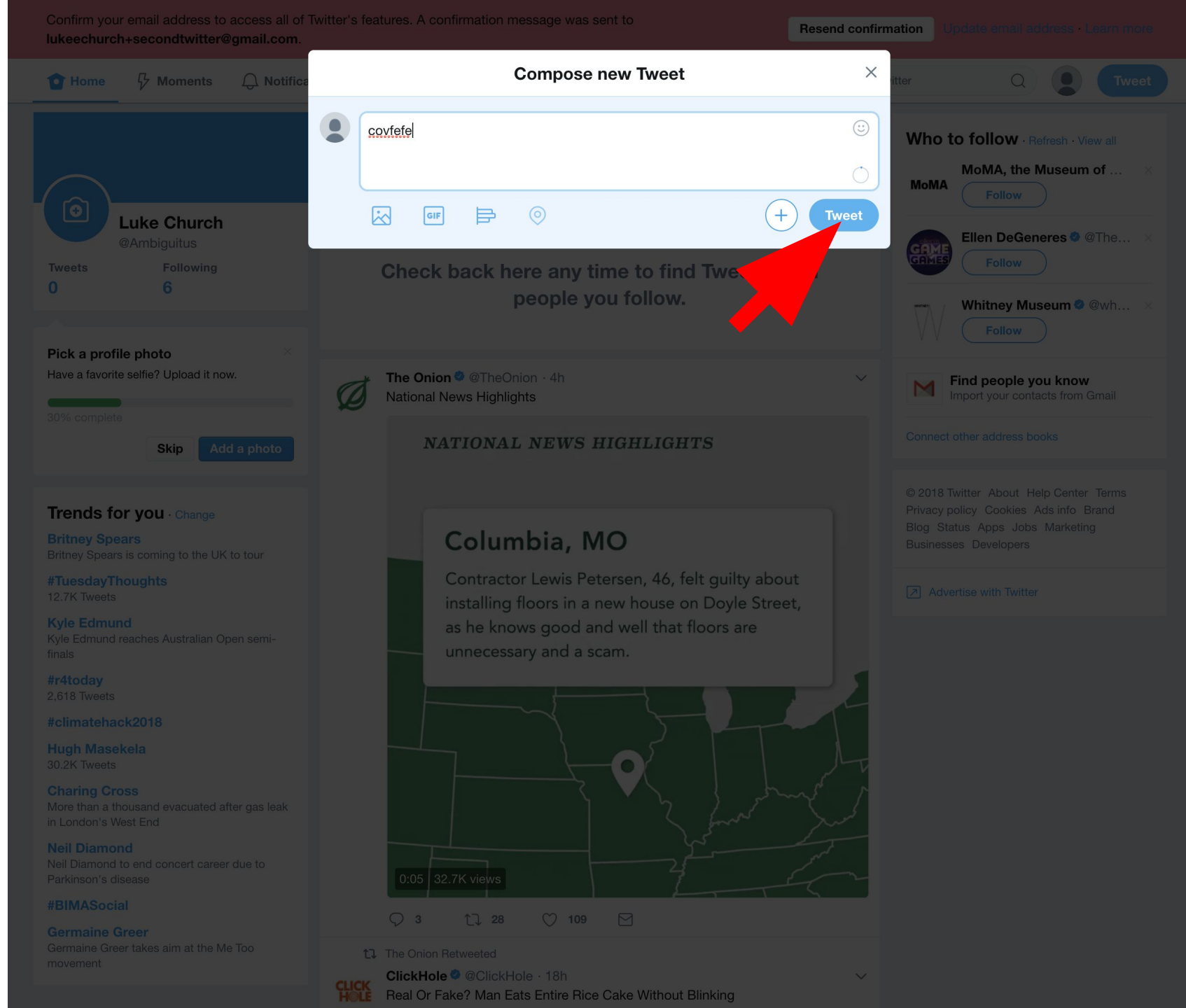
MKKKKKKKMPKR

$1.35 + 0.4 + 1.1 + 0.2 + \sim 0.2$

$1.35 + 1.1 + 0.2$

$1.35 + 7 \cdot 0.2 + 1.35 + 1.1 + 0.2 + \sim 0.2$

= 11.5s



# Keyboard shortcuts

Keyboard shortcuts

Actions

n

New Tweet

l

Like

r

Reply

t

Retweet

m

Direct message

u

Mute User

b

Block User

Enter

Open Tweet details

o

Expand photo

/

Search

Cmd

Enter

Send Tweet

Navigation

?

