Digital Disruption in Cognitive Psychology

Tom Stafford
8th Sept 2022

tomstafford.staff.shef.ac.uk
@tomstafford

These slides: http://bit.ly/tom-talks
Abstract

Digital technology has changed how the world works, from the music industry to democracy. These changes have also come to cognitive psychology, challenging us to adopt new methods, learn new analysis techniques and adapt to new ways of relating to the participants in our experiments and to the users of our research. Using examples from my own explorations in open research, online research, public engagement and data intensive methods I will sketch the opportunities and challenges cognitive psychology faces.

These slides: http://bit.ly/tom-talks
Motivation
On a september day...
On a september day...22 years ago


These slides: http://bit.ly/tom-talks
On a september day...22 years ago

What are the rules of this game?
Cognitive Section founded 1978

\[ S_{S_3} = \sum_{j=1}^{k} (\bar{x}_j - \bar{x})^2 \]
Digital disruption

UG psychologists:
2000: 30k.
2011: 60k
2014: 100k

(and massification, and hypercompetition)

Image: Anthony Poynton “Retro Vinyl Record” CC0 Public Domain

These slides: http://bit.ly/tom-talks
“...in this field of behavioural science, there are a lot of charlatans...the whole field is riddled with duff studies and memes that people believe are true but are not true.”
Psychological science has low reproducibility

Replication of 100 studies all published in 2008, from three important psychology journals: Psychological Science (PSCI), Journal of Personality and Social Psychology (JPSP), and Journal of Experimental Psychology: Learning, Memory, and Cognition (JEP:LMC)

(mal)adaptations

underpowered studies  data intensity
p-hacking              power simulation
competition            team science
extractive relations   engagement
insularity             interdisciplinarity

These slides: http://bit.ly/tom-talks
Why? Low statistical power

Studies in psychology are chronically under-powered, reducing the chances of findings true effects, shifting the balance of in favour of false positives.

<table>
<thead>
<tr>
<th></th>
<th>$p &lt; 0.05$</th>
<th>$p &gt; 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True effects</strong></td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td><strong>Null effects</strong></td>
<td>False Alarm</td>
<td>Correct Rejection</td>
</tr>
</tbody>
</table>
We have always known this
e.g. Cohen, 1962
Average power was 0.48 for medium effects

“These values are deemed to be far too small and suggest that much research in the abnormal-social area has lead to the failure to reject null hypotheses which are in fact false. This in turn may have lead to frequent premature abandonment of useful lines of investigation”

These slides: http://bit.ly/tom-talks
It’s not getting any better

Szucs & Ioannidis, 2017 analysed 26,841 statistical records from 3,801 cognitive neuroscience and psychology papers published 2011 - 2014

Average power was 0.44 for medium effects

“our data suggest that power in cognitive neuroscience and psychology papers is stuck at an unacceptably low level. This is so because sample sizes have not increased during the past half-century”

Low power: consequences

Overestimate effect size
Shift ratio of Hits:False Alarms
-> significant results are less indicative
-> Ultimately: wasted time and effort


These slides: http://bit.ly/tom-talks
Data Intensive Psychology


An idea!

These slides: http://bit.ly/tom-talks
These slides: http://bit.ly/tom-talks
These slides: http://bit.ly/tom-talks
Example player data for people who played more than 15 times
n=854064

These slides: http://bit.ly/tom-talks
Not just data...need theory
Practice: amount
"at least 10,000 hours of dedicated practice (about 6 years of playing chess 5 hours a day) are required to attain the highest levels of performance" (Kahneman, 2011, p238).


The “Ten Thousand Hours Rule”


The “Ten Thousand Hours Rule”


These slides: http://bit.ly/tom-talks
These slides: http://bit.ly/tom-talks
These slides: http://bit.ly/tom-talks

Destiny

~120 million active players


These slides: http://bit.ly/tom-talks
Performance functions

These slides: http://bit.ly/tom-talks
Stafford & Haasnoot (2017)
A graph showing average score as a function of time with different learning strategies:

- **Practice (1-15 attempts)**
- **Spacing (0-60 minutes)**
- **Initial performance (percentile)**

Stafford & Haasnoot (2017)
Lots more to do

Does it matter how you practice?
  - spacing? variability? exploration?

Can we predict the level of skill someone will eventually acquire?
  - How? How soon?

