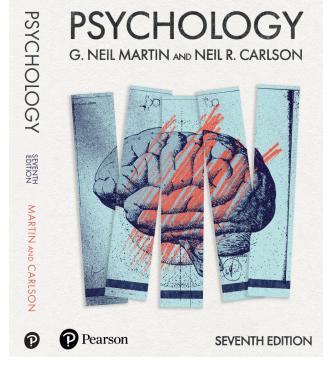
Practical ideas for introducing new & controversial research to the introductory psychology programme

Professor G. Neil Martin, FRSA





How much of a psychologist are you?





True or False?

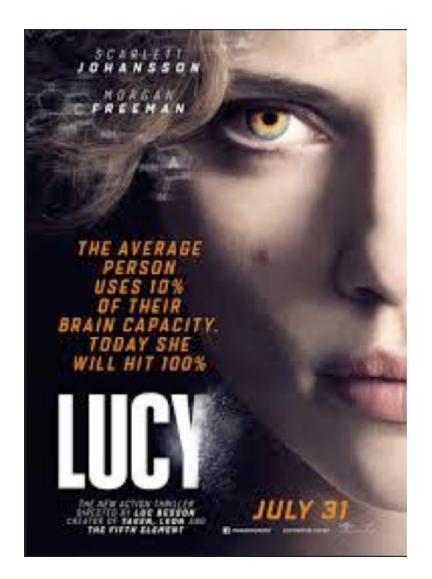
- A person with schizophrenia has a split personality
- Women are more likely to conform than men
- People who threaten to commit suicide rarely do
- We only use 10% of our brain
- Astrology is a good predictor of personality
- "Opposites attract" (in romantic relationships)
- Neuroscience has concluded that people are left-brained or right-brained
- People exposed to words about old people leave a laboratory walking more like an old person
- People who suffer from amnesia typically cannot recall their own name or identity
- Human memory works like a video camera, accurately recording events
- Once you have experienced an event and formed a memory of it, that memory does not change.



They are all...

FALSE







MAKING INTRO PSYCH (EVEN MORE) INTERESTING

- Replication (and well-known psychological phenomena)
- Questionable research practices
- Deception
- Face recognition
- Knowledge of memory
- Sense of smell (and taste)
- Cognitive psychology and classroom learning
- Some highlights of the 7th edition of psychology



PSYCHOLOGY & REPLICATION/REPLICABILITY

"the integrity of research, especially medical and social science research, is at risk from what is known as the 'reproducibility crisis' (i.e. it being very difficult or impossible to replicate a scientific study)." (<u>https://</u> <u>committees.parliament.uk/committee/135/science-and-technology-</u> <u>committee/news/156859/reproducibility-of-research-inquiry-launched/</u>)

UK Parliament's Science and Technology Committee, 2021



6th edition-2018

WHAT IS PSYCHOLOGY? 7

Controversies in psychological science: Replication

The issue

What a time to be alive. If you read certain journals, news media and websites and listen to certain researchers, you might be vaguely aware that some psychologists think the discipline is undergoing a crisis. But why? And what is the cause of the crisis? At the heart of the problem is this: replication failure. Replication refers to the process whereby a study is

repeated in a way that is as close to the original as possible with the prediction that the same results will be produced as those found in the original study. A direct replication will match the original study almost identically (but the participants will differ, the building will differ, the experimenter will differ. but these are sources of variance which should be negligible if an effect is strong). A 'conceptual' replication will reproduce the conditions of the original experiment only loosely - various parts of the methodology might differ, for example, or there may be the addition of different types of group. What has made news in Psychology recently is that psychologists have been unable to replicate some well-known effects and findings in the discipline. That is, they have not been able to find the same results as early studies and these findings were thought to be robust and reliable (Pashler and



Harris, 2012; Pashler and Wagenmakers, 2012; Laws, 2013; APS, 2015). Many of them led to theories and many citations and became classic in the field being widely cited in textbooks, including this one. Some of the more notable replication failures are summarised in Table 1.2. The problem is not specific to psychology - only 11

per cent of 53 landmark preclinical cancer trials have been found to replicate (Begley and Ellis, 2012) with 35 per cent of pharmacology studies replicating (Prinz et al., 2011). Of the 49 most widely cited papers in medicine, only 44 per cent were replicated (lonnidis, 2005). Sixty per cent of 55 studies failed to replicate in finance (Hubbard and Vetter, 1991), 40 per cent in advertising (Reid et al., 1981) and 46 per cent in accounting, management, finance, economics and marketing (Hubbard and Vetter, 1996). The situation is somewhat better in education (Makel and Plucker, 2014), human factors (Jones et al., 2010) and forecasting (Evanschitzky and Armstrong, 2010).

In psychology, priming research involving social factors has been particularly susceptible to non-replication (Earp and Trafimow, 2015). In priming research, participants' behaviour is influenced by exposure to stimuli of which they are not consciously aware or which influences them in a way in which they were not aware. For example, presenting someone with words related to being a professor might lead to participants performing well on a later IQ test (because their schema of a professor has been activated)

Klein et al. (2014) found that in 13 replication attempts of ssic and contemporary findings using 36 samples comprising 6.344 participants, 10 were successful, one was weakly replicated and two sets of findings were not. Both failures to replicate involved social priming. The latest last set of replication attempts, however, found that of 100 experiments published in journals from the year 2008, fewer than half were successfully replicated (Open Science Collaboration, 2015), as you can see in Figure 1.1.

