Opening up the science:

Methodological transparency and reproducibility in second language acquisition research

Emma Marsden

University of York, UK

BRITISH ACADEMY

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

THE UNIVERSITY of

Towards Methodological Transparency & Reproducibility in SLA research

1. Open science & the IRIS digital repository of materials
2. Relationship between transparency and quality
3. Replication & the importance of transparency
4. Challenges and Recommendations

The sticks and carrots of open science

ublicly funded research **should** be made freely available

PEN PUBLICATION

> cited more

Wagner, A. B. (2010).

PEN DATA

- Corpora of oral production data, L1 and L2: FLLOC, SPLLOC, CHILDES
- Trønso
- CLARIN

> increased citations

Piwowar & Vision (2013)

> stronger evidence and better statistical reporting

Wicherts, J., Bakker, M., Molenaar, D. (2011).

OPEN METHODS

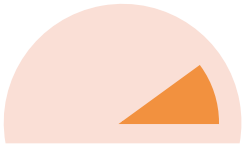
Making materials available

Materials = Data collection tools, instruments, stimuli, scoring and coding procedures, analysis protocols

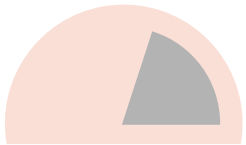
Part 1: Open science: METHO

aterials design for L2 data collection: a lonely, mysterious business

- Some problems:

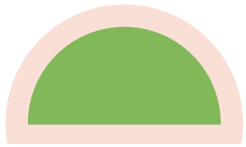


Researchers create and keep own instruments
-> re-invention of wheel, poor systematicity of research



Maintenance and access to instruments is *ad hoc*

Publications just have brief descriptions with short samples



What a contrast with full instrument!

How much of the methods do we see in journals?

(17) Experimental: *[−past] *shì* ... *de* (4 tokens)

*Xiǎo Wáng	shì	míngtiān	chī-wán	nà	gè	dàngāo	de,	bú	shì	jīntiān.
Xiao Wang	COP	tomorrow	eat-up	that	CL	cake	DE	not	COP	today

Intended: It is tomorrow that Xiao Wang (is going to) eat up that cake, not today.

(18) Control A: [+past] *shì* ... *de* (4 tokens)

Xiǎo Wáng	shì	zuótiān	chī-wán	nà	gè	dàngāo	de,	bú	shì	jīntiān.
Xiao Wang	COP	yesterday	eat-up	that	CL	cake	DE	not	COP	today

It was yesterday that Xiao Wang ate up that cake, not today.

(19) Control B: [−past] canonical (4 tokens)

Xiǎo Wáng	dǎsuàn	míngtiān	chī-wán	nà	gè	dàngāo.
Xiao Wang	intend	tomorrow	eat-up	that	CL	cake

Xiao Wang intends to eat up that cake tomorrow.

Mai, Z., & Yuan, B. (2016). Uneven reassembly of tense, telicity and d
features in L2 acquisition of the Chinese shì … de cleft construction b
English speakers. *Second Language Research,* 32(2), 247-276.

IRIS
A digital repository of data collection instruments for research into second language learning and teaching

The University of York
Georgetown University

Part 1: IRIS:
Transparency of metho

BRITISH ACADE

What is IRIS?

Instruments for Research Into Second Languages

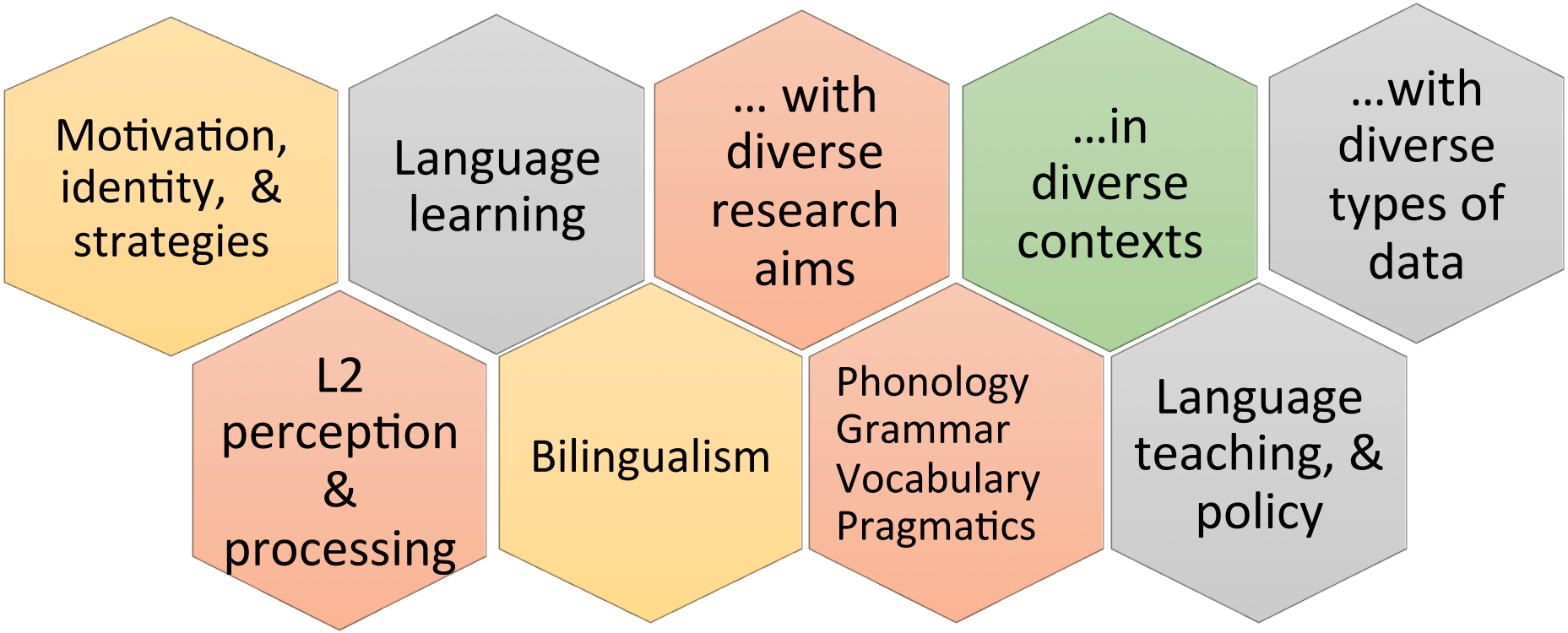
- **A sustainable digital repository**
- **2200+ materials used to collect data**:
  - e.g. questionnaires, grammaticality judgment tests, observation & interview schedules, word lists, sound & video files, language tests, pictures, teaching materials, software scripts,
  - **And ANALYSIS PROTOCOLS and DATA**
- **Downloadable**
- **Uploadable**
- **Searchable**

my Research Project' AN110002

Marsden, E., Mackey A., & Plonsky, L. (2016). The IRIS Repository: Advanc
research practice and methodology. In A. Mackey & E. Marsden (Eds.),
Advancing methodology and practice: The IRIS Repository (pp. 1-21). Rout

Scope of content

as wide as the field of L2 research…



Rationale behind IRIS

- Materials can be adapted to suit different contexts, learners, languages
- Promotes transparency
- Easier to evaluate quality and reliability
- Quality assurance: only peer-reviewed publications
- Stimulates & facilitates replication

