Why is so much ecological research done wrong?

Tomáš Grim

Palacký University, Olomouc, CZ

Because we make lots of *basic* mistakes

The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies

LEIF ENGQVIST

Conclusions beyond support: overconfident estimates in mixed models

Holger Schielzeth and Wolfgang Forstmeier

Citing practices in ecology: can we believe our own words?

Peter A. Todd, Darren C. J. Yeo, Daiqin Li and Richard J. Ladle

Why Most Published Research Findings Are False

John P. A. Ioannidis

Etc., etc., etc. ...

Because we make lots of *basic* mistakes

Simple means to improve the interpretability of regression coefficients

Holger Schielzeth*

Effect size, confidence interval and statistical significance: a practical guide for biologists

Shinichi Nakagawa[1,*] and Innes C. Cuthill[2]

Sample Size in the Study of Behaviour

Michael Taborsky[1]

How To Avoid Seven Deadly Sins in the Study of Behavior

MANFRED MILINSKI

Forstmeier, Nakagawa, Schielzeth – read whatever they wrote, please ☺

A model example: cuckoos and hosts

Teckentrup 2017: The egg

Cuckoos and hosts: is a larger host better?

Behav Ecol Sociobiol (1999) 47:41–46

ORIGINAL ARTICLE

Oddmund Kleven · Arne Moksnes · Eivin Røskaft
Marcel Honza

Host species affects the growth rate of cuckoo (*Cuculus canorus*) chicks

Abstract The cuckoo (*Cuculus canorus*) is an obligate interspecific brood parasite. When about to lay an egg, the female must decide which nest to parasitise. A high-quality host species should be preferred, to enhance the possibility of producing a viable offspring. In this study, we investigated the effects of two closely related host species, the great reed warbler (*Acrocephalus arundinaceus*) and the reed warbler (*A. scirpaceus*) on the growth rate of cuckoo nestlings. We found that cuckoo nestlings raised by the larger host species, the great reed warbler, grew significantly faster and became statistically significantly larger at fledging than nestlings raised by the smaller host, the reed warbler. Our results indicate a qualitative difference between the two host species. The great reed warbler, considered to be the best host, was parasitised at a higher rate than the reed warbler.

- 2 host species

Kleven et al Behav Ecol Sociobiol 1999

Previous comparative studies

Behav Ecol Sociobiol (2006) 60: 716–723
DOI 10.1007/s00265-006-0215-z

ORIGINAL ARTICLE

Tomáš Grim

Cuckoo growth performance in parasitized and unused hosts: not only host size matters

- 3 host species

Grim Behav Ecol Sociobiol 2006

Correlation between 2 data points!

Behav Ecol Sociobiol (1999) 47:41–46 © Springer-Verlag 1999

ORIGINAL ARTICLE

Oddmund Kleven · Arne Moksnes · Elvin Roskaft
Marcel Honza

Host species affects the growth rate of cuckoo (*Cuculus canorus*) chicks

Behav Ecol Sociobiol (2006) 60: 716–723
DOI 10.1007/s00265-006-0215-z

ORIGINAL ARTICLE

Tomáš Grim

Cuckoo growth performance in parasitized and unused hosts: not only host size matters

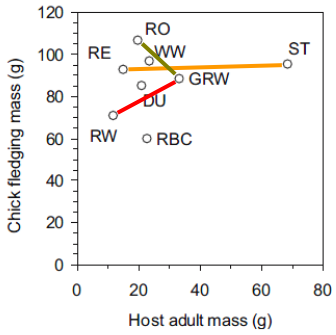
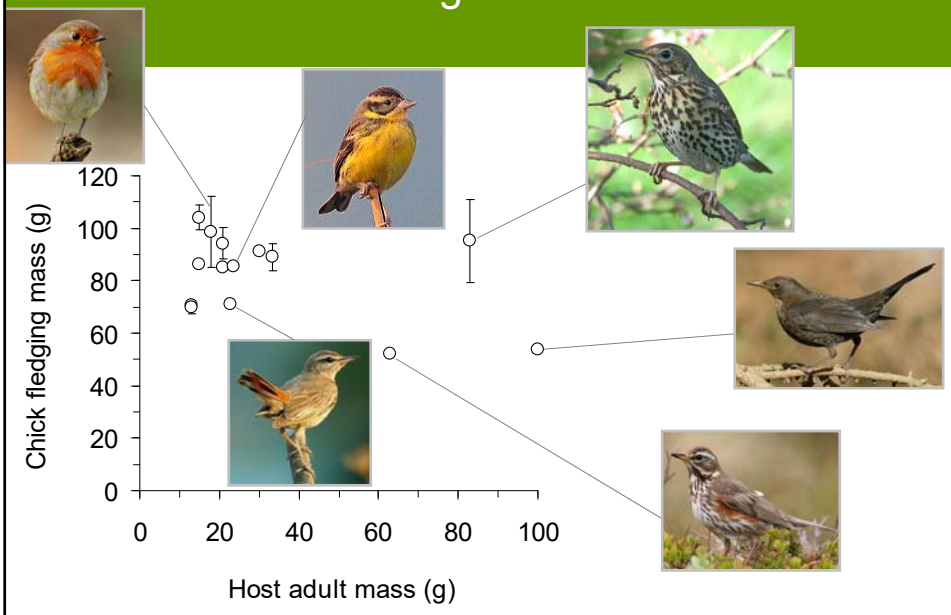