These, and similar results, suggest that some research may be susceptible to Questionable Research Practice (QRP) which leads to positive findings that may ultimately be nothing more than well-camouflaged Type 1 errors (Simmons et al., 2011; John et al., 2012; Makel, 2014). In statistics and psychology a Type 1 error is one where your statistical analysis yields a result - say a difference between two groups - that has an extremely low probability of being found by chance but is actually false (the opposite is a Type 2 error where a statistical test suggests that there is no difference between two groups or conditions when there actually is).

8 1 THE SCIENCE OF PSYCHOLOGY

Controversies in psychological science: Continued

Table 1.2 Some recent failures to replicate well-know	wn findings In psychology	
Finding	Author/s	Author/s of failed replication
Exposure to high achievement words leads to better cognitive task performance	Bargh et al (2001)	Harris et al (2013)
Exposure to honesty-related words males people disclose alcohol-related behaviour	Rasinski et al (2005)	Pashler et al (2013)
People who wrote about secrets estimated hills to be steeper	Slepian et al (2012)	LeBel & Wilbur (2013)
Using imagery to reduce prejudice	Birtel & Crisp (2014)	McDonald et al (2014)
Priming with religion affects hand grip endurance	Hone & McCullough (2015)	Hone & McCullough (2015)
Priming with 'elderly'-related words makes people walk like old people	Bargh et al (1986)	Shanks et al (2013)
Priming with words about professors makes people work harder	Dijksterhuis & Van	Doyen et al (2012)
	Knippenberg (1998)	
Evidence of precognition	Bem (2011)	Ritchie et al (2012)
We feel psychologically closer to people physically close to us	Bash & Shalev (2012)	Nosek & Lakens (2014)
Holding a pen lengthways in the mouth increases happiness	Strack et al (1988)	Wagenmakers et al (2016)
Priming moral purity leads to cleaning-related thoughts/ behaviours	Zhong & Liljenquist's (2006)	Earp et al (2014)
Stereotype threat during maths tests in women	Spencer et al (1999)	Ganley et al (2013)
Power posing	Carney et al (2010)	Ranehill et al (2015)
Reminders of money alter people's political views	Vohs et al (2006)	Rohrer et al (2015)
Avoidant individuals will respond to social warmth with intimacy	MacDonald & Borsook (2010)	Philipp-Muller & MacDonald (201

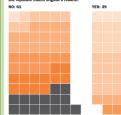
John et al. (2012) found that more than 50 per cent of the 2,000 psychologists they asked wished to extend their study until a sufficient sample was recruited which produced selectively publishing studies that produced significant results. significant finding or set of findings which are, to all intents and purposes, Type I errors. Type 1 errors in psychology and statistics are findings that are spurious - a study may show a statistically significant difference between two groups, for example, but this difference is spurious because it is based on problems with data ollection, the method or data analysis

P-hacked publications arise from studies with small samples and power, which are selective in the sample they recruit and retain, which engage in post-hoc data analysis selection, a statistically significant result. Forty per cent had reported and which may exclude negative results, conditions, experiments and even participants. Simmons et al. (2011), for exam-Both are examples of p-hacking, the publication of a statistically ple, have demonstrated how easily a surprising and logically derived finding or phenomenon can turn out to be utterly bogus because of the methodological and statistical design and analysis decisions made by the experimenter. In their study, they were able to demonstrate that people who listened to a song about getting older made people feel younger. In the lowing table, they report (in bold) what was included in the

Controversies in psychological science: Continued

RELIABILITY TEST

An effort to reproduce 100 psychology findings found that only 39 held up.* But some of the 61 non-replications reported similar findings to those of their original papers. Did replicate match original's results?



original report and what was excluded from the written-up report (in grey). It shows how excluding data and being flexible with your data

WHAT IS PSYCHOLOGY? 9

analysis can lead to false positives. For example, they included a number of factors in their study but did not include others the greater the number of factors (or variables) in your study. the easier it will be to find a significant result. You have an approximately 1 in 20 chance of finding a significant result regardless of whether that result is genuine. It is unacceptably easy', they conclude, 'to publish 'statistically significant' evidence consistent with any hypothesis. And reporting dramatic single effects like these reflects another problem: publication bias. This refers to journals' tendency to publish only results which are positive and statistically significant and a reluctance to publish negative results.