- 2400 files -> **over 1100 data collection instruments**
  - new materials contributed almost daily
- From 1300 researchers
- Searchable across 330+ parameters,
  - ✓research area
  - ✓type of instrument
  - ✓language feature
  - ✓L1, L2
  - ✓participant characteristics: age, proficiency
  - ✓author
- Materials qualify for IRIS if used for:
  ***peer-reviewed* publications**
  or PhDs

Part 1: IRIS: Transparency of ME
Marsden, Plonsky & M

- Submission to IRIS supported by 37 journals
  - In acceptance letters from editors to authors, and in author guidelines

- Upload to IRIS in publication guidelines of the *American Association of A*
- Formally endorsed by *British Association of AL*

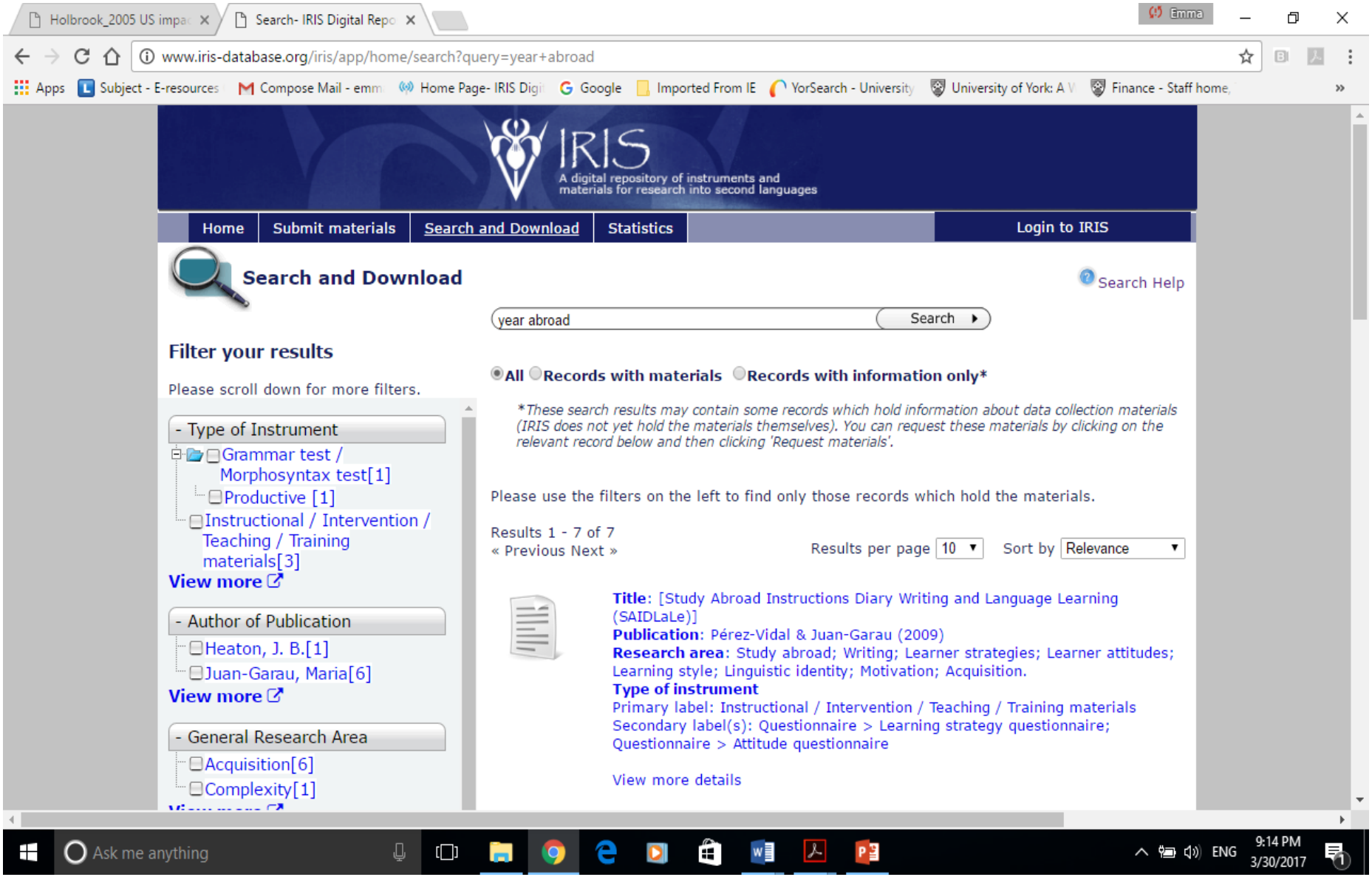
- Only venue in AL that qualifies articles for Center for Open Science badge

- 16,500+ downloads of research materials

- Cited in handbooks, methods books, position pieces & syntheses

- Used in many research methods training courses

Materials used for ‘year abroad’ research?



art 2: Open methods to improve rigour and replicability…

The "methodological turn" in applied linguistics

"**methodological issues** …demand a kind of **professional scrutiny** that goes directly to the core of what we do and what we know…"

Byrnes, 2013, p. 825

"Methodological practices and study quality **need to be measured, not assumed**"

Plonsky, 2013

Benefits of transparent instrument design

Part 2a: Grammaticality judgment tests

Indicate how acceptable the following sentences are

completely unacceptable			***completely acceptable***
1	2	3	4

1) I don't know what is he eating for dinner tonight.
2) I don't know what he is eating for dinner tonight.

equent use of judgment tests because *assumed* to be…

Easy to…

- develop
- administer
- score

Common conventions

→ greater comparability across studies

JTs are controversial

Tests of ***explicit*** or ***implicit*** knowledge?
(e.g., Ellis, 2005; Gutierrez, 2013; Vafaee et al., 2016)

Theoretical perspectives on linguistic knowledge-> methodological decisions

- Do instructions say 'how *acceptable*…' or 'how *correct*…' ?
- Scaled vs. dichotomous response
- Oral vs. written
- …

Methodological synthesis of JTs

How do we design, administer and report on JTs?