Fig. 2 Cuckoo chick fledging mass plotted against the host adult mass (sexes averaged). *DU* dunnock, *GRW* great reed warbler, *RBC* rufous bush chat, *RE* redstart, *RO* robin, *RW* reed warbler, *ST* song thrush, *WW* white wagtail

Grim Behav Ecol Sociobiol 2006

More data do not change the conclusions



Grim (in neverending prep.), see also Grim & Samaš Acta Ornithol 2016

What is muddled in Kleven et al. 1999? Corellation between *potential* predictors

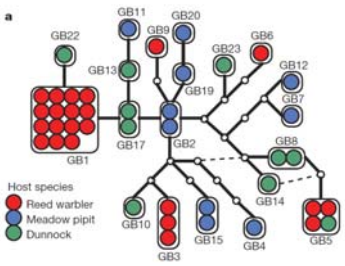
- body size
- only correlate of quality
- diet quantity
- diet quality
- nest defence
- mating system
- habitat ... everything ➔ impossible to infer causality
- "nature is the supermarket of evidence – buy what you like"

> Several studies on birds have shown great intraspecific variation in growth rate among nestlings (e.g. Magrath 1991; reviewed by Gebhardt-Heinrich and Richner 1998), mainly as a consequence of differences in the quantity or quality of food given to the young (Krebs and Avery 1984; reviewed by Martin 1987). The quantity and/or quality of food brought by the host species may

Kleven et al Behav Ecol Sociobiol 1999; Grim Behav Ecol Sociobiol 2006

What is muddled in Kleven et al. 1999? Corellation between *potential* predictors

- **host races** – maternal lines
- **cross-fostering**
- eliminates cuckoo genes X host quality correlation
- does not eliminate correlations among host traits!



Gibbs et al Nature 2000

Kleven et al Behav Ecol Sociobiol 1999; Grim Behav Ecol Sociobiol 2006

... thus it is useless to compare 2 species

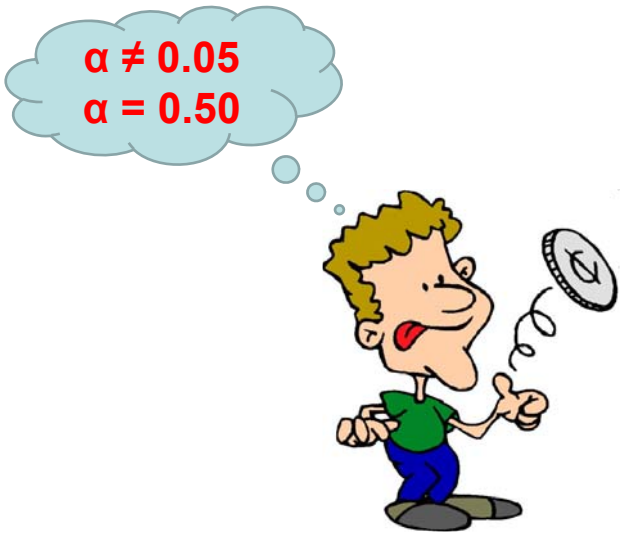
Why Not to Do Two-Species Comparative Studies: Limitations on Inferring Adaptation

Theodore Garland, Jr.[1]
Stephen C. Adolph[2]

ference with respect to any measure of the phenotype. But, for two species, the appropriate null hypothesis is more likely to be that a *difference* exists for any phenotypic trait.

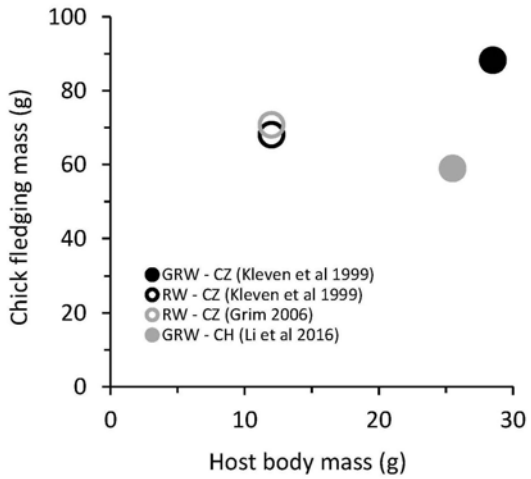
Garland & Adolph Physiol Zool 1994

... thus it is useless to compare 2 [...]