Why has some research in psychology been so difficult to replicate? One reason might be that the original result was not robust or valid in the first place. In statistics, this is called a false positive - claiming a result or finding exists when it doesn't, lournals have been criticised for publishing and prioritising novel, creative, significant results and for not publishing negative results or direct replications. This is called a publication bias - publishing only positive results. The result of this is that non-supporting research does not get published and that positive, but spurious, results are embedded in the literature in perpetuity. Martin and Clarke found that of 1,551 psychology iournals, only 3 per cent accepted replications. Many journals emphasised the importance of new and original research.

original study

Virtually identical Extremely similar Very similar tely similar Somewhat similar Slightly simila Not at all similar

* based on criteria set at the start of each stud

Figure 1.1 Replication graph http://www.nature.com/news/firstresults-from-psychology-s-largest-reproducibility-test-1.17433

Table 3. Study 2: Original Report (in Bolded Text) and the Requirement-Compliant Report (With Addition of Gray Text)

Using the same method as in Study I, we asked 34 University of Pennsylvania undergraduates to listen only to either "When I'm Sixty-Four" by The Beatles or "Kalimba" or "Hot Potato" by the Wiggles. We conducted our analyses after every session of approximately 10 participants; we did not decide in advance when to terminate data collection. Then, in an ostensibly unrelated task, they indicated only their birth date (mm/dd/yyyy) and how old they felt, how much they would enjoy eating at a diner, the square root of 100, their agreement with "computers are complicated machines," their father's age, their mother's age, whether they would take advantage of an early-bird special, their political orientation, which of four Canadian quarterbacks hey believed won an award, how often they refer to the past as "the good old days," and their gender. We used father's age to control for variation in baseline age across participants.

An ANCOVA revealed the predicted effect: According to their birth dates, people were nearly a year-and-a-half younger after listening to "When I'm Sixty-Four" (adjusted M = 20.1 years) rather than to "Kalimba" (adjusted M = 21.5 years), F(1, 17) = 4.92, p = .040. Without controlling for father's age, the age difference was smaller and did not reach significa (Ms = 20.3 and 21.2, respectively), F(1, 18) = 1.01, p = .33.

7th edition-2023

Replication and reproducibility

In psychology precision of the sections in this book, this is prohibity one of the sections in this book, this is prohibity one of the psychologius were analytic orginate some well-known effects and funding that were considered notaxe and reliable prohibits were enalted by the sections in the section of t ist important you war near. There is were considered to near a structure (state in transposine were considered to near a structure (state in transposine were considered to near a structure (state in transposine were considered to near a structure (state ware) and transposine were considered to near a structure (state in the structure) and the structure (state is and transposine ware) and transposine were considered to near a structure (state is a structure) and the structure (state is a structure) (state to near a structure) (structure) (structure)

mands, as an do from what is known, phenomeno here also been found are security from the constraints of the down and the

loosely; aspects of the method might be modified, such as the

<text><text><text><text><text>

<text><text><text><text><text> The second example has a rechas been activated). Early Labs 4 project, found Replication failure in psychology and classic and contemporary affect disclosings

Moderator variables

of 46 per cure – with 51 out of 112 mmd ton et al. 2021). Of the 1,000 most wide sites, only 44 per cent twee found to re Stary per cent of 35 strades fielded to rep band 62. Verner, 1991), 40 per cent in a 64. Winnner, 1981), 40 per cent in ment, finance, economics and marketic 1996). The situation is somewhat better

Banks et al. 2019. Banks of 2019. Banks to eque data is open access that is, marking years marker than the second second second second second second second second second paper arailable nultur for anyons to read. "Wave a jurnal marker than following principles dynamic linear that the second second second second paper arailable nultur for anyons to read. "Wave a jurnal marker than following principles dynamic linear that the second second second second read of the second read to the article's paper can download or read the increasion of the second read to the second I ends to be pure a densero of the end y given Hamilton and the strength of the matrix of the strength of the strength of the matrix of the strength of the s

at be regarded as correct, sporting arrors in davign, rets of a wady. The issue g the pundemic when very

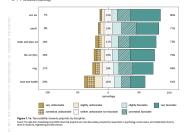
The present quark large transfer spectra of the present spectra of t

2010) and forecasting (32-maschinky & Armstrong, 2010). These, and similar results, suggest that some research my se susceptible to publication bias or Questionable Research Individual (2014). This is a superior for high part of the superior activity (2014). Which leads to promoving it D Publics. The CVPID-10 and Initiativity bias confirming more than well-accounting and the superior activity of the superior activity (2014). Which leads to promoving it D Publics. 2014 Annue and the superior activity (2014). Subd. (2014). Dist. (2014). Dist. 2014 Annue and the superior activity (2014). Subd. (2014). Dist. (20

r. QRPs which are historically well-known of simply poor design and method; for

Research practices	
p-hadong	 Andrybring data sets well a significant result is found (multiple testing) also known as data-divelgate. Benoval of data to ensure p < 0.6 Conducting updatened statistical lests well a significant p value is found Benoval of conditionary groups from analysis in celefor to achieve a significant p value. Changing analysis stategy after the side has have been conduction analysis
Omitting data	Omission of null findings from the study of more participants until a significant result is found
HARGing or H (ypothesising) A (fter) the R (exults) are K (nown)	Data are reinterpreted according to what was found rather than what was predicted in the hypotheses; post-hoc interpretation of results
Publication bias	The tendency for journals to publish research reporting only statistically significant results; this can lead to a balance to publish similar studies which do not find significant results; i.e., where null finding user reported
Reporting 'trends'	Probability values above 0.05 reported as 'trends towards significance' or as 'marginally significan