This menu

j

Next Tweet

k

Previous Tweet

Space

Page down

.

Load new Tweets

Timelines

g

h

Home

g

o

Moments

g

n

Notifications

g

r

Mentions

g

p

Profile

g

l

Likes

g

i

Lists

g

m

Messages

g

s

Settings

g

u

Go to user...

# Example

M (become command)  
K ('n')  
R (wait for response)

The screenshot displays the Twitter mobile app interface. At the top, navigation icons for Home, Moments, Notifications, and Messages are visible, along with a search bar and a 'Tweet' button. A red arrow points to the 'Tweet' button.

The main content area shows a tweet from 'The Onion' (@TheOnion) posted 4 hours ago. The tweet is titled 'NATIONAL NEWS HIGHLIGHTS' and features a video player showing a map of Missouri with a location pin in Columbia, MO. The video description reads: 'Contractor Lewis Petersen, 46, felt guilty about installing floors in a new house on Doyle Street, as he knows good and well that floors are unnecessary and a scam.' The video has 32.7K views and 3 comments.

Below the tweet, there is a section for 'Trends for you' with several trending topics listed, including #TuesdayThoughts, Britney Spears, Kyle Edmund, #r4today, #climatehack2018, Hugh Masekela, Charing Cross, Neil Diamond, and #BIMASocial.

On the right side, there is a 'Who to follow' section with a list of suggested accounts: Whitney Museum, J. Paul Getty Museum, Hillary Clinton, and a 'Find people you know' section with a 'Connect other address books' link.

# Example

M (Prepare to type)