How to escape plateaux in learning?
Learning curves as theoretical anchor

Games tremendous opportunity for studying skill acquisition, but the huge, messy, data must be probed with

- formal fitting of learning curves
- experiments for strong causal inference


Curve fitting code in R and Python: https://osf.io/fvm8s/.

These slides: http://bit.ly/tom-talks
Learning curves as theoretical anchor

This figure shows a simple, three-parameter, power-law learning curve: Performance, \( f(t) \), is a function of practice, \( t \); an upper limit, \( u \); the learning gain, \( a \), which defines how far initial performance is from the upper limit; and the learning rate, \( c \). The notation follows Steyvers and Benjamin (2019).

Code for implementing this learning-curve function, and fitting it to data, is available on OSF, at https://osf.io/fvm8s/.
These slides: http://bit.ly/tom-talks

- asymptote
- learning rate
- initial performance

Fit params: $u = 35719, a = 26175, c = -0.24$
Maximizing the Potential of Digital Games for Understanding Skill Acquisition

Tom Stafford and Nemanja Vaci
Department of Psychology, University of Sheffield

Abstract
Gaming is a domain of profound skill development. Players’ digital traces create data that track the development of skill from novice to expert levels. We argue that existing work, although promising, has yet to take advantage of the potential of game data for understanding skill acquisition, and that to realize this potential, future studies can use the fit of formal learning curves to individual data as a theoretical anchor. Learning-curve analysis allows learning rate, initial performance, and asymptotic performance to be separated out, and so can serve as a tool for reconciling the multiple factors that may affect learning. We review existing research on skill development using data from digital games, showing how such work can confirm, challenge, and extend existing claims about the psychology of expertise. Learning-curve analysis provides the foundation for direct experiments on the factors that affect skill development, which are necessary for a cross-domain cognitive theory of skill. We conclude by making recommendations for, and noting obstacles to, experimental studies of skill development in digital games.


Curve fitting code in R and Python: https://osf.io/fvm8s/.

These slides: http://bit.ly/tom-talks
The general lessons
(mal)adaptions

underpowered studies  data intensity
p-hacking  power simulation
competition  team science
extractive relations  engagement
insularity  interdisciplinarity

These slides: http://bit.ly/tom-talks
(mal)adaptions

underpowered studies    data intensity
p-hacking               power simulation
competition            team science
extractive relations   engagement
insularity             interdisciplinarity

These slides: http://bit.ly/tom-talks
Research on Research
Research on Research Institute (RoRI)

http://researchonresearch.org

These slides: http://bit.ly/tom-talks
We’re hiring (soon!)

- From October, at least 1 PDRA, contract until 2027, full time
- would suit researcher with a background in experiment design, quantitative analysis, and/or metascience (research on research)
- based at RoRI and The Department of Psychology, University of Sheffield, UK
- opportunities to design & lead projects, journal publications; grow collaborations; mentoring
- details https://tomstafford.staff.shef.ac.uk/?p=954
- informal enquiries welcome t.stafford@sheffield.ac.uk
Strong Foundations

- Fundamentals
- Theory
- Computation
- Collegiality

Image: Stable Diffusion (via huggingface.co),
prompt “ancient foundations building towards a sky of stars”
END

Tom Stafford, t.stafford@sheffield.ac.uk
http://tomstafford.staff.shef.ac.uk
@tomstafford

These slides: http://bit.ly/tom-talks
RESERVE SLIDES
FOLLOW
(not for show)
Learning
Learning curves as theoretical anchor

Fig. 1. The learning curve as a theoretical anchor for studies of skill acquisition in games. This figure shows a simple, three-parameter, power-law learning curve: Performance, $f(t)$, is a function of practice, $t$; an upper limit, $u$; the learning gain, $a$, which defines how far initial performance is from the upper limit; and the learning rate, $c$. The notation follows Steyvers and Benjamin (2019).

Code for implementing this learning-curve function, and fitting it to data, is available on OSF, at https://osf.io/fvm8s/.
Fit params: \( u = 35719, a = 26175, c = -0.24 \)

 Observed

 fit \( y = u - ax^c \)

 osf.io/fvm8s/
Learning curves as theoretical anchor

Experiments with games need:

- Entertainment value / intrinsic motivation
- Challenge / modifiable difficulty
- Scoring / clear outcome measure
- Benchmarks / absolute or relative
- Components / identifiable and quantifiable
- Out of game measures / e.g. Kokkinakis et al. (2017)


Curve fitting code in R and Python: https://osf.io/fvm8s/.