K = 299 studies

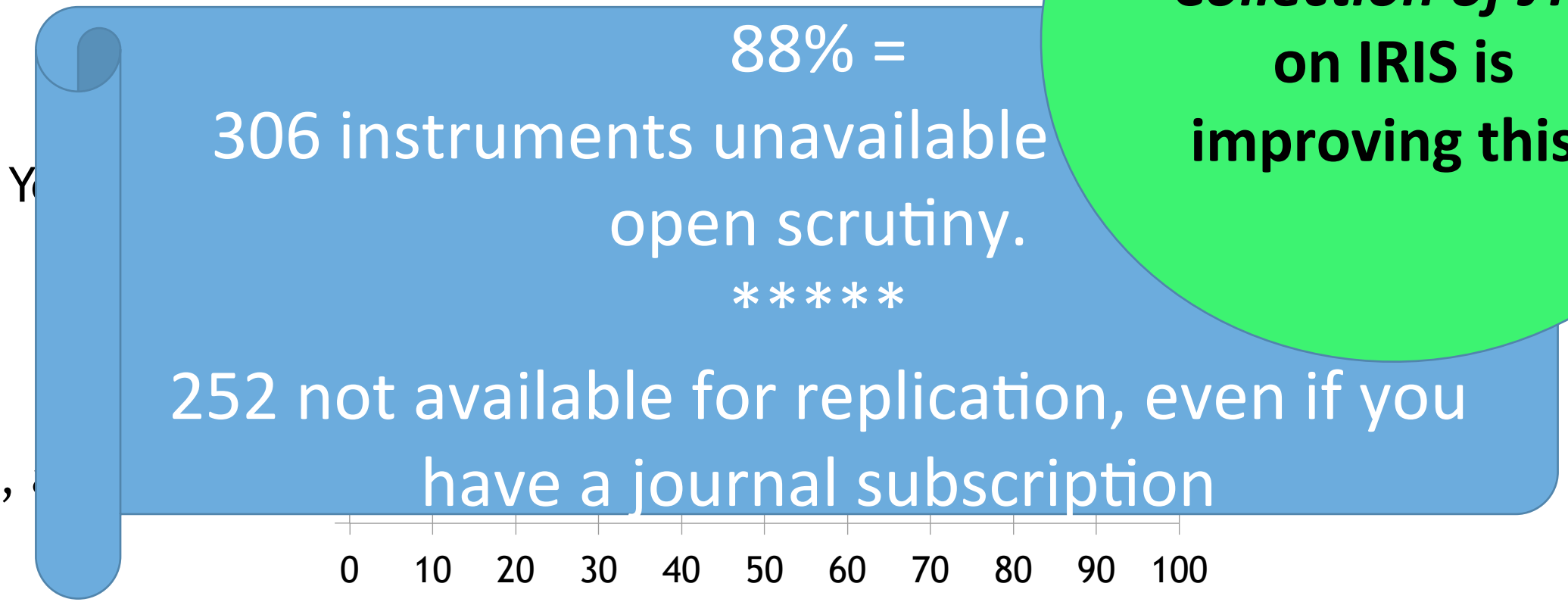
382 JTs

Total sample size in our synthesis = 24,679 !

art 2a: Instrument transparency: **JTs.** *Plonsky, Marsden, Gass, Spinner, Crowther (in prog*

Transparency?

% of JTs available



A *'Special Collection of JTs'* on IRIS is improving this!

Part 2a: Instrument transparency: **JTs.** *Plonsky et al. (in prog*

Consequences

- Median *N whole study*
  - L2 learners = 47
  - NS controls = 20
  - Overall study =60
    - Field median = 60 (Plonsky, 2013)
- L1s
  - 45% = English
  - 15% = Chinese
  - 9% = Japanese
  - 8% = French
- TLs
  - English = 59%
  - Spanish = 17%
  - French = 10%
- 77% one-shot design -> non-developmental, not within-subject over time

Imagine the '*n*' if JTs available across the globe?

Think of the L1 – L2 combinations!

Existing JTs used to investigate change over time, e.g. after teaching or year abroad

Transparent design -> clearer operationalisation of knowl e type

-> the reader has to *infer* what kind of knowledge the researchers elicited …
without full reporting of design features,
without <u>*seeing*</u> the instrument!

Consequences of lack of full transparency: Construct validity

Time pressure matters.

52% we don't know whether timed or not

Modality matters.

18% we don't know whether written or aural

Breadth of construct - types and tokens of linguistic feature tested?

51% we don't know if there was only one 'version'

If one version, narrows breadth of construct being elicited

art 2a: Instrument transparency: **JTs.** *Plonsky, Marsden, Gass, Spinner, Crowther (in prog*

Conseq[illegible]ncy: **Reliability**

Reliability coefficients can be given alongside instruments on IRIS.

New coefficients posted every time that JT used?

[illegible]efficient

[illegible] in SLA (Plonsky, 2013)

0.80 (=field average, Plonsky & Derrick, 2016)

[illegible]e instrument

[illegible] ungrammatical

Part 2b: Instrument transparency: Self-paced reading

Similar aim to JTs

Eliciting sensitivity to grammatical structure and norms

But, SPRs identify **precise point** of difficulty *during* parsing,

without an explicit judgment

Self-paced reading: an exam

Try to understand this sentence.
Click for each word to appear.
You will be asked a question at the end!

I

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

don’t

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

know

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

what

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

is

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

she

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

having

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

for

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

dinner

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Self-paced reading: an example

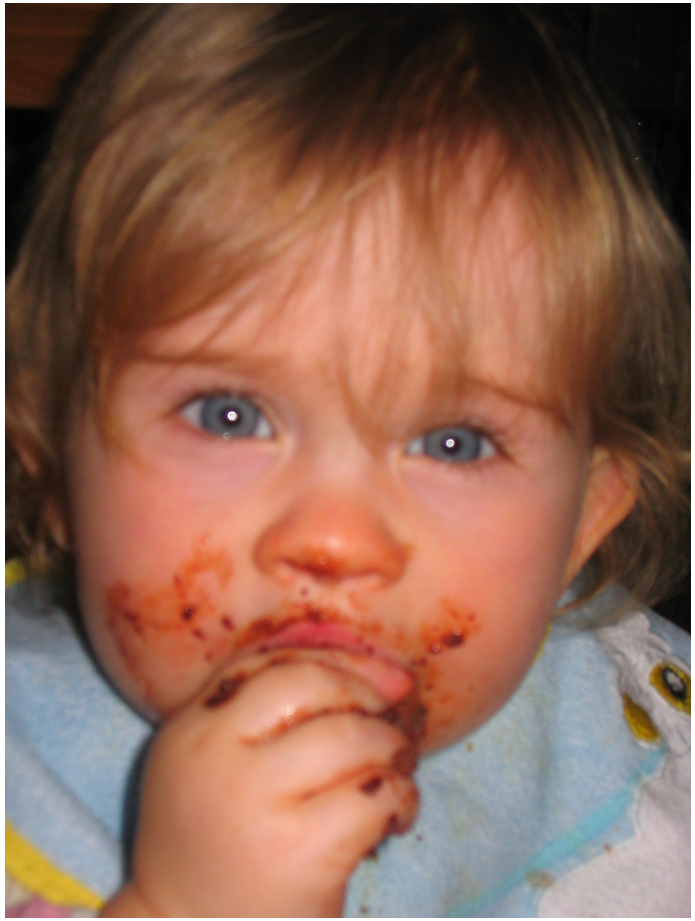
tonight

Stimuli taken from Marsden, Cruickshank, Roberts (in progress) *Can we train processing routines for abstract synt*

Which picture best matche

The word ‘dinner’ - **after** the critical region

A



B

Transparency of instrument design: SPR

Methodological synthesis

63 studies

reporting a total of 71 SPR tests

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Transparency and reproducibility of SPR tests in SLA research

Availability	No. of studies
Full stim[illegible]	3
Another	
Exa[illegible]	4
Full	16
Ju[illegible]	28
No[illegible]	16

96% of <u>full</u> materials not yet available on <u>permanent</u>, <u>open</u> repository

A '*Special Collection of SPRs*' on IRIS is improving this!

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

onsequences of poor availability: Agenda limiting - L2s?

Language	No. studies: L1	No. studies: target language
English	19	43
Chinese	9	2
Spanish	5	11
Greek	4	1
German	3	6
Korean	3	1
Dutch	3	2
Japanese	3	0
French	1	2
Russian	1	0
Arabic	1	0
Multiple languages	18	

Part 2b: Instrument transparency: SPRs. *Marsden, Thompson, Plonsky (subm*

onsequence of poor availability:

genda limiting: Cross-linguistic influence?