Garland & Adolph Physiol Zool 1994

... thus it is useless to compare 2 [...]



Li et al Evol Ecol 2016

What is muddled in Kleven et al. 1999? Levels: prediction *vs*. measurement

- predictor – host quality – ***species***
- response – chick mass – ***individual***

- dichotomization of continual predictor:
- low quality host: n = 1 (!)
- high quality host: n = 1 (!)

- **treatment replicates** (species) vs.
- **repeated measures** (individuals)

Kleven et al Behav Ecol Sociobiol 1999

The most common error in ecological studies

Ecological Monographs, 54(2), 1984, pp. 187–211

PSEUDOREPLICATION AND THE DESIGN OF ECOLOGICAL FIELD EXPERIMENTS[1]

STUART H. HURLBERT
Department of Biology, San Diego State University, San Diego, California 92182 USA

Abstract. Pseudoreplication is defined as the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent. In ANOVA terminology, it is the testing for treatment effects with an error term inappropriate to the hypothesis being considered. Scrutiny of 176 experimental studies published between 1960 and the present revealed that pseudoreplication occurred in 27% of them, or 48% of all such studies that applied inferential statistics. The incidence of pseudo-

> *No one would now dream of testing the response to a treatment by comparing two plots, one treated and the other untreated.*
>
> —R. A. Fisher and J. Wishart (1930)

Hurlbert Ecol Monogr 1984

The most common error in ecological studies

> *I don't know how anyone can advocate an unpopular cause unless one is either irritating or ineffective.*
>
> —Bertrand Russell (in Clark 1976:290)

Pseudoreplication is a Pseudoproblem

Jeffrey C. Schank, University of California
Thomas J. Koehnle, Hiram College

Logic of experiments in ecology: is pseudoreplication a pseudoissue?

Lauri Oksanen

On misinterpretations of pseudoreplication and related matters: a reply to Oksanen

Stuart H. Hurlbert, Stuart H. Hurlbert, Dept of Biology and Center for Inland Waters, San Diego State Univ., San Diego, California 92182, USA. (shurlbert@sunstroke.sdsu.edu)

ЖУРНАЛ ОБЩЕЙ БИОЛОГИИ

Zhurnal Obshchei Biologii (Journal of General Biology). Abstracts

Volume 64, 2003. N 4, pp. 292–307

Pseudoreplication in ecological research: the problem overlooked by Russian scientists

M. V. Kozlov

The most read (?) and misunderstood concept

Why Does the Frequency of Nest Parasitism by the Cuckoo Differ Considerably Between Two Populations of Warblers Living in the Same Habitat?

Andrzej Dyrcz & Konrad Halupka

Parent birds were not marked individually. To avoid pseudoreplication (i.e. testing the same individual more than once), we studied each species only within a single season and refrained from repetitive nest searching in the same reed-bed.

➢ **multicollinearity**
(time [year] x species)

> Authors who cite Hurlbert would do better if they had read his paper!
>
> – A.J. Underwood (1998:344)

Dyrcz & Halupka Ethology 2007 Graham Ecology 2003

The most important paper in biology

Pseudoreplication and the design of ecological field experiments
SH **Hurlbert** - Ecological monographs, 1984 - Wiley Online Library
Abstract Pseudoreplication is defined as the use of inferential statistics to test for treatment effects with data from experiments where either treatments are not replicated (though samples may be) or replicates are not statistically independent. In ANOVA terminology, it is
☆ ❞ Počet citací tohoto článku: 7779 Související články Všechny verze (počet: 69)

236x/yr

Molecular structure of nucleic acids
JD **Watson**, FHC Crick - Nature, 1953 - nature.com
A structure for nucleic acid has already been proposed by Pauling and Corey 1. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the
☆ ❞ Počet citací tohoto článku: 13978 Související články Všechny verze (počet: 19)

218x/yr

Google Scholar (Dec 17, 2017)

The most common error in ecological studies

FOLIA ZOOLOGICA - 45(1): 31-34 (1996)

EFFECT OF HABITAT ON THE DIET OF REED WARBLER (ACROCEPHALUS SCIRPACEUS) NESTLINGS

Tomáš GRIM and Marcel HONZA

Received August 21, 1995
Accepted January 16, 1996

Institute of Landscape Ecology, Academy of Sciences of the Czech Republic, Brno

Abstract

The diet of reed warblers (*Acrocephalus scirpaceus*) was studied in reed-beds in the Southern-eastern part of the Czech Republic. Food samples were collected when placing neck collars on nestlings in the breeding season of 1994. From 94 food samples were obtained 708 prey items. Diptera (66.5%), Homoptera (12.7%) and Aranea (7.2%) were the predominant food types observed, and the average body length of the prey was found to be 8.0 mm (1.9-21.0 mm). The total diversity of food was H' = 1.109, and J = 0.446. Food composition between two studied plots was significantly different.