These slides: http://bit.ly/tom-talks
Power simulation
Decision models

https://github.com/tomstafford/DDMgif
Decision models


Decision models

Key parameters

Drift rate = participant sensitivity / stimulus strength

Response threshold = participant bias / their speed-accuracy trade-off


Hit Rate
No Sato

Parameters:
drift: 1 vs 1.1
boundary: 2 vs 2
intersubj var = 0.05
trials/ppt = 40

These slides:
Hit Rate

No Sato

Parameters:
- drift: 1 vs 1.1
- boundary: 2 vs 2
- intersubj var = 0.05
- trials/ppt = 40

These slides: http://bit.ly/tqs-talks
ppts/group for 80% power:

RT: ~140
Acc: ~115
Drift: ~55 !!

trials/ppt = 40

These slides: http://bit.ly/tqs-talks
Paper: Quantifying the benefits of using decision models with response time and accuracy data

- pre-print
- interactive data explorer: sheffield-university.shinyapps.io/decision_power

These slides: http://bit.ly/tqs-talks
Team Science
Research question: “Are soccer players with dark skin tone more likely to receive red cards from referees than players with light skin tone?”

Four leagues: English, Germany, France and Spain

2012-13 Season

2,034 Players

3,147 Referees*

These findings suggest that significant variation in the results of analyses of complex data may be difficult to avoid, even by experts with honest intentions.”

https://fivethirtyeight.com/features/science-isnt-broken/#part1
“161 researchers in 73 research teams and observed their research decisions as they used the same data to independently test the same prominent social science hypothesis: that greater immigration reduces support for social policies among the public”

High volatility of results
AND
sources of volatility not easily characterised
Engagement
Dialogues with the vaccine hesitant


Dr Charlotte Brand
lottybrand.wordpress.com
"Scientific consensus" is not to be trusted when so many are reliant on funding from pharmaceutical companies. They have much to lose if they do not follow the narrative.

From an individual perspective, I have no issues surrounding vaccinations. With an extensive travel history I have received more vaccinations than the average individual, however my concern regarding the Covid vaccination lies in just how much of a threat is Covid to the general population? I do not trust the Media narrative or figures hence it is a person’s free will choice whether they should accept this vaccine or not. Much more research needs to be done into just how dangerous Covid 19 actually is.

I had one vaccine, felt obligated to do so and everyone I spoke to was having it. I had the Astra Zeneca vaccine, cheap one. Then all the news was on about blood clots and countries blocking the AZ vaccine as it wasn’t safe. I then cancelled my 2nd vaccine appointment after that. The vaccine is untested, too many intelligent people are against it, the government tells lies and too many of them for me to believe a word they say, to the point I will never vote in an election ever again. For such a killer virus going around there has been far too much normality and the vaccine rules and protocol continually change so why should people trust it, if it’s a killer virus and everyone has to be vaccinated then why are people allowed to fly in and out of the country since we first knew about it. We don’t even know if the vaccine stops you being a virus carrier, it’s clearly an untested vaccine. Also since when does the Government care about old people, disabled, sick people but now they are the important ones at the top of the list to be vaccinated.

Information about vaccines seems v biased towards having it hence the distrust I’m not anti vaccines - i just want to be fully informed and resent being referred to as an ‘anti vaxer’ (you haven’t! :-( )
Interdisciplinarity
Stepping back, my take on the 'replication crisis' in psychology is that it really isn't about replication. Instead, this is what digital disruption looks like in a culture organised around scholarly kudos rather than profit. We now have the software tools to coordinate data collection, share methods and data, analyse data, and interact with non-psychologists, both directly and via the media, in unprecedented ways and at an unprecedented rate. Established scholarly communities are threatened as "the way things are done" is challenged.
Image: Stable Diffusion (via huggingface.co),
prompt “a ruined fortress representing the discipline of cognitive psychology”
Image: Stable Diffusion (via huggingface.co),
prompt “crowd of scholars on a spaceship”
Image: Stable Diffusion (via huggingface.co),
prompt “ancient foundations building towards a sky of stars classic painting style”
Image: Stable Diffusion (via huggingface.co), prompt “lonely academic sitting among ruins”
Image: Stable Diffusion (via huggingface.co),
prompt “lonely academic sitting among ruins”
Image: Stable Diffusion (via huggingface.co),

prompt “joyful team building a futuristic tower, scifi pop art”