K

K

K

K

K

K

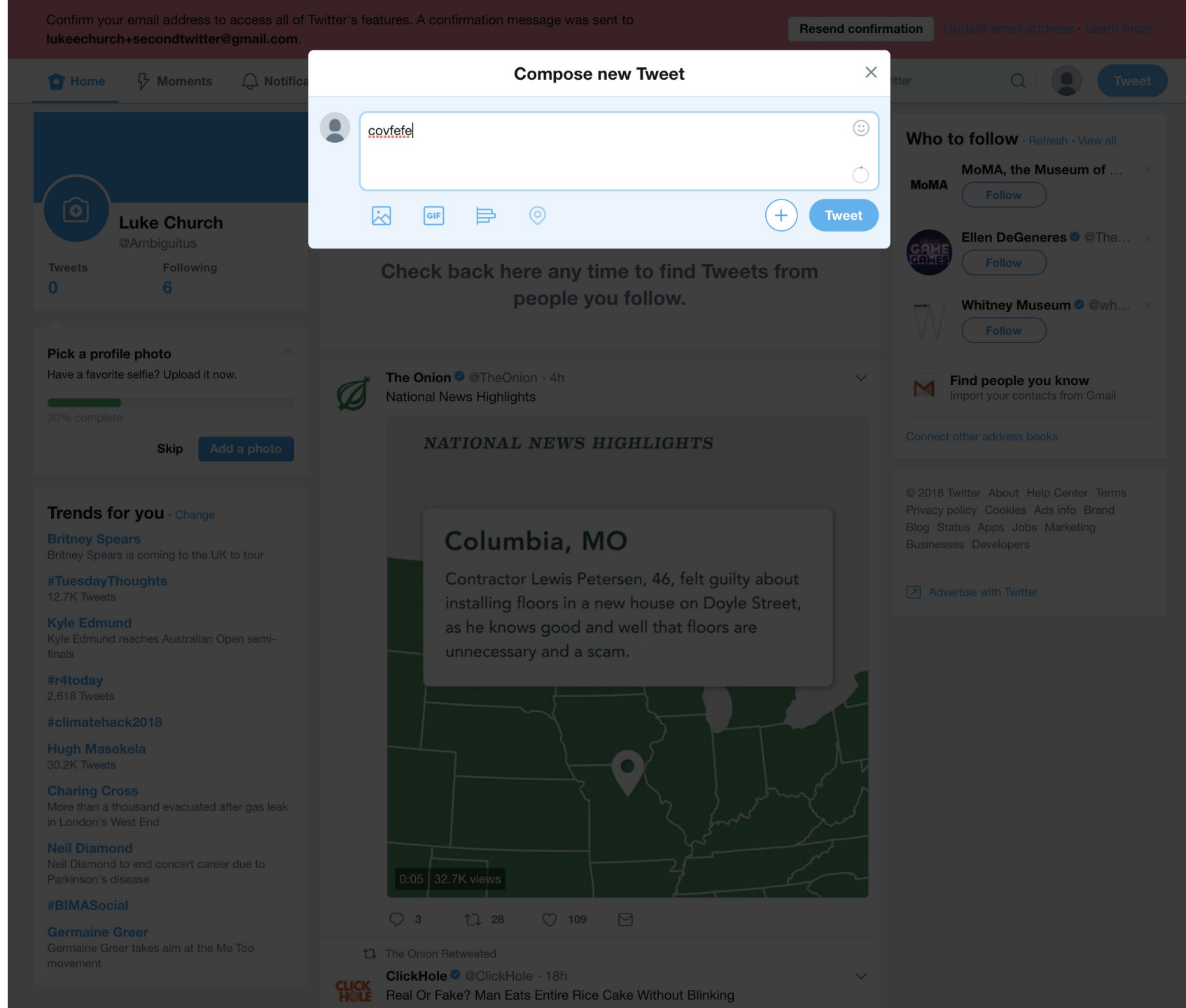
K

M (Prepare to click)

K (cmd)

K (enter)

R (Wait for response)