The most common error in ecological studies

Arrival timing in subadult and adult Black Redstart males: competition-dependent behaviour?

Author(s): Schwarzova, L (Schwarzova, L.)[1]; Stros, P (Stros, P.)[1]; Frynta, D (Frynta, D.)[1]; Fuchs, R (Fuchs, R.)[1,2,3]

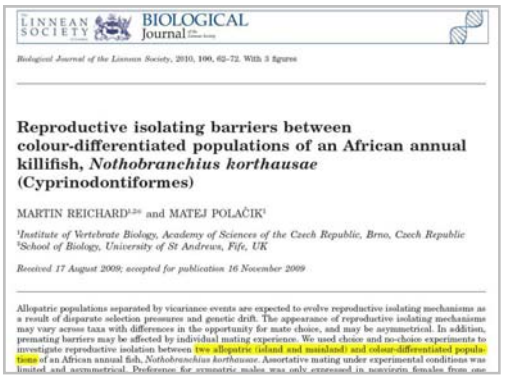
Source: ETHOLOGY ECOLOGY & EVOLUTION **Volume:** 22 **Issue:** 2 **Pages:** 111-118 **Article Number:** Pll 922021483 **DOI:** 10.1080/03949371003707588 **Published:** 2010

Times Cited: 1 (from Web of Science)

Cited References: 34 [view related records] Citation Map

Abstract: The different arrival times of 1-year-old and older males is a widely recognised phenomenon in most migrating passerines. The converse pattern, i.e. the yearlings arriving at the breeding grounds at the same time as adults, has been reported only exceptionally. Two hypotheses have been formulated to explain the delayed arrival of yearling males: investment reduction, and energetic constraint hypotheses, respectively. However, which of these is the more appropriate to explain the phenomenon remains unclear. We studied arrival timing in two distinct urban populations of Black Redstart (Phoenicurus ochruros), a small passerine exhibiting delayed plumage maturation, whereby dull-coloured yearlings are easily distinguishable from black adults. The birds inhabiting South Moravian villages did not show differential timing of arrivals between yearlings and older birds in the resource-rich environment. On the other hand, under suboptimal conditions in Prague the yearlings were delayed by three days on average compared to the adult males. We discuss the possibility that the young birds of this species sample the situation on a potential breeding site during the autumnal singing period and flexibly adapt their migrating behaviour according to the circumstances. In such a case the investment reduction hypothesis is the best explanation for different arrival timing of yearling and adult males in Black Redstart.

The most common error in ecological studies



LINNEAN SOCIETY BIOLOGICAL Journal

Biological Journal of the Linnean Society, 2010, **100**, 62–72. With 3 figures

Reproductive isolating barriers between colour-differentiated populations of an African annual killifish, *Nothobranchius korthausae* (Cyprinodontiformes)

MARTIN REICHARD[1,2*] and MATEJ POLAČIK[1]

[1]*Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, Brno, Czech Republic*
[2]*School of Biology, University of St Andrews, Fife, UK*

Received 17 August 2009; accepted for publication 16 November 2009

Allopatric populations separated by vicariance events are expected to evolve reproductive isolating mechanisms as a result of disparate selection pressures and genetic drift. The appearance of reproductive isolating mechanisms may vary across taxa with differences in the opportunity for mate choice, and may be asymmetrical. In addition, premating barriers may be affected by individual mating experience. We used choice and no-choice experiments to investigate reproductive isolation between two allopatric (island and mainland) and colour-differentiated populations of an African annual fish, *Nothobranchius korthausae*. Assortative mating under experimental conditions was limited and asymmetrical. Preference for sympatric males was only expressed in nonvirgin females from one

available at tropical and temperate locations (Martin and Fitzgerald 2005), while a comparison of only two populations does not provide any power to distinguish between alternative hypotheses. Moreover, a comparison of established populations from a single species says nothing about why some species are more likely

Blackburn et al 2009: Avian Invasions, p. 99

The most common error in ecological studies

Effects of Exotic Habitat on Nesting Success, Territory Density, and Settlement Patterns in the Blackcap (*Sylvia atricapilla*)

VLADIMÍR REMEŠ

*tat quality. I monitored the nesting success of Blackcaps (*Sylvia atricapilla*) in two types of forest in southern Moravia in the Czech Republic. I assessed their breeding density and territory size using a territory mapping method and the minimum convex polygon method. I determined spring arrival through direct observations and measured vegetation characteristics and pattern of spring leafing of shrubs in both forests. I show that Blackcaps preferentially settled in a plantation of introduced black locust (*Robinia pseudoacacia*) upon their return from spring migration. In this plantation, they reached twice the density as that observed in a natural floodplain forest nearby. However, they had significantly lower nesting success (15.5%) than in the floodplain forest (59%). Returning migrant Blackcaps may be lured by early-leafing shrubs in the exotic planta-*