14 The science of psychology



Do registered reports offer an improvement over the tradi- Lay people's predictions of replicability

10 total model of publicate generalized in the darky strates of pro- tone strates and the strates of the strates of the strates of the methylicate strates of the strates of the strates of the strates of processing strates of the strates of the strates areas a dimensional by the strates of the strates of the strates of processing strates of the strates of the strates areas a dimensional by the strates of the strates of the strates of processing strates of the strates of the strates of processing strates of the strates of the strates of processing strates of the strate	When experts may added to produce the extensions of our of the major regritations projects of the prace for practice the Replacements Property studies, Many Lab 2 and the size of the studies of the studies of the studies of the studies of the content and its per occur of mergedown structures with add the theories et al. 2011, Frendi et al. 2019, Cannerse et al. 2019 the studies of the studies of the studies of the studies of the next high performance of the studies of the studies of the next high performance of the studies of the studies for a studies of the studies of the studies of the studies of the first performance of the studies of the studies of the studies for a studies of the studies of the studies of the studies for a studies of the studies of the studies of the studies for a studies of the studies of the studies of the studies of the first performance of the studies of the studies of the studies for a studies of the studies of the studies. The studies of the studies of the studies of the studies of the studies of the studies of the studies of the studies of the studies. The studies of the studies of the studies of the studies of the studies of the studies of the studies of the studies. The studies of the studies of the studies. The studies of the st
---	--

10 | 1 The science of psychology

Error type	Most frequent errors (N = 50)
Data naming/definition ambiguity	86
incorrect version control	62
Miscalculated data processing/analysis	47
Inacourate data coding	45
Error with data entry	32
Material, record or data loss	80
Programming error	15
Inaccurate data transmission	41
Error in data merging and selection	21
Infrastructural or technological issue	17
Oversight in planning/assessment of the study	14
Inadequate/faality documentation	21
Participant assignment error	15
Improper data point connections	
inacounte reporting	7
Incorrect software or hardware configuration	5
Misoganized data or files	6
Non-compliance with rules	5
Insuch of athics	,

otlight on. . . Students' QRPs

linis to any evaliable study data 20% inis to any evaliable one/jets scripts 20%

info about index rearraductions 11%

about indep repeakations 11% 11 indep relaxitiess checks 12% COI disclosures 21% paramatoria or pre-analysis 10% paramatoria or

underco of the research 27%

eathors' previous work 30%

advor/instations 375

anonymous user comments 52%

solified endomeratest by users 597

Figure 1.1b Importance of information I

es of the preprint 27% inka (k.g. OROD, GoogleSchelar) 29% open acholarahip 30%

is a student, do you think you would be immune to com-	that participants were asked to judge as acceptable or couldn't
initing QBN:7 A study of 252 (mainly) German psychol-	reach ajugment (on a 1-5 scale) Studietts were more likely to
syn students, gandama and their upsychiast found that all	contact QRPs in data analysis, nather than in meansch data
props shared similar views about the acceptability of QBPs	Supervision' attitudes to QRPs had a direct influence on the
(bishea & Peter, 2018). Table 1.4 shows you the statements	students' bolieti.

Practice	Participant attitudes		Supervisor attitudes	
		M (SD)		M (SD)
Selectively reporting studies	199	2.18 (0.90)	121	2.03 (1.22
Deciding whether to exclude data after looking at the results	189	1.96(0.94)	119	2.13 (1.25
Changing or formulating new hypotheses after analysing the data	201	1.92 (1.00)	140	1.97 (1.21
Rounding off p values	196	2.02 (8:97)	120	1.97 (1.19
Claiming to have predicted an unexpected result	199	1.82(0.87)	134	1.76(1.06

27% 20% 24%

20%

28% 33% 28%

25%

important in aligney incomentary way important intervely

the unity would be unccossfully applicand. the unity descriptions. "When they we of 100, how confident they were in their and the coldence, accuracy was of per-tion unity descriptions. To Description in thus applications of the 22 million register, 12 from the SSRP were formally the theory they apply the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of the theory of the theory end to the theory of theory of the theory of the t

Table 1.4 Continued				
Table 1.4 Continue				
Practice	Participant attitudes		Supervisor attitudes	
		M (SD)		M (50)
failing to report all relevant conditions	201	1.65 (0.73)	129	1.66 (1.01)
failing to report all relevant dependent measures	190	1.79 (0.79)	127	1.75 (.97)
Tubifying data	202	1.08 (2.47)	163	1.07 (.39)
falsely claiming that results are unaffected by demographics	188	1.57 (0.74)	126	1.52 (.86)
Collecting more data in order to achieve significance	199	2,21 (1.09)	128	2.34(1.32)
Stopping data collection after achieving the desired result	199	2.04(0.94)	115	1.70 (.96)
Reporting effect sizes	192	4.61 (2.74)	137	464(79)
Conducting a power analysis	160	4.47 (2.65)	99	4.37 (.05)
Uting Dayesian analysis	67	3.57 (1.02)	37	3.3*(1.35)
Utilizing sequential analysis	100	2,23 (1.06)	52	1.97 (1.25)

69X

208 198

168

16 | 1 The adiance of psychology means, integrity, and reprodu Why did you study psychology?