# Example

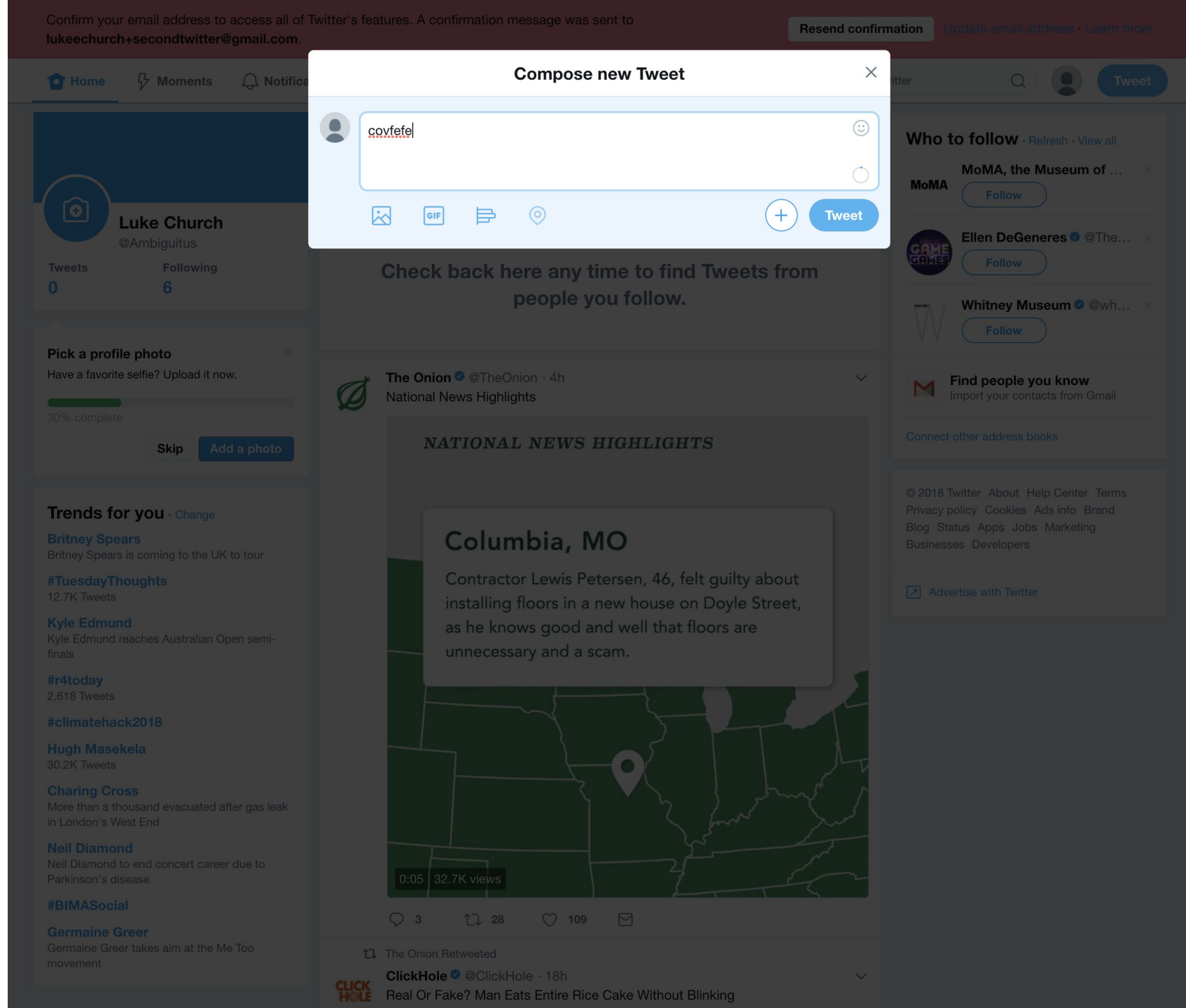
MKR  
MKKKKKKKMKKR

$1.35 + 0.2 + \sim 0.2$

$1.35 + 7 \cdot 0.2 + 1.35 + 0.2 + 0.2 + \sim 0.2$

= 6.45s

(Compared to 11.5s before)