We know the L1 matters for online processing

52 / 71 with learners with *one* L1

14 compared across different L1s

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Consequence of poor availability:

Agenda limiting. Who are the participants?

Beginner	Interm	Advanced	Near Native	Bilingual
7	18	50	3	8

56 / 71 = university students

- Higher meta-linguistic knowledge could affect reading times

(Keating & Jergerski, 2015)

Part 2b: Instrument transparency: SPRs. *Marsden, Thompson, Plonsky (subm*

Consequence of poor availability:

Agenda limiting. Wh mprehension only?

To date, resear rehension is *high*

- advanced l
- only analy

But lea comprehension is poo

Existing SPRs could be used with other learner populations

ting, Jegerski & VanPatten, 2016; Xu, 201

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Consequence of poor availability:

Agenda limiting. Inv… …over time?

Growing int…

(Altmann… …& Garod, Cha… et a… *Approaches… …gualism, 5,*

YET, S…

7…

24%… …ble

One a wit…

Existing SPRs could be used in longitudinal designs

e.g. before, during, after year abroad

Though see: McManus, K., & Marsden, E. (2016). L1 explicit instruction can impro… L2 online and offline performance. *Studies in Second Language Acquisition*, 1-…

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Consequences of lack of transparency:

Construct validity - Comprehension of *what*?

Central tenant of SPR: processing *during reading comprehension*

Comprehension question u

Question should not

56 / 71 used CQs

BUT what do these CQs fo

17 – no example

34 - *one* example

One example of a CQ does not tell u **where** attention is repeatedly focussed during reading

Only 5 studies provided multiple examples of comprehension questions

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Consequences of lack of full transpar ncy: Construct validity

Stimuli design (*k* items wor equency, length)

Length matters: affects sier, 2011).

Yet…

- 25 / 71 length
- 5 / 71 re
- 19 / 29 using se ot specified

A bank of stimuli helps instrument design and research training

Part 2b: Instrument transparency: **SPRs.** *Marsden, Thompson, Plonsky (subm*

Consequence of lack of transparency:

Data cleaning and analysis protocols

Outlier removal?

- 12 did
- 44
- 15

Out

- 14
- 5 u
- Wi

er).

Having analysis protocols available would make analyses more systematic and reproducible

Analyse all data or just correct responses?

- 7/56 studies analysed all
- 28/56 analysed only correct responses
- 21/56 didn't report

Summary of Part 2:

Greater transparency and

Agenda setting and research aims

Scope and power of studies (differ

Operationalise our constructs (des

Comparability across studies (data

‘Better reporting’ unlikely to fully address all these problems

(across all journals & techniques…)

The actual materials necessary

More transparency and systematicity of methods ->

Benefits for all types of validity & reliability

Part 3: Replication research in SLA

"Conducting a research study again, in a way that is either identical to the original procedure or with small changes (e.g., different participants), to test the original findings" (Mackey & Gass, 2005: 364).

"essential … support for theory" (Porte, 2012)

Strong replication movement in Psychology

- "Many Labs" & Reproducibility projects
- ***Pre-registration*** *of materials & analyses*

Commentaries and calls for replication

46 published commentaries & calls for replication in L2 research

Santos 1989…Polio & Gass 1997… Porte 2012…Vandergirft & Cross 2017

23 + from other disciplines: Psychology, Education, Sociology, Business, Marketing, Organisation Science

Replications… lacking prestige, originality, or excitement

(Makel et al 2012, citing Lyndsay & Ehrenberg, 1993; Neuliep &

Other syntheses of replications:
for Psychology: Makel, Plucker & Hegarty (2012)
for Education research: Makel & Plucker (2014)

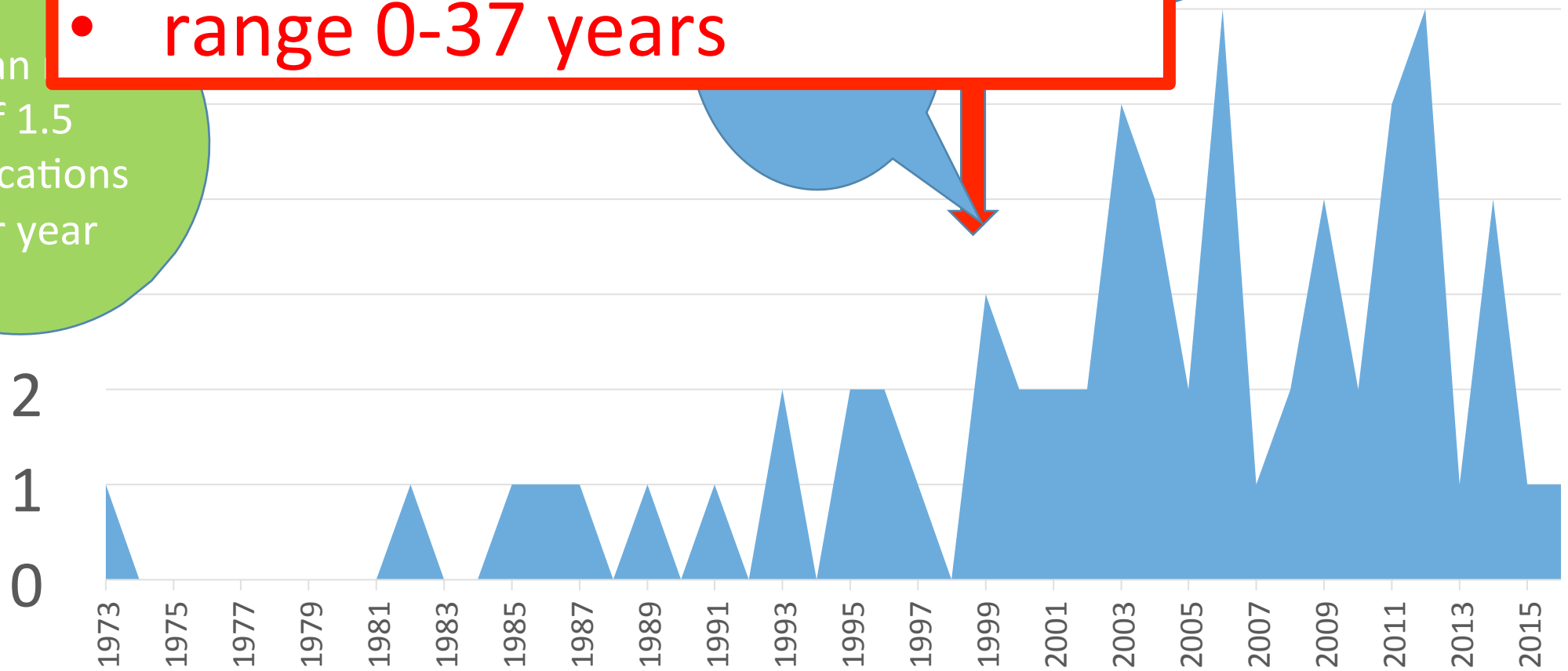
rep 0 initial studies

How long does it take for a study to be replicated?

- mean 6.64 years (sd 6.16)
- range 0-37 years

Norris & Ortega 2000?

Mean of 1.5 replications per year



roportion of replication in journals that published mo

	SSLA	*The MLJ*	*FLA*	*LL*	*AP*	Mean rate	
TOTAL articles	562	1009	1528	855	1030		
k replications	13	8	5	5	4		
% of total	2.31%	0.79%	0.33%	0.58%	0.39%	0.88%	**0.70%**

across the 26 journals that *have* published replications = **0.26%**

1973 – 2015, last complete year before synthesis

ompared to other disciplines?