- not two “types” but two “locations”!
- the language obscures: sweeps the problem under the carpet

The most common error in ecological studies

Duration of sympatry and coevolution between the great spotted cuckoo and its magpie host

M. Soler & A. P. Møller*

Departamento Biología Animal, Ecología y Genética, Facultad de Ciencias, Universidad de Granada, E-18001 Granada, Spain
* Department of Zoology, Uppsala University, Box 561, S-751 22 Uppsala, Sweden

NATURE · VOL 343 · 22 FEBRUARY 1990

inherently erroneous, because in the absence of replication, statistics can only tell us that there are spatial differences in nature – which we know anyway. Curi-

Oksanen Oikos 2001

What pseudoreplication is and what it is not?

- **repeated measurements of individuals ≠ pseudoreplication**
- paired study desing (→ paired t-test, etc.)
- repeatability, personality, growth
- random effects – mixed models

- **no repeated measurements of indiv. ≠ correct replication**

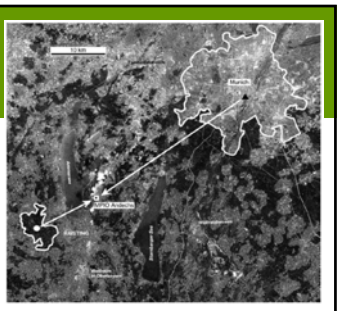
- pseudorepl. is a statistical/inferential error (not a design error)
- **structure: hypothesis → design → data → statistical model**

Nakagawa & Schielzeth Biol Rev 2010

Inappropriate generalizations

INCREASED SEDENTARINESS IN EUROPEAN BLACKBIRDS FOLLOWING URBANIZATION: A CONSEQUENCE OF LOCAL ADAPTATION?

JESKO PARTECKE[1] AND EBERHARD GWINNER[2]



Correct title:
Blackbirds from Munich differ from those living in a nearby forest

Assuring that the replicate samples or measurements are dispersed in **space** (or time) in a manner appropriate to the specific hypothesis being tested is the most critical aspect of the design of a mensurative experiment.

(Hurlbert 1984, p. 189)

Inappropriate generalizations

Antioxidants in the egg yolk of a wild passerine: Differences between breeding seasons

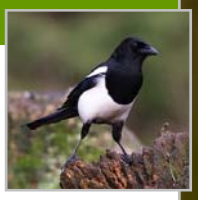
Rita Hargitai [a,*], Zoltán Matus [b], Gergely Hegyi [a], Gábor Michl [a], Gyula Tóth [b], János Török [a]

. Carotenoid concentration was higher in 2000 than in 2001. As

Assuring that the replicate samples or measurements are dispersed in space (or **time**) in a manner appropriate to the specific hypothesis being tested is the most critical aspect of the design of a mensurative experiment.

(Hurlbert 1984, p. 189)

Inappropriate generalizations



Behavioral Ecology (2015), 00(00), 1–8. doi:10.1093/beheco/arv206

Original Article

Birds use eggshell UV reflectance when recognizing non-mimetic parasitic eggs

Michal Šulc,[a,b] Petr Procházka,[a] Miroslav Capek,[a] and Marcel Honza[a]

Correct title:
Great reed warblers use eggshell UV reflectance when recognizing non-mimetic parasitic eggs

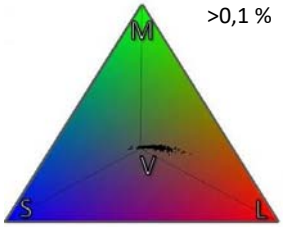
Assuring that the replicate samples or measurements are dispersed in space (or time **[or phylogeny]**) in a manner appropriate to the specific hypothesis being tested is the most critical aspect of the design of a mensurative experiment.
(Hurlbert 1984, p. 189)