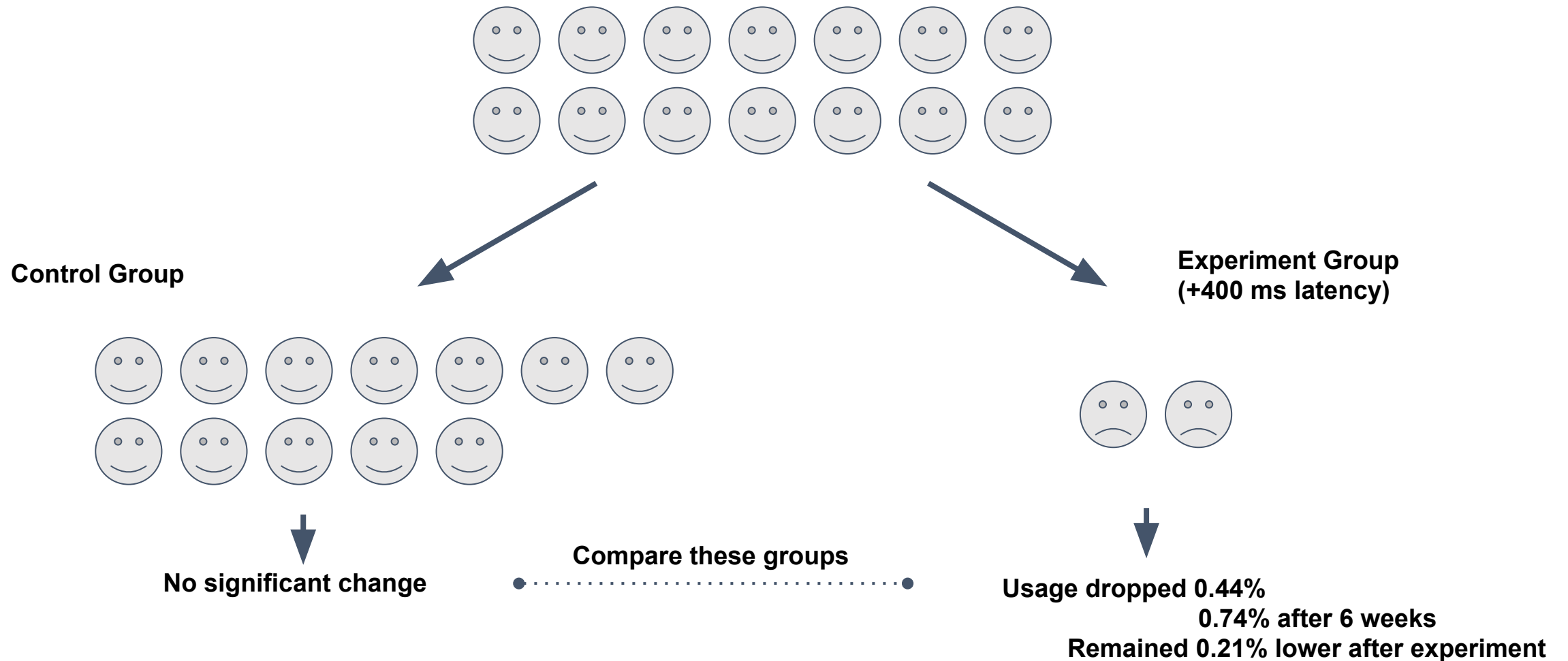


Experiments: Measuring time/usage

# How many links should be on a search result page? (10, 20 or 30?)

- User studies: More is better
- When given 30, usage fell - why?
  - Analysis showed 400ms extra latency