You've pioneered the way psychologist: do and think about psychology with the Many Labs projects - what motivated you to start this project?

computer engineering student tohed to psychology and worse

Dr Brian Nesek is Prefesser of Psychology at the University of Virginia, USA. He on-founded the Society for the Inprovement of Psychological Science (http:// impervingspcch.org/), and the Center for Open Science (OOS), http://co.id) watons aim is to improve rigour, transparency, Integrity, and reproducibility across research disolytics. He established the Mary Labs series of ity of the published literature. I have always been making science as a system as effective as possib

Inc. Be exhibited the Mary Lubs series of jets in psychology and also co-developed what were the easiest and the most difficult things about setting up difficult things about setting up Many Labs? you tell us a bit about you

In you tell us a Dreadout your own of the sckground, and what you do? rea up in California advent to Cal Poly. San Luis Obispo for a Walk of the standard of the standard school at Walk sock together on these huge projects. The harde managing the social reaction to what we found in Labs projects. The findings suggested that we ho net the downserved of regist meansure, transa sandhar downserved the present own data for sandhar and the present own data for sandhar The sandhar and the downserve to register and the present down data for the present present suit of broken downserve to register and the present down data for the present the downserve downserve register and the present down data for the present the downserve downserve register and the present down data for the present the downserve downserve register and the present downserve downserve the downserve downserve register and the present downserve do



were mixed: five were replicated, eight were pondents gave less clear/accurate predictions

thedr and Pateries in Psychological sibility of publishing, registered psychologicalscience, org/ arguing, hould emphasise finding that e hould emphasise finding that the section has highlighted eral problems in publication of science. Researchers and see hould emphasise finding that the section has highlighted and problems problems in the section has highlighted eral problems in publication of science. Researchers and see hours, at least, began to recognise and adams and section has highlighted eral problems in publication of science. Researchers and see hours of section has highlighted eral problems in publication of science. Researchers and see hours of section has highlighted eral problems in publication of science. Researchers and see hours of section has highlighted eral problems in hours of section hours highlighted eral problems in hours of section hours of section hours of section hours hours of section hours of section hours of section hours of section hours hours of section hours of section hours of section hours of section hours hours of section hours of section hours of section hours of section hours hours of section hours hours of section hours of sect erhaps half-hazerdelly, although there continues and for a special issue a call for a special issue and for a special issue and for a special issue the spec

The way forward: solutions

cjennal. Accountifully in Ream⁴/₃ is C "Connec. 2019, Vicane-Ease 1: on and mixial analysis of systems for conduct of research' (although that institution to referse DDA's sequence that the system of the system of the system of the system 2021]. Psychology Papar its own 1 2021]. Psychology Papar its own

as contractly applications, the density α_{ij} as the system of the strategy contract, α_{ij} and α_{ij} and sost-noc data analysis selection; ithe results, conditions, experi- 'Open data' is a simp Simmons et al (2011), for exam- fully into practice. I Simmons et al (2011), for exam-naily a logically derived finding to be utrefty bogis depending on researchers may be estir decisions made by the experi-bet listening to children's song et als whether listening to song the solution of the song the solution of the song the solution of the solution of

and another based must be determined by the regime to the second second

ORIGINAL RESEARCH article

Front. Psychol., 11 April 2017 Sec. Quantitative Psychology and Measurement Volume 8 - 2017 | https://doi.org/10.3389/fpsyg.2017.00523

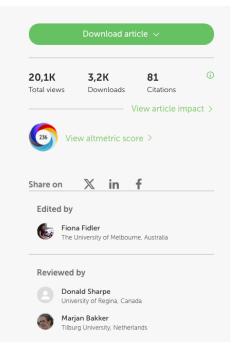
Are Psychology Journals Anti-replication? A Snapshot of Editorial Practices



Richard M. Clarke²

¹ School of Psychotherapy and Psychology, Faculty of Humanities, Arts and Social Sciences, Regent's University London, London, UK
 ² Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

Recent research in psychology has highlighted a number of replication problems in the discipline, with publication bias – the preference for publishing original and positive results, and a resistance to publishing negative results and replicationsidentified as one reason for replication failure. However, little empirical research exists to demonstrate that journals explicitly refuse to publish replications. We reviewed the instructions to authors and the published aims of 1151 psychology journals and examined whether they indicated that replications were permitted and accepted. We also examined whether journal practices differed across branches of the discipline, and whether editorial practices differed between low and high impact journals. Thirty three journals (3%) stated in their aims or instructions to authors that they accepted replications. There was no difference between high and low impact journals. The implications of these findings for psychology are discussed.





CONVERTING THIS INTO A CLASS ACTIVITY

What is replication?

Why is it important?

What are the Open Science Framework and Many Labs and how have these contributed to our understanding of how psychology research has been conducted and published?