.07% Psychology, but NB: pre-replication-boom & **top 100** journa

% - 3% Business, Marketing, Communication

.13% Education (**top 100**, 1938-2014)

What do we replicate? Study design and findings

ticipants? The **WEIRDest**

stern Educated Industrialized Rich Democratic (Mishra et al. 2012) …
speaking (1/4) or learning (1/2) English

'ages' and 'proficiency' ***not reported*** (67%, 37% studies) (Thomas 1994) -> replicability??

f replications whose initial findings were…

Null	Null but trend	Stat sig differences	other
3	3	83	7

need to replicate null findings too

when power (n) is low, "null findings" ≠ "no effect"

Schmidt, F.. & Oh, I.-.S. (2016) The Crisis of Confidence in Research Findings in Psychology: Is Lack of Replication the Problem? Or Is It Something Else? *Archives of Scientific Psychology* 4, 3

How much do we change when we replicate?

% studies with changes between I and R…

Number of changes	changes = reasons for replication *% studies*		
0	33		
1	28		
2	21		
3	9		
4	7		
5	1		
Mean per study (st dev.)	1.34 (1.31)		

How much do we change when we replicate?

% studies with changes between I and R…

Number of changes	changes = reasons for replication	changes acknowledged, but not reason for replication *% studies*	
0	33	45	
1	28	31	
2	21	13	
3	9	9	
4	7	1	
5	1	0	
Mean per study (st dev.)	1.34 (1.31)	0.91 (1.04)	

How much do we change when we replicate?

% studies with changes between I and R…

Numbe chang	s = reasons for replication	s acknowledged, but not reason for replication	changes not acknowledged uthors *tudies*
0			46
1	28	31	25
2	21	13	15
3	9	9	9
4	7	1	4
5	1	0	0
Mean per study (sd)	1.34 (1.31)	0.91 (1.04)	1 (1.18)

no relationship to what the replication calls itself (e.g., 'partial replication,' just 'replication', extension, conceptual)

Do “authorship overlaps” relate to whether findings are supportive of initial study?

% “supportive of Initial”, as function of authorship independence

	Findings in relation to initial study (% of replication studies)	
Author overlap? (% total R studies)	**% not or partially not supportive**	**% partially or very supportive**

How do relations h[illegible]s relate to the na[illegible]

Questionable Research Practices (bias, p-hacking)
Or
Materials availability and fidelity to initial study
??

sychology:

2% supportive if [illegible]

5% supportive if [illegible]

ducation:

9% supportive with overlap, in same publication

1% with overlap, in new publication

4% when no author overlap

How do studies use the *findings* and the *data* of the Initial study?

Only 6% of Replications provided Initial study's effect size.

Only 6% of Replications used Initial's raw data in a new analysis

Data Sharing!

Open data is associated with strength of evidence and quality of reporting

Wicherts JM, Bakker M, Molenaar D (2011) Willingness to Share Research Data Is Related to the Strength of th
Evidence and the Quality of Reporting
Statistical Results. PLoS ONE 6(11): e2682

How do Replicators compare their findings to Initial Study?

% of Replication Studies that compared their findings to Initial Study using…

% narrative Comparisons	% mentioned original findings	% dichotomous decision from Null Hypothesis Significance Test	% descriptive statistics	% unclear	% effect sizes
93	90	84	34	6	1

How do Replicators compare their findings to Initial Study?

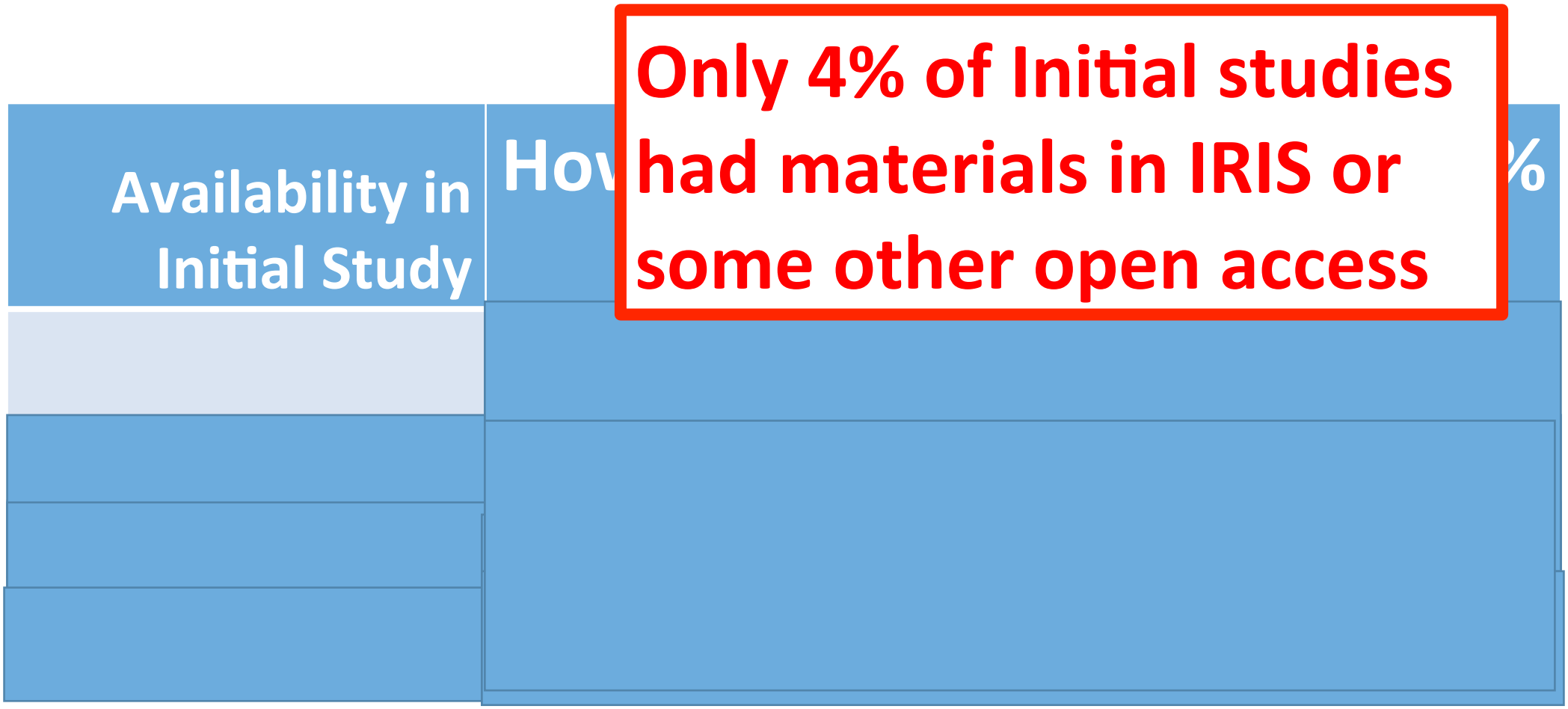
earning from the Reproducibility Project in Psychology:
nass coverage:

"in only 36% of the studies were the original results replicated" [because $p > 0.05$]

BUT… 77% of replication effect sizes were within a 95% prediction interval of original effect size

(Patil et al. 2016)

Transparency: How do replicators get hold of materials?