# Latency experiment

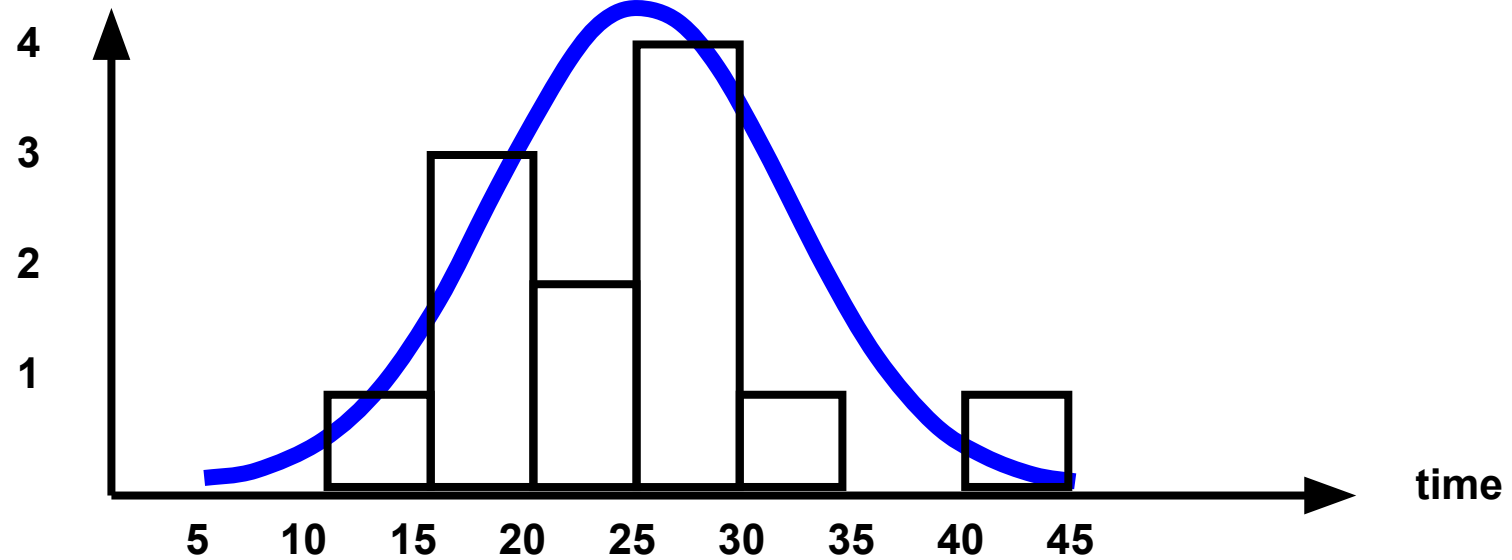




These are A/B experiments

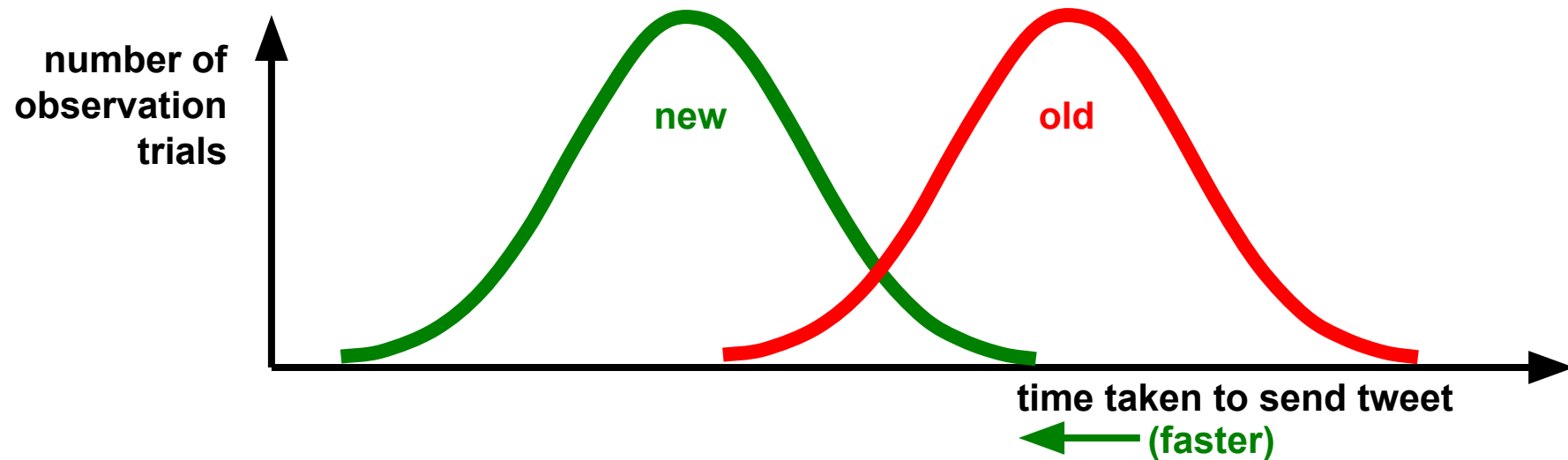
(statistics: histograms & distributions)

number of  
observations



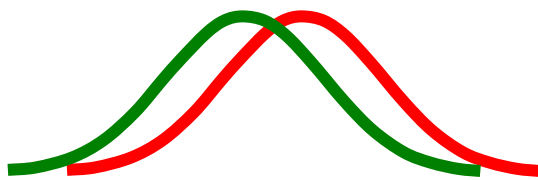
# Experimental treatments

- A *treatment* is some modification that we expect to have an effect on usability:
  - How long does Donald take to send his tweet using this great new interface, compared to the crummy old one?
  - Expected answer: *usually* faster, but not *always*

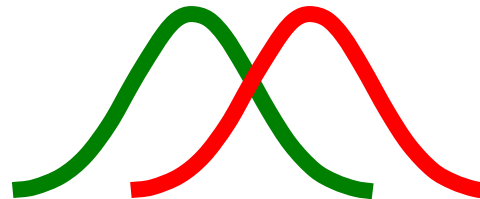


# Hypothesis testing

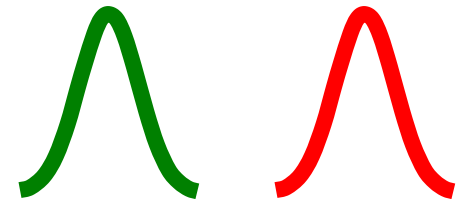
- *Null hypothesis:*
  - What is the probability that this amount of difference in means could be random variation between samples?
  - Hopefully very low ( $p < 0.01$ , or 1%)
  - Use a statistical *significance test*, such as the *t-test*.