Šulc et al Behav Ecol 2016

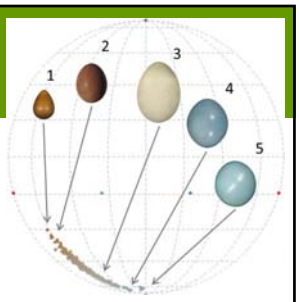
Inappropriate generalizations

Biology = variation
Variation constrains potential for generalizations

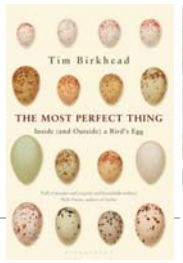
~~In~~appropriate generalizations

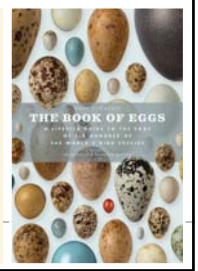


>0,1 %



- 5,604 eggs
- 636 species
- all avian orders





Hanley et al Biol Lett 2015

~~In~~appropriate generalizations

The Many Faces of Replication

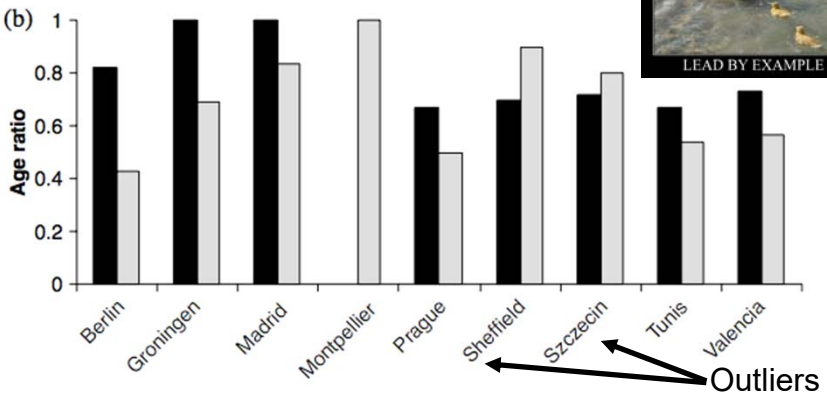
Douglas H. Johnson*

ABSTRACT

Replication is one of the three cornerstones of inference from experimental studies, the other two being control and randomization. In fact, replication is essential for the benefits of randomization to apply. In addition to ordinary replication, the repetition of treatments within a study, two other levels of replication have been identified. Pseudoreplication, a termed coined by Stuart Hurlbert, generally involves making multiple measurements on experiment units (which is commendable) and treating them as if they reflected independent responses to treatment (which is erroneous). Metareplication is a higher level of replication in which entire studies are repeated. Scientists are too much concerned about analysis of data within studies and too little concerned about the repeatability of findings from studies conducted under a variety of conditions. Findings that are consistent among studies performed at different locations at different times with different investigators using different methods are likely to be robust and reliable.

Johnson 2006 USGS Paper 34

Without metareplication any generalization is just a wishful thinking



- Scientific "truth" is polled: 6:2

Evans et al Oikos 2009 (+ many other great studies by his team)

Without metareplication any generalization is just a wishful thinking

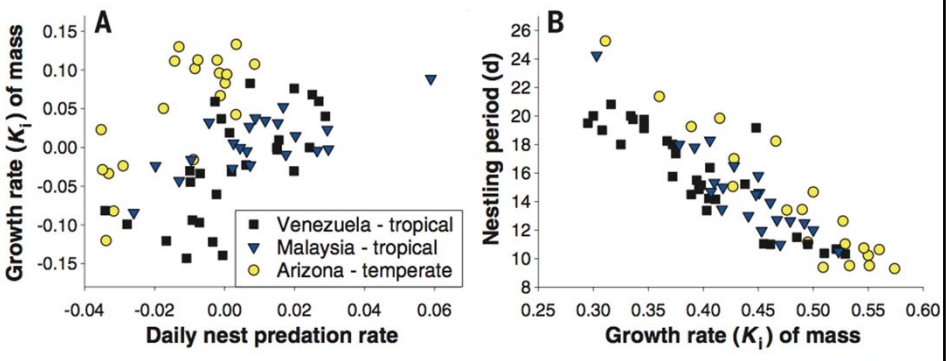
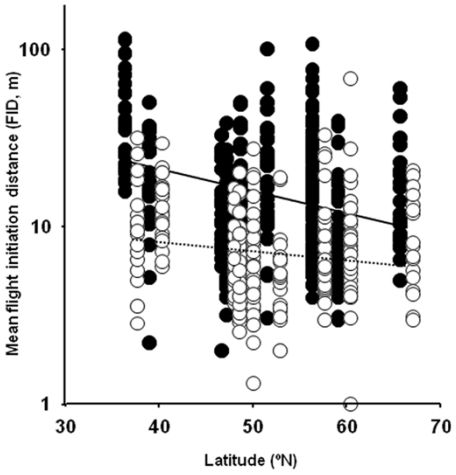


Fig. 2. Growth and nest predation on three continents. (**A**) Peak growth rate is faster in species with

Martin Science 2015 (+ many other great studies by his team)