1) Replication: cultural changes in academia

Replication *not* an easy route; it's an *essential* route

Facilitate student apprenticeship model

"a lot is hidden behind the final [published article]."

Roxana, a replicator, SLA grad student/ trainee teacher from Vasquez & Harvey 2010 p.436

"Ralph is doing a preliminary study of re-inventing the wheel."

Effects of publication bias on replication effort

nly 4/67 tried to replicate a study that had null findings

ear that replication won't 'replicate findings of original'

v to make 'null findings' more publishable?

nsparent replication e.g. via pre-registration should help

thods fully ***reviewed*** and ***approved*** ***->***

rinciple Acceptance ***IPA***

N data collected

iewers cannot 'argue out' on basis of methodological flaw

pectives in Psychological Science, Cortex, Journal of Child Language

Part 3: Example of a multi-site replication

Morgan-Short, Marsden, Heil, et al. (manuscript)

https://osf.io/

The study: Is comprehension affected by attending to grammatical or lexical forms?

EITHER read a written text (and spot forms):

El pueblo azteca, como pueblo primitivo, podía encontrar u
problemas presentados por las fuerzas de la (La1) naturaleza.
importancia a su religión. En ella su Dios principal y todopo
sol) (Sol1) . Se lo admiró mucho. Tonatiuh tenía las bondades y los
humanos, pero con un gran poder sobrenatural. Según la (La2) r
sol (Sol2) Tonatiuh necesitaba que lo alimentaran (- n1) con una susta a...

Available on IRIS
www.iris-database.org

OR hear the same in the oral modality (and spot forms)

THEN comprehension measured by 10 multiple choice questions

from Leow et al. (20

Participants: Sites by Modality

704 participants across 7 research sites

- 4 sites ran **listening version**
  - University of Illinois at Chicago*
  - University of Oregon
  - Southampton University
  - Kazimierz Wielki University & Adam Mickiwicz University

- 3 sites ran **reading version**
  - University of York*
  - Georgetown University
  - University of South Carolina

Demonstration of the need for replication:

- Clearly replicated findings in 4 sites
- Not clear in 2 sites
- Not replicated in 1 site

Challenges for multi-site replication

1. Seeking collaborators
2. Parity in proficiency at different sites (Thomas, 1994)
3. Compatibility of software (E Prime / superlab scripts)
4. Responsibilities for data entry & analysis

we provided detailed protocols

Benefits of multi-site replication!

1. N=704

2. Protocols ready to be used again

   Four further replications ongoing:

   1 in China, 1 Native speakers, 1 Heritage learners, 1 after year abroad

3. Having different sites tempered our claims:

   One individual study could have *concluded* on basis of site with differenc

   Similar findings across ***6*** sites suggest something special in one 'odd' s

   = More reliable reason for generating new hypothesis

CONCLUSIONS:
Challenges for open methods and Recommendations

ome challenges for open materials

ncern 1: Open materials will encourage bad use of terials

still need critical thinking about…

urpose of materials

se or adaptation of materials

nalysis and *interpretation*

"I thought I was on to som
but I can't figure out hov
move it."

Osborne, R. (2013). Why open access makes no sense.
In N. Vincent & C. Wickham (Eds.) Debating Open Access. (pp 96-105). London: The British Academy.

oncern 2: Reluctance to share

thers might misuse my materials"

ut notes about use on IRIS

Ve gatekeep to reduce bad science: peer eview

ly next study might be 'scooped'"

he existing study should be open to full scrutiny

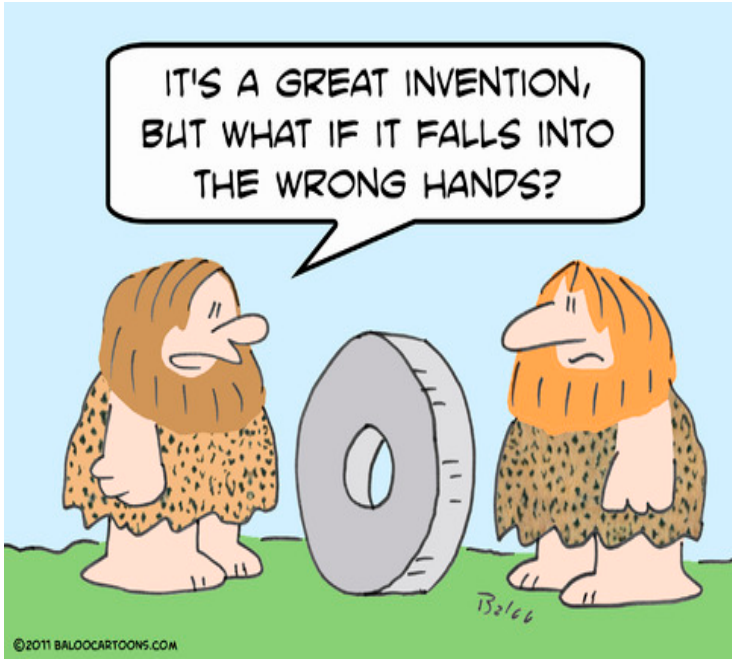
our plan won't be same as others'

might be proved wrong"

ood!

ut… unlikely to be so clear-cut

>citations!



"I don't have time"

- 15 minutes of your time *vs* 3 years of PhD student'
- Sharing *magnifies* impact of time & public money

"I can't find my materials or data"

- All the more reason for IRIS to exist!

Concluding remark (1)

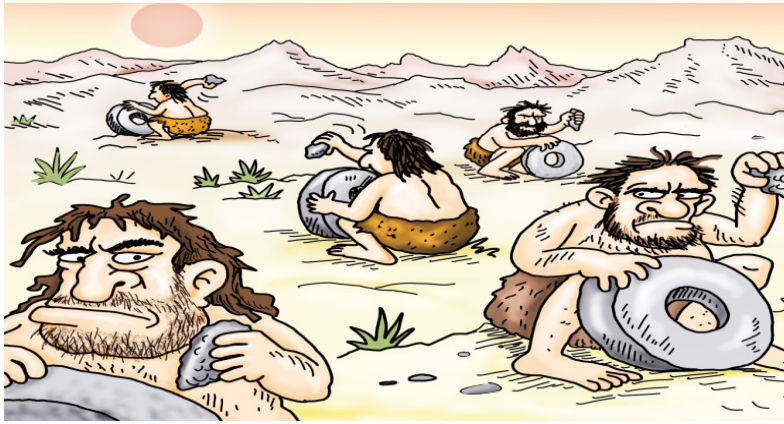
Rather than:
I'll share with *you (theoretical ally),*
but I won't share with *them*

researchers:

- Ethics: sharing with *all*

- Etic: one study individualism *vs synthetic, across contexts and over time*
(Plonsky 2012, Norris & Ortega 200

- a *collective methodological* memory



oncluding remark (2)

you, future reviewers and editors of journals and books:

Without seeing FULL materials, can ***reviewers*** *properly* evaluate?