What is the difference between a direct and conceptual replication?



SEMINAR/CLASS ACTIVITY: STATE/CONTEXT DEPENDENT MEMORY

Compare and contrast Godden & Baddley's (1975) scuba diving study with a "replication" by Murre (2021; 2022)

 Ask the students to describe what Godden and Baddeley did
 Then, ask them to compare that study's method with those of the Murre study.

3. How do they differ?

4. In what way is the latter study not a direct nor a conceptual replication?

5. What are the issues with describing this study as a replication?



SEMINAR/CLASS ACTIVITY: STRACK ET AL (1988)- FACIAL FEEDBACK HYPOTHESIS



Figure 13.13 Examples of the facial expressions created in Strack et als experiments when people held a pen between the lips or the teeth- note the position of the muscles.

Source: Strack, F., Stepper, S. and Martin, L. L., Inhibiting and facilitating conditions of the human smile. Journal of Personality and Social Psychology, 1988, 54, 768–777. © Fritz Strack. Reproduced with permission. https://psycnet.apa.org/record/1988-25514-001

You can repeat this study in the session using a selection of cartoons and rating scales

A failure to replicate the result of Strack et al directly in 17 laboratories can be found here: https://journals.sagepub.com/doi/full/10.1177/1745691616674458

Discussion of the controversy here which suggests that variations in procedure can affect whether the finding is replicated -<u>https://psycnet.apa.org/record/2018-16714-001</u>



CLASS ACTIVITY: QUESTIONABLE RESEARCH PRACTICES

- Match the term with the definition
- The exercise provides opportunity to discuss the nature of QRPs, why they are committed and the consequences of committing them.
- You could refer to case studies in psychology, such as Diederik Stapel and Brian Wansink- and, possibly, Eysenck & Grossarth-Maticek work on cancer-prone personality.
- cf Gino, Ariely.
- Fabrication in a study about honesty
- https://www.informahealthcare.com/doi/metrics/10.1080/08989621.2024.2329265



SEMINAR/CLASS ACTIVITY: QUESTIONABLE RESEARCH PRACTICES

- A. p-hacking
- B. Omitting data
- C. Additional recruitment
- D. HARKing or H (ypothesising) A (fter) the R

(esults) are K (nown)

- E. Publication bias
- F. Reporting "trends"
- G. Small samples



1. The tendency for journals to publish research reporting only statistically significant results; this can lead to a failure to published similar studies which do not find significant results, i.e., where null findings are reported

2. Analysing data sets until a significant result is found (multiple testing); (also known as data-dredging); removal of data to ensure p<.05; conducting unplanned statistical tests until a significant p value is found; Removal of conditions/groups from analysis in order to achieve a significant p value; changing analysis strategy after the data have been <u>collected/ analysed</u>

3. Insufficiently powered studies leading to inappropriately applied statistical tests

4. Leaving out null findings from the study

 Recruiting more participants until a significant result is found
 Data are reinterpreted according to what was found rather than what was predicted in the hypotheses; post-hoc interpretation of results

7. Probability values above .05 reported as "marginally significant"

How might psychologists prevent questionable research practices and improve replicability? What is the nature and purpose of a registered report?

How they differ from other types of report?

Why have psychologists chosen to adopt this as one model of publishing their data? How might students <u>adopt</u> registered reports in their own study? Would this be practical? What would be the obstacles to adopting this approach?

What is open science and open data? Do students believer there are negative aspects to open data, as well as positive, and if so what are they and how might these be addressed or mitigated? You can refer to national research concordats here, which stipulate that institutions such as Universities adopt open science and open data practices (such as making data and code freely available)

KRISHNA & PETER (2018): QRPS

Practice	<u>]</u>
Selectively reporting studies	
Deciding whether to exclude data after looking at the results	
Changing or formulating new hypotheses after analyzing the data	
Rounding off p values	
Claiming to have predicted an unexpected result	
Failing to report all relevant conditions	
Failing to report all relevant dependent measures	
Falsifying data	
Falsely claiming that results are unaffected by demographics	
Collecting more data in order to achieve significance	
Stopping data collection after achieving the desired result	
Reporting effect sizes	
Conducting a power analysis	
Using Bayesian analysis	
Utilizing sequential analysis	

Ask students to respond on a five point scale whether they think these practices are acceptable (where 1= unacceptable).

•You could also ask students how they think their tutor/supervisor might respond, using the same scale.



SEMINAR/CLASS ACTIVITY: DECEPTION!

- •A test of students' knowledge about deception and the behavioural cues to deception
- •Ask the students to indicate what verbal and non-verbal behaviour they think would identify a liar.
- •Are there any consistent cues, according to them, that liars use/ exhibit when they lie?
- •Are there any verbal and non-verbal cues that signify truth-telling? Ask them why they think these cues are indicative of deception (and truth-telling)?



ARE THESE INDICATIVE OF PEOPLE NOT TELLING THE TRUTH & LYING?