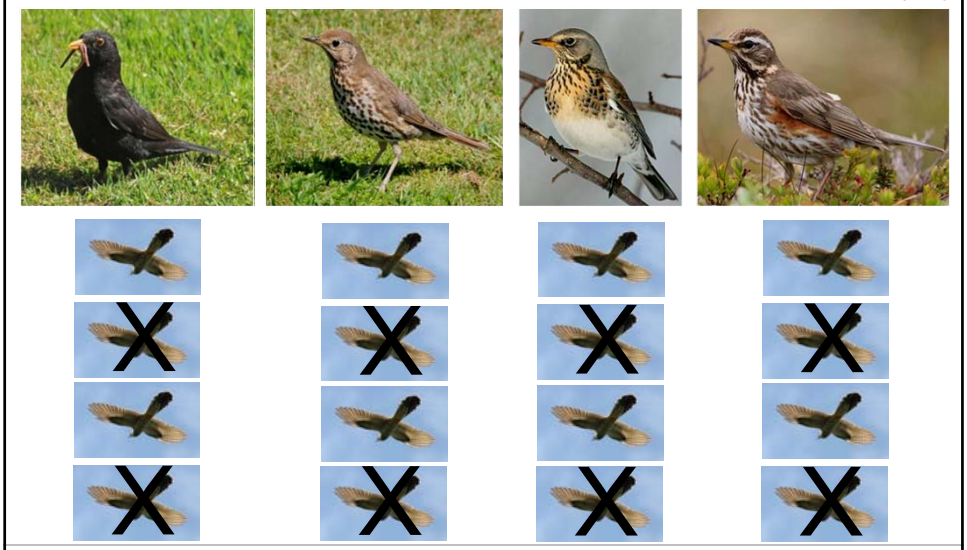
Without metareplication any generalization is just a wishful thinking



- urban – rural replicates
- paired design
- allows for additional questions

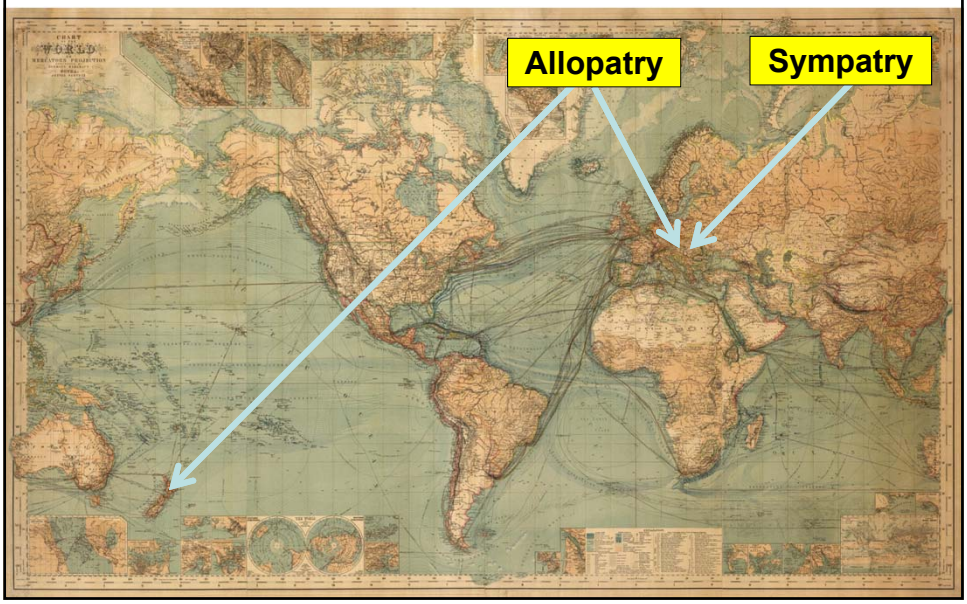
Díaz et al PLoS ONE 2013, Behav Ecol 2015, Møller et al Oecologia 2012, 2015

An example: sympatry *vs*. allopatry
Our study: "The black**bird** is not **birds**"



Grim et al J Anim Ecol 2011

A fundamental problem of allop./symp. studies: The length of allopatry unknown



Avian invasions = natural experiments

In the Light of Introduction: Importance of Introduced Populations for the Study of Brood Parasite–Host Coevolution

Tomáš Grim and Bård G. Stokke

Cambridge Univ Press 2016

Metareplication *vs.* pseudoreplication

Sympatry			Micro-allopatry		Macro-allopatry			Sympatry			Micro-allopatry		Macro-allopatry		
Ch.	Gr.	Lu.	Ol.	Br.	Au.	Ha.	Ta.	Ch.	Gr.	Lu.	Ol.	Br.	Au.	Ha.	Ta.
y1	y1	y1	**y1**	y1	y1	y1	y1	y1	y1	y1	y1	y1	y1	y1	y1
y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2	y2
y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3	y3
…	…	…	…	…				…	…	…	…	…			

A typical brood parasitism study (= pseudoreplication)

Samas et al Front Zool 2014, Ethology 2011; Grim et al Behav Ecol Sociobiol 2014

Metareplication *vs.* pseudoreplication

- comparisons **between** 2 species = **pseudo**replication
- comparisons **within** 2 species = **meta**replication

Table 2 Egg rejection response and latency to rejection by thrushes

Type of parasitism	Blackbird			Song thrush		
	ddf	F	P	ddf	F	P
CONSPECIFIC						
(R)ejection						
Breeding density	105	↔	.02	41	↔	0.09
Clutch	101		0.24	38		0.38

Samas et al Front Zool 2014, Ethology 2011; Grim et al Behav Ecol Sociobiol 2014

Metareplication makes a change: basic “dogma” of BP studies challenged

Journal of Animal Ecology

Journal of Animal Ecology 2011, **80**, 508–518 doi: 10.1111/j.1365-2656.2010.01798.x

Constraints on host choice: why do parasitic birds rarely exploit some common potential hosts?