only  
random  
variation  
observed



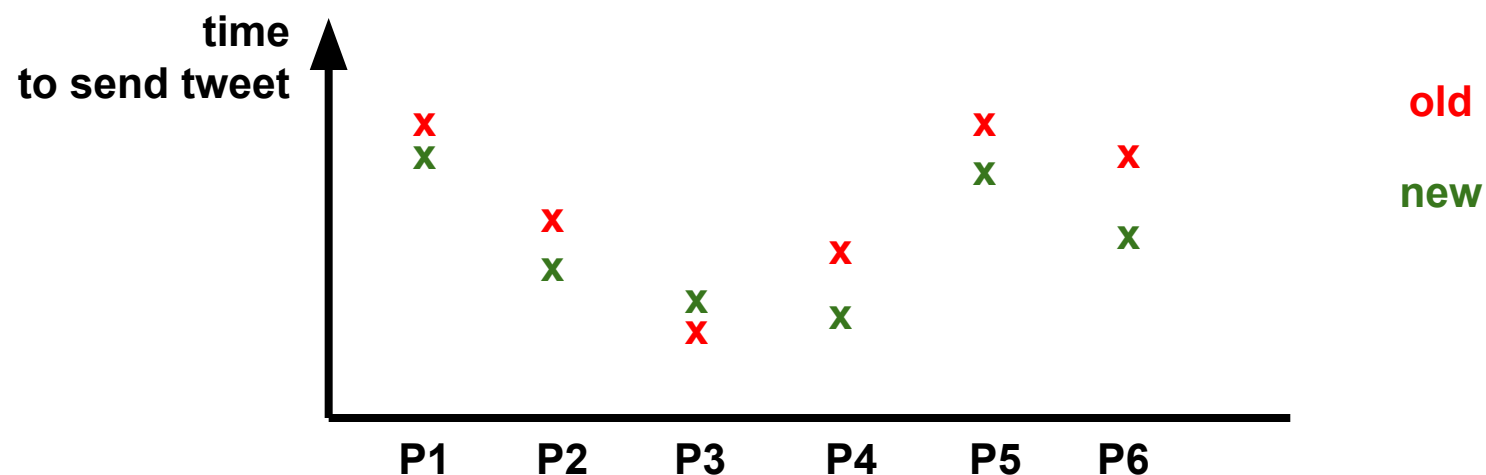
observed effect  
probably does  
result from  
treatment



very significant  
effect of  
treatment

# Sign tests

- In a within subjects experiment it's possible to compare the results
  - Explores the [null] hypothesis that the median of the pairs is zero
  - Means might not be significant, but the sign can be
  - This is a non-parametric test, so doesn't depend much on the data, but not very powerful (use a paired t-test, or Wilcoxon rank test instead)