Without **<u>open</u>** access to materials and data, can ***researchers*** have *properly*:

- built on previous methods *systematically*?
- compared their data to previous data?
- reduced potential bias? (given that independence helps)

Ask your journal editors…

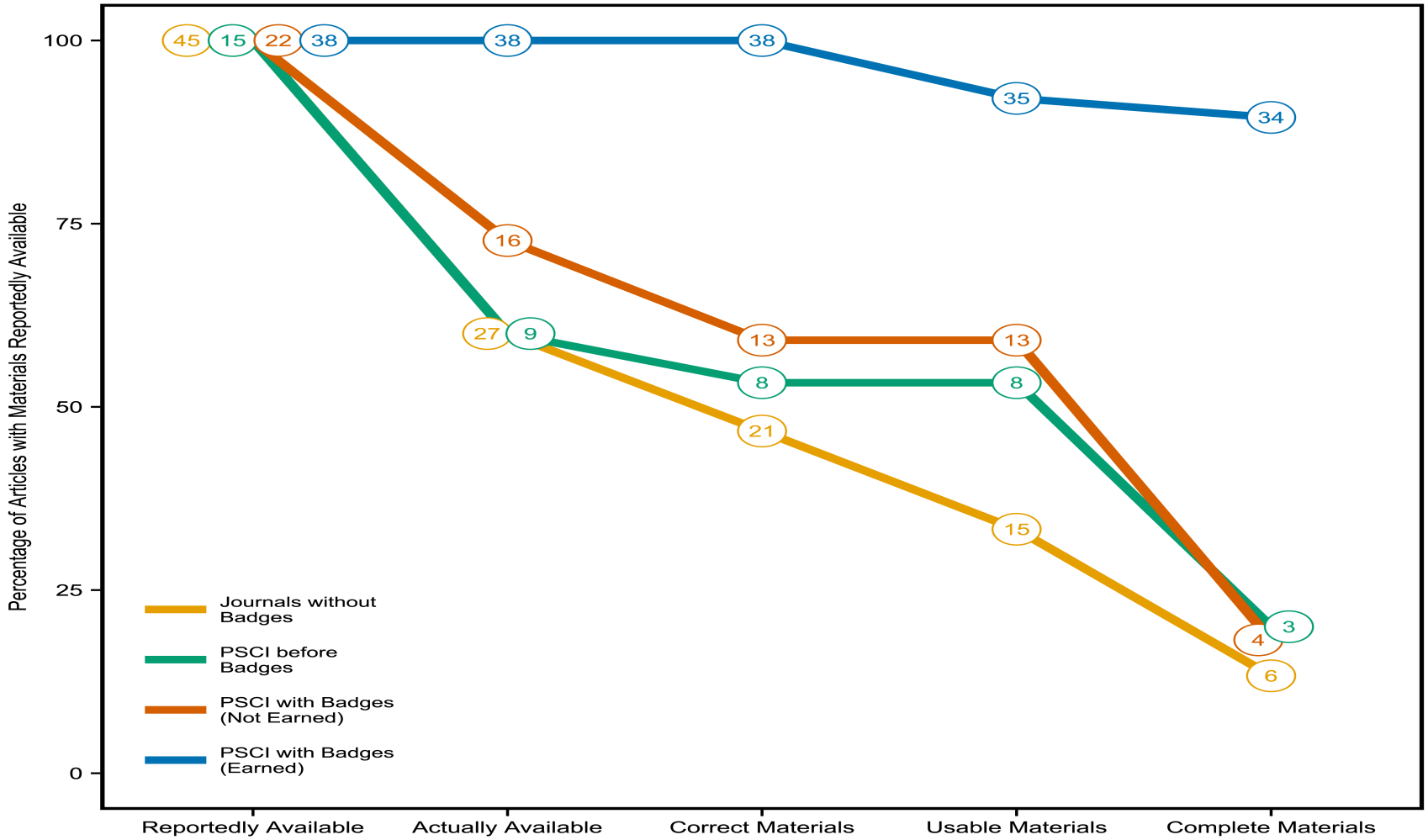
incentivise and recognise open science

Open Science Badges (from the Center for Open Science)

nguage Learning adopted the scheme in 2015.
udies in Second Language Acquisition, available from now
plied Linguistics, board approved
e Modern Language Journal, board approved
nguistic Approaches to Bilingualism, pending
nguage, Interaction and Acquisition, pending

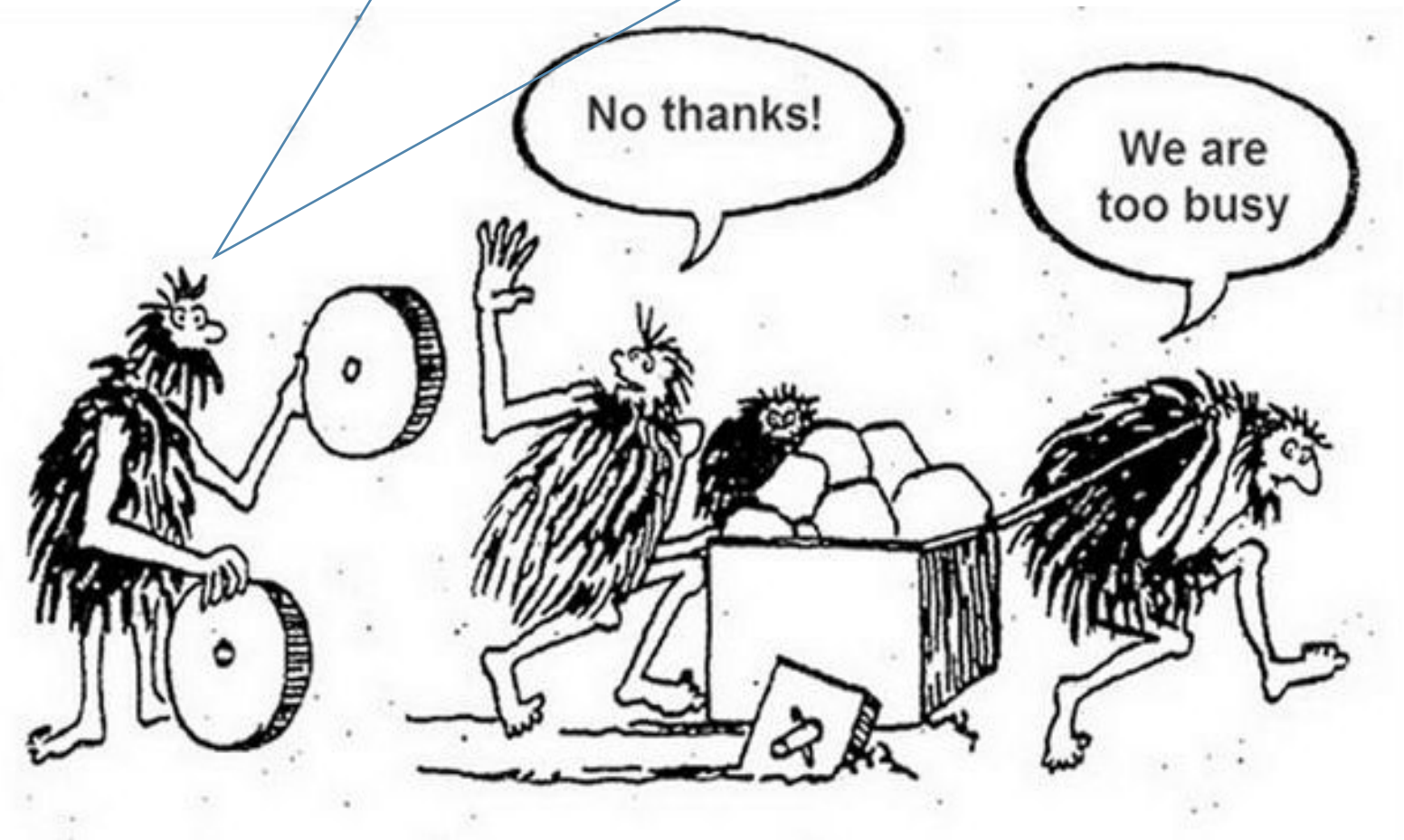
- Trofimovich & Ellis (2015) Editorial *Language Learning.* [Open Science Badges]
- Blohowiak, B., Cohoon, J., de-Wit, L.., Farach, F., Hasselman, F., & DeHaven, A. (2016). Badges to acknowledge open practices. osf.io/tvyxz

Effectiveness of journals recognising open materials with badges



dwell MC, Lazarević LB, Baranski E, Hardwicke TE, Piechowski S, Falkenberg L-S, et al. (2016) Badges to Acknowledge Open Practic
mple, Low-Cost, Effective Method for Increasing Transparency. PLoS Biol 14(5): e1002456. doi:10.1371/journal.pbio.1002456

Heh guys, have you seen these on
www.iris-database.org ?