Tomáš Grim[1*], Peter Samaš[1], Csaba Moskát[2], Oddmund Kleven[3], Marcel Honza[4], Arne Moksnes[5,6], Eivin Røskaft[5,6] and Bård G. Stokke[5,6]

Samas *et al. Frontiers in Zoology* 2014, **11**:34
http://www.frontiersinzoology.com/content/11/1/34

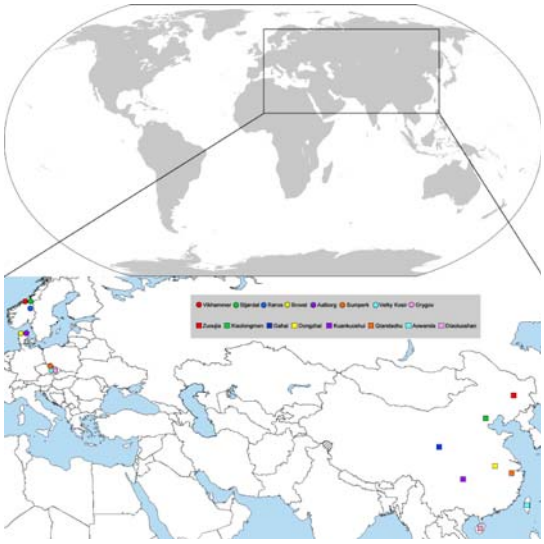
FRONTIERS IN ZOOLOGY

RESEARCH Open Access

Host responses to interspecific brood parasitism: a by-product of adaptations to conspecific parasitism?

Peter Samas[1], Mark E Hauber[2], Phillip Cassey[3] and Tomas Grim[1*]

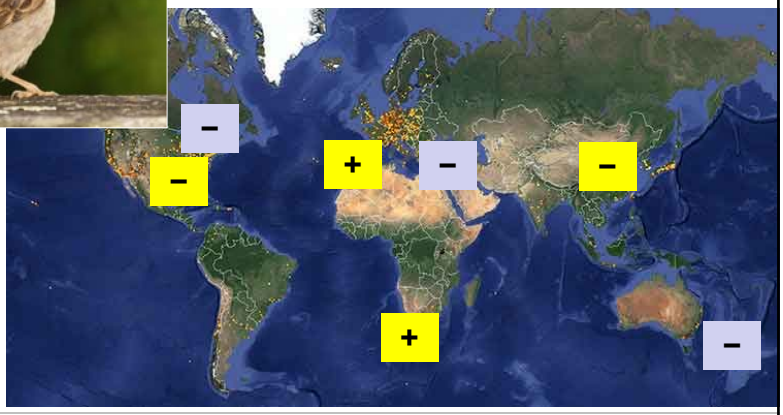
Metareplication makes a change: textbook examples of unsuitable hosts rejected



Liang et al Behav Ecol 2016

Metareplication makes a change: a potential global model for CBP studies rejected

- **Scientific "truth" is polled: 5:2**



Manna et al J Avian Biol 2016

Metareplication makes a change: common knowledge?

Vol. 66: 185–190, 2015 | CLIMATE RESEARCH Clim Res | Published December 3

Effects of urbanization on bird phenology: a continental study of paired urban and rural populations

Anders Pape Møller[1,2], Mario Díaz[3], Tomáš Grim[4], Alena Dvorská[4], Einar Flensted-Jensen[5], Juan Diego Ibáñez-Álamo[6,7], Jukka Jokimäki[8], Raivo Mänd[9], Gábor Markó[10,11,12], Paweł Szymański[13,*], Piotr Tryjanowski[14]

	All previous studies (n = 17)	This single study (n = 137)
N spatial (= no. urban-rural study site pairs)	1	8
N biological (= no. species, not data points!)	1	54
N statistical (= urban-rural spp. pairs, not breeding pairs!)	1 (= pseudoreplication)	137 (= metareplication)

Møller et al Clim Res 2015

Lessons
(we know them ... but humans forget)

- "think before u collect data"
- hypothesis → design → data → statistical model
- fundamental errors vs. polishing-nitpicking
- = get **the <u>basics</u>** right

Lessons
(we know them ... but humans forget)

- don't pseudoreplicate – metareplicate
- avoid unsubstantiated generalizations
- the secretary**<u>bird</u>** is not **<u>birds</u>**
- the great tit is not birds too!

Lessons
(we know them ... but humans forget)

Metareplication across
- **phylogeny**
- **space**
- **time**

So why is so much ecological research done wrong?