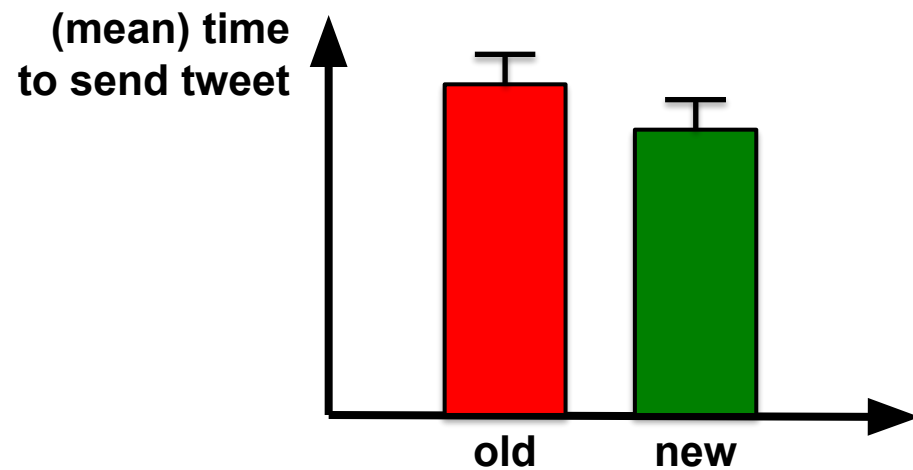
Experiment A: 'significant' but boring

# Sources of variation

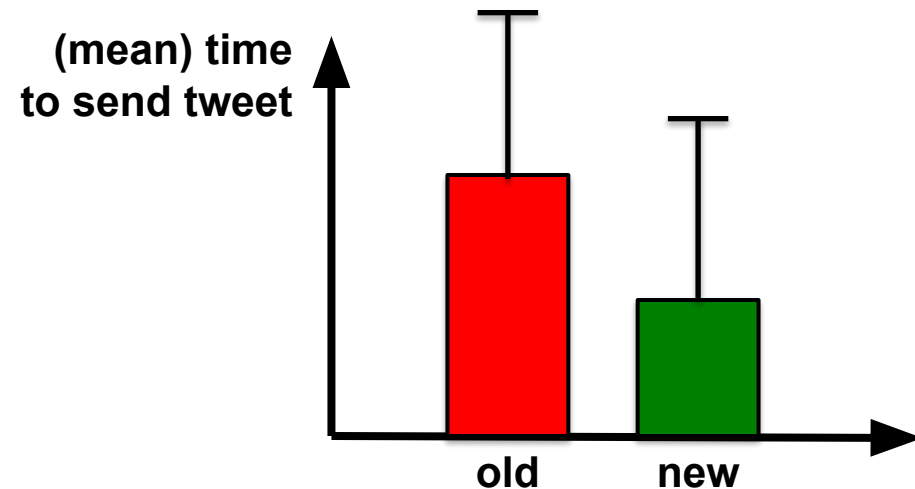
- People differ, so quantitative approaches to HCI must be statistical.
- We must distinguish sources of variation:
  - The effect of the treatment - what we want to measure.
  - Individual differences between subjects (e.g. IQ).
  - Distractions during the trial (e.g. sneezing).
  - Motivation of the subject (e.g. Mondays).
  - Accidental intervention by experimenter (e.g. hints).
  - Other random factors.
- Good experimental design and analysis isolates these.

# Effect size – means and error bars

- Difference of two means may be statistically significant (if sample has low variance), without being very interesting.
  - But mean differences must *always* be reported with a confidence interval, or plotted with 'error bars'



Experiment A: 'significant' but boring



Experiment B: interesting, but treat with caution

# Problems with controlled experiments

- Huge variation between people (~200%)
- Mistakes mean huge variation in accuracy (~1000%)
- Improvements are often small (~20%)
- ... or even negative (because new & unfamiliar)
- ... and may result from something unrelated to your design!



# The Hawthorne Effect



- Studies on productivity in 1924-1932
  - Do lighting levels affect productivity?
  - Studies appeared to show improvements in both directions
  - Results show the motivational effect of being studied, not of the change

# Is efficiency always a design goal?

- What if you wanted to encourage thoughtfulness? Creativity?

# Taylorism

- F.W. Taylor (1856-1915)
  - Engineer who invented scientific management
  - Measure workers as if parts in a machine
  - Optimise by measurement and correction
- Not so popular with trade unions!
  - Note that 2nd wave HCI (the turn from human factors to social science) involved working closely with trade unions, especially in Sweden and Denmark



# Discretionary use systems

If you are not working to someone else's goal, you can decide whether or not to be efficient (or whether you want to use the system at all)



Simone Giertz: "Queen of Shitty Robots"

# Efficient creativity?

- What if there isn't a good measure of productivity?
  - Maximise output of poetry-lines?
  - Maximise musical notes played per second?
  - Maximise Cambridge graduates per year?
- Optimum User Experience
  - What if you wanted people to enjoy what they did?