Thank you for listening

With thanks to:

y collaborators

usan Gass, Patti Spinner, Dustin rowther (JT)

lison Mackey (IRIS)

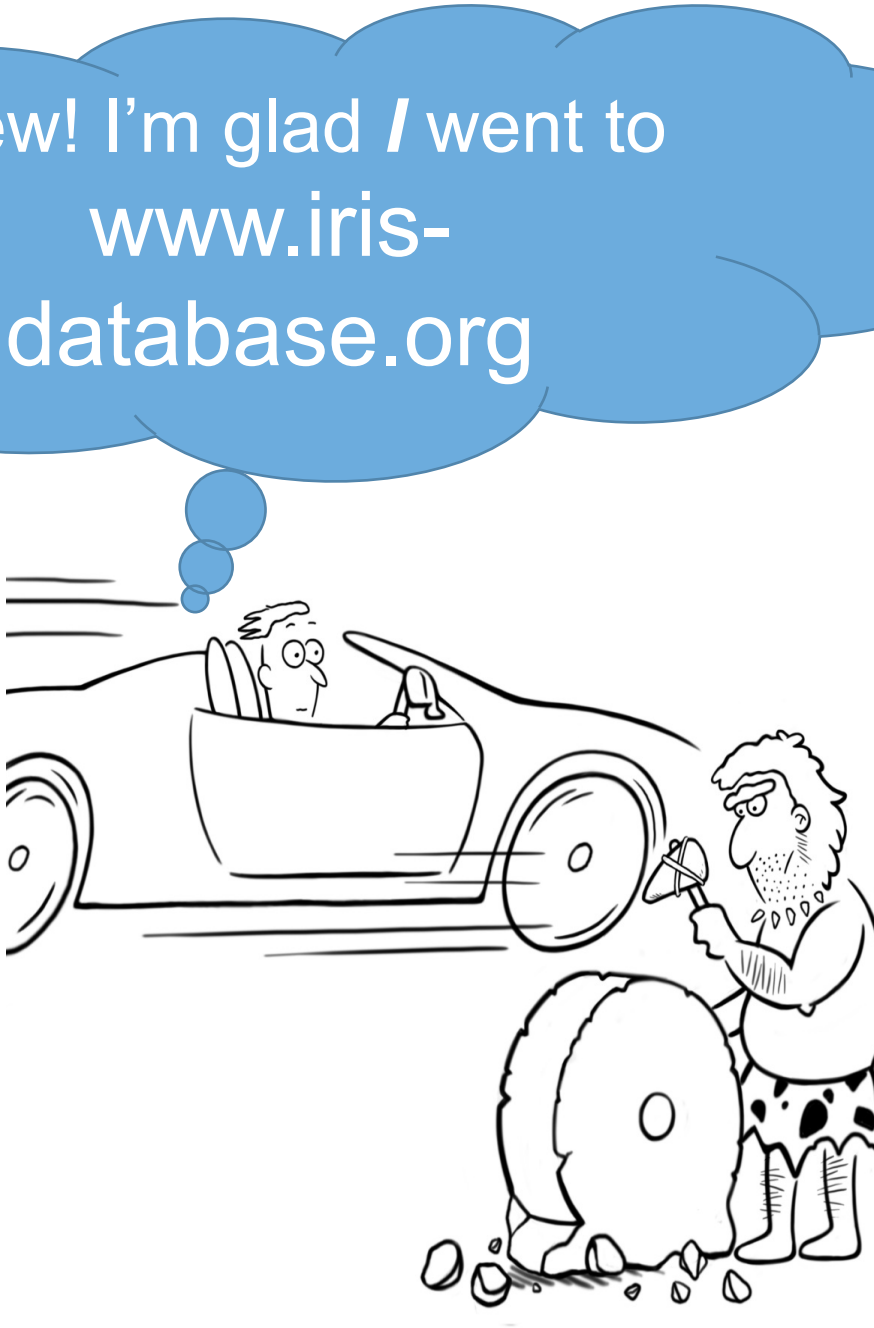
os Mitchell & Florence Myles earner language corpora projects)

ara Morgan-Short, Jeanne Heil, avid Abugaber (Replication)

uke Plonsky (IRIS, JT, SPR, Replication)

ophie Thompson (SPR, Replication)

nders



Selected references

Kidwell MC, Lazarević LB, Baranski E, Hardwicke TE, Piechowski S, Falkenberg L-S, et al. (2016) Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. PLoS Biol 14(5): e1002456. doi:10.1371/journal.pbio.1002456

Morgan-Short, K., Heil, J., Botero-Moriarty, A., & Ebert, S. (2012). Allocation of Attention to Second Language Form and Meaning. *Studies in Second Language Acquisition*, *34*(04), 659-685.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. Doi: 10.1126/science.aac4716

Plonsky, L. (2012). Replication, meta-analysis, and generalizability. In G. Porte (Ed.), *Replication research in applied linguistics* (pp. 116-132). New York: Cambridge University Press.

Polio, C., & Gass, S. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition, 19* (4), 499-508.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 12*, 287-301.

Wagner, A. B. (2010). Open access citation advantage: An annotated bibliography. *Issues in Science, Technology & Librarianship*. 60: Winter

Louise Corti - Veerle Van den Eynden - Libby Bishop - Matthew Woollard - 2014 Managing and Sharing Research DataA Guide to Good Practice, Sage

UK Data Archive MANAGING AND SHARING DATA a best practice guide for researchers 2nd ed
http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/researchdatamanagement/documents/managingsharing.pdf

Klein, R. et al. (2014). Investigating variation in replicability: A ‘‘Many Labs’’ replication project. *Social Psychology* 45 (3), 142–152.

Leow, R. P., Hsieh, H., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning, 58*(3), 665-695.

Marsden, E., Mackey, A., Plonsky, L. (2016) The IRIS repository: Advancing research practice and methodology. In Mackey, A. & Marsden E. (Eds.) *Advancing Methodology and Practice: The IRIS repository of Instruments for Research into Second Languages.* Taylor & Francis.

Makel, M. & Plucker, J (2014). Facts are more important than novelty: Replication in the Education sciences. *Educational Researcher,* 43 (6), 304 –316.

Makel, M., Plucker, J. & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives in Psychological Science,* 7 (6) 537 –542.

Wicherts, J., Bakker, M., Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One* , 6, 11, e26828