Cues

Vocal

Hesitations (use of speech fillers, e.g. 'ah', 'um', 'er', and 'hmmm' Speech errors (grammatical errors, word or sentence repetition, false starts, sentence change, sentence incompletions, slips of the tongue, etc. High-pitched voice Speech rate (number of spoken words in a certain period of time) Latency period (period of silence between question and answer) Pauses (silent, filled or mixed) Visual Gaze aversion (looking away from the conversation partner) Smiles (smiling and laughing) Facial fidgeting (face touching or rubbing hair) Self-fidgeting (touching, rubbing, or scratching bady or face) Fidgeting (undifferentiated) Illustrators (hand and arm movements designed to modify or supplement what is being said verbally) Leg and foot movements Posture shifts (movements made to change seating position) Head movements (head nods and head shakes) Eye blinks (blinking of the eyes)

Cues	Actual relationship	Assumed relationship
Vocal		
Hesitations (use of speech fillers, e.g. 'ah', 'um', 'er', and 'hmmm'	0.04	Associated with lying
Speech errors (grammatical errors, word or sentence repetition, false starts, sentence change, sentence incompletions, slips of the tongue, etc.	0.00	Associated with lying
High-pitched voice	0.21	Associated with lying
Speech rate (number of spoken words in a certain period of time)	0.07	No assumed relationship
Latency period (period of silence between question and answer)	0.02	No assumed relationship
Pauses (silent, filled or mixed)	0.02	Associated with lying
Visual		
Gaze aversion (looking away from the conversation partner)	0.03	Associated with lying
Smiles (smiling and laughing)	0.00	No assumed relationship
Facial fidgeting (face touching or rubbing hair)	.08	Associated with lying
Self-fidgeting (touching, rubbing, or scratching bady or face)	-0.01	Associated with lying
Fidgeting (undifferentiated)	0.16	Associated with lying
Illustrators (hand and arm movements designed to modify or supplement what is being said verbally)	-0.14	No assumed relationship
Leg and foot movements	-0.09	Associated with lying
Posture shifts (movements made to change seating position)	0.05	Associated with lying
Head movements (head nods and head shakes)	-0.02	Associated with lying
Eye blinks (blinking of the eyes)	0.07	Associated with lying

Note. Data on actual relationships between cues and lying is taken from DePaulo et al (2003). Data on assumed relationships is taken from Vrij (2008). Positive scores indicate increase in liars; negative-scores indicate decrease in liars; significant relationships are indicated in bold.



SEMINAR/CLASS ACTIVITY: FACE RECOGNITION HOW MANY DIFFERENT FACES CAN YOU SEE? (JENKINS ET AL, 2011)



QUESTIONS

- Why were they unable to identify the correct number?
- What does this say about our ability to identify faces accurately?
- Do they think that facial recognition software would perform better?



FACE RECOGNITION: SOME SIMPLE EXPERIMENTS

- Present students with images of facial features (eyes, noses, mouths, chins) cut out from magazines and newspapers; ask them to identify the famous faces from which they were cut. Which features provide the best clues to the identification of the individual?
- Present an unfamiliar face briefly as a target and ask students to spot this target face from within a collection of other unfamiliar faces presented serially. How many people are able to identify the face correctly? How did they remember the face? What features did they look for as they searched through the distracter faces?
- Present unfamiliar faces (preferably without the hair, which would provide an additional clue) very briefly and ask students to identify the sex/age of each one. How accurate are they? Are students better able to identify the sex/age of faces of their own ethnic origin than of faces of different ethnic origins?
- Present famous faces briefly, some in the correct orientation and some upside down. Is
 identification as accurate when the faces are inverted as when they are the right way up?
 Why is this?



SEMINAR/CLASS ACTIVITY: Beliefs about memory

 Table 8.4
 Percentage of respondents agreeing with

 each memory statement, along with the average rate of
 agreement across items.

Table 8.5The mean number of endorsements formetaphors of memory used in Brewin et al's (2019) study(7 = strongly agree).

		SurveyUSA	Diary entry	5.:
Item	(%)	(%)	Video camera	5.
Amnesia: People suffering from amnesia	81.4	69.6	Library	5.
ypically cannot recall their own name or dentity.			Rooms in a house	4.
Confident testimony: In my opinion, the tes-	22.1	32.9	Storehouse	4.
mony of one confident eyewitness should		52.9	Computer storage	4.
e enough evidence to convict a defendant			Painting	4.
f a crime.			Collage	4.
ideo memory: Human memory works	46.9	52.7	Made-up story	4.
ke a video camera, accurately recording			Perfume	4.
ne events we see and hear so that we can eview and inspect them later.			Artist's impression	4.
nexpected events: People generally notice	77.4	65.0	Jigsaw puzzle	4.
hen something unexpected enters their			Melodies on a piano	4.
eld of view, even when they are paying tention to something else.			Filing cabinet	4.
0	20.0	20.0	Muscle	4.
ermanent memory: Once you have experi- nced an event and formed a memory of it,	28.0	39.9	Lock and key	4.
hat memory does not change.			Subway map	3.
ypnosis: Hypnosis is useful in helping wit-	46.4	44.6	Conveyor belt	3.
esses accurately recall details of crimes.			Aviary	3.
verage agreement rate (out of 6 items).	50.33	50.83	Compost heap	3.
	(3.02)	(3.05)	Bottle	2.

Note: SurveyUSA data are from Simons and Chabris (2011).

doi: 10.1371/journal.pone.0051876.t003



•When participants presented with:

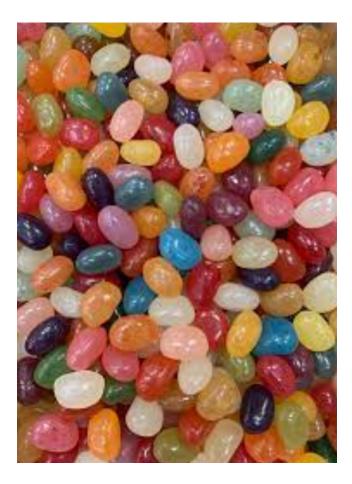
'Human memory works partly like a video camera, accurately recording some of the events we see and hear so that we can review a simplified version later' and 'Human memory is not like a video camera because we cannot play back events exactly as happened'

they were less likely to accept the video camera metaphor (Brewin et al, 2019)

Importance of wording in influencing people's beliefs- discussion of validity



SEMINAR/CLASS ACTIVITY: THE JELLY BEAN TEST





SEMINAR/CLASS ACTIVITY: THE JELLY BEAN TEST

A good way of showing students the difference between <u>retronasal</u> and <u>orthonasal smelling</u> -and the importance of olfaction to <u>flavour</u> is the Jelly Bean test.

For this you will need some physical items: a packet of differently flavoured jelly beans (with a map on the back identifying which colour is which <u>flavour</u>- this is very important), a blind-fold (option) and nose clip (optional). You can also run this exercise remotely by asking students to prepare in advance and obtain the jelly beans. The instructions below are the same.

- 1. Ask the student to put on a blind-fold (or to close their eyes)
- 2. Ask them to pinch their nose (or apply the nose clip)
- 3. Ask them to open their mouth and place a jelly bean on their tongue (the student must not be able to see the colour nor to breathe through the nose)
- 4. Ask them to chew the bean and try to identify the flavour
- 5. Note the flavour
- 6. Remove the nose clip or ask them to release their fingers and ask them to identify again
- 7. Note any differences between the two.
- 8. Repeat for as many flavours as you want

How good or poor were they when their nose was pinched? Discuss with the students why they identified in the way that they did, using the exercise to illustrate the function of <u>retronasal</u> smelling. You could bring in real-life examples, such as having a <u>cold-which</u> impairs olfaction but leaves gustation relatively intact (you can taste food but not its odour)





Journal of Applied Research in Memory and Cognition



Volume 1, Issue 4, December 2012, Pages 242-248

Target Article

Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice

Henry L. Roediger III 🙎 🖾 , Mary A. Pyc

Optimizing Learning in College: Tips From Cognitive Psychology

Adam L. Putnam¹, Victor W. Sungkhasettee², and Henry L. Roediger, III² ¹Department of Psychology, Carleton College, ²Psychological & Brain Sciences Department, Washington University in St. Louis



Perspectives on Psychological Science 2016, Vol. 11(5) 652–660 © The Author(s) 2016 Reprints and permissions: sagepub.com/journals/Permissions.nav DOI: 10.1177/1745091616645770 pps.sagepub.com SAGE



JUST *SOME* IDEAS

•INSTRUCTOR'S MANUAL accompanying textbook has over 300 of these types of exercises

•PLUS MCQs & powerpoint slides for each chapter



NEW IN THE 7TH EDITION

- All 17 chapters fully revised
- 57 Spotlight on...sections highlighting new and/or unusual research in psychology
- 26 Controversies in Psychological Science sections
- 39 Psychology in Action sections
- 36 '... an international perspective' sections
- Interviews with 14 psychology practitioners or leading researchers in psychology, including educational and clinical psychologists, a member of the UK SAGE behaviour sub-committee, the co-ordinator of a fear lab, the training manager for the National Autistic Society, the co-founder of an online psychology experiment business, a psychologist who runs fMRI experiments, and the co-founder of the Open Science Framework and ManyLabs
- 1017 new references
- All new further reading/viewing suggestions
- A major new section on replication in psychology in Chapter <u>1-the</u> issues, the problems and possible solutions- and an interview with Professor Brian <u>Nosek</u>
- The mental disorders chapter fully revised to include ICD-II

NEW IN THE 7TH EDITION

- · Expanded sections on topical and psychologically relevant topics including
- vaccine acceptance
- attitudes to climate change
- the rise of fake news and how to fight it
- nudging and its problems
- conspiracy theories
- · face recognition technology and its use and effectiveness
- the use of phonics to teach reading
- · sex differences and the brain
- machine learning
- screen and social media use in teenagers
- predicting your personality from your digital footprint
- how behaviour and sensation changed during COVID-19 and lockdown.

Thank you for listening!

Practical ideas for introducing new & controversial research to the introductory psychology programme

Professor G. Neil Martin, FRSA NeilOnComedy@outlook.com

