

The Power and the Pitfalls of Large-scale, Unreplicated Natural Experiments

Shanta C. Barley* and Jessica J. Meeuwig

School of Animal Biology and the Oceans Institute, University of Western Australia, 35 Stirling Highway, Crawley, Perth, Western Australia 6009, Australia

ABSTRACT

Large-scale, unreplicated natural experiments (LUNEs) have a unique power to test hypotheses at ecologically realistic scales and have delivered insights of great power into cosmology, evolution and geology. Yet, LUNEs are relatively rare in the field of ecology and continue to meet resistance due to their lack of replication. However, in the vast majority of cases, large-scale experiments cannot be replicated for practical and ethical reasons. Here, we make the case that LUNEs have had a disproportionately positive effect on conservation policy and are a crucial next step in the extrapolation of

our understanding of ecological processes from small-scale experiments to relevant scales, particularly in the context of the current “replication crisis” affecting many sciences. Greater inclusion of LUNEs in mainstream ecology will help humanity to solve global problems as human transformation of the planet accelerates in coming decades.

Key words: ecosystem scale; ecological processes; large-scale experiment; natural experiment; pseudo-experiment; replication.

We must rely more on long-term, ecosystem-scale experiments and real case histories, and less on small-scale experimentation.—David Schindler (2012)

THE MODERN CHALLENGE

Since the beginning of the Anthropocene, *Homo sapiens* has transformed the planet’s ecosystems via exploitation, deforestation, pollution, eutrophication and climate change (Estes and others 2011;

Ehrlich and Ehrlich 2013). Not only have we eradicated apex predators in many parts of the world (Myers and others 2007; Atwood and others 2015), but we are a uniquely voracious “hyper-keystone” species and super-predator, consuming adult prey at a rate 14 times that of any other predator (Darimont and others 2015; Worm and Paine 2016). Humanity failed to anticipate many of the ecological problems that it has caused, including the depletion of the ozone layer and bio-accumulation of DDT (Kates and Clark 1996). It has been argued that this lack of foresight was partly due to an overreliance on small-scale, short-term experiments that failed to capture the complex nature of large-scale perturbations and net ecosystem responses (Carpenter 1989, 1990; Pace and others 1998; Schindler 1998; Table 1). Although small-scale, controlled studies provide a useful means to test hypotheses, there is an urgent

Received 10 February 2016; accepted 25 June 2016;
published online 29 August 2016

Author contributions JM/SB conceived of the review, JM/SB contributed to structuring the manuscript, SB wrote the paper and JM/SB edited the paper.

*Corresponding author; e-mail: shanta.barley@gmail.com

Table 1. Contrasting Outcomes from Small- and Large-scale Experiments

Subject	Small-scale results	Large-scale results	Confounding factors	References
Carbon capture and storage	"Pilot" projects do not cause seismicity	Large-scale wastewater injections can cause earthquakes	Changes in pressure at different spatial scales	Zoback and Gorelick (2012)
Invasive cane toads <i>Bufo marinus</i>	>30% of all snake species in Australia vulnerable to toad invasion	Most species unaffected; some taxa increased in abundance	Toads targeted lizards; stochastic weather events	Brown and others (2011)
Effect of elevated CO ₂ on temperate forests	25% reduction in water loss	<3% increase in catchment run-off	Soil/atmosphere feedback, plant-animal interactions	Leuzinger and others (2011)
Species diversity	Higher diversity = higher rates of ecological processes	Species composition, not diversity per se, determines ecosystem properties	Species interactions	Wardle and Barker (1997)
Soil respiration	Effect of plant roots insignificant compared to microbial heterotrophs	Plant roots contribute >54% of soil respiration	Edge effects	Högberg and others (2001)
Influence of climate change on physiological variability	NA	Strong effect on mammalian metabolic rates/insect thermal tolerances	Undetectable at small scale	Chown and others (2004)
Effects of livestock grazing on grassland diversity	Grazing maintains/enhances plant diversity on mesic grasslands	Grazing may reduce biodiversity	Biological effects	Lezama and others (2014)
Influence of windstorms on forests dynamics	Highly unpredictable - wind gusts are more random at smaller scales	Effect correlated with wind speed, topography and land-use history	Topography	Weimin and Peet (2011)
Atlantic salmon <i>Salmo salar</i> dynamics	No effect of temperature on physiology; no effect of prey density or current speed on foraging rate	Temperature is a strong determinant of growth rate; prey density/current speed influence foraging rate.	Large-scale variation in currents, temperature regimes	Folt and others (1998)
Abundance of Atlantic cod and its prey, capelin, in the North Atlantic	Prey-predator abundance negatively correlated or uncorrelated	Abundances positively correlated; both species concentrated in large-scale oceanic fronts	Small-scale predator avoidance by capelin and small-scale prey depletion by cod	Rose and Leggett (1990)
Relationship between abundance of marine organisms and seagrass	Significant effect of seagrass height/density on abundance of fish and decapod species	Of the 23 significant responses on small scales, 15 were no longer significant at large scales	Stochastic settlement of larvae among seagrass beds	Bell and Westoby (1986)

Table 1. Contrasting Outcomes from Small- and Large-scale Experiments

Subject	Small-scale results	Large-scale results	Confounding factors	References
Effect of larval supply on reef fish abundance	Post-settlement mortality density dependent	Post-settlement mortality density independent	Predation risk and distribution/abundance of spatial refuges	Hixon and others (2012)
Impact of fish herbivores on macroalgae	Where macroalgal herbivores are present, they cull macroalgae	Intact and structurally intact herbivore communities may fail to consume algae	Behavioural variation in herbivores possibly induced by habitat differences	Bennett and Bellwood (2011)
Effects of <i>Daphnia</i> on microzooplankton under eutrophication	NA	Large-enclosure results agreed more closely with whole-lake dynamics than with small-enclosure results	Competition	Sarnelle (1997)

need to replicate findings at the level of ecosystems to better anticipate ecological “surprises” (Knowlton 1992; Clark and Gelfand 2006; Miao and Carstenn 2006; Chave 2013). However, a significant challenge remains: how does one robustly test hypotheses at large scales?

THE MODERN SOLUTION?

The solution to this problem may lie in ecologists making greater use of large-scale but unreplicated “natural experiments” (LUNES). LUNES have been generated by natural phenomena such as El Niño, hurricanes and tsunamis (Kessler 2011; Zoback and Gorelick 2012). However, these “accidental” experiments (*sensu* Hillerislambers and others 2013) are also a by-product of scientific and economic activities, including fisheries, large-scale infrastructure projects such as the Suez canal, carbon capture and storage projects and catastrophic oil spills such as Deepwater Horizon and the fur trade (Estes and Palmisano 1974; Smith and others 2004; Madin and others 2010). LUNES, regardless of their origin, can provide unique insights into ecosystem-level processes that complement the findings of classical laboratory and field experiments (Hillerislambers and others 2013).

However, LUNES constitute a Faustian bargain: in exchange for rare insights into net ecosystem responses, the ecologist may have to sacrifice treatment replication (Hargrove and Pickering 1992; Johnson 2006). Replication has traditionally been considered a prerequisite for extrapolating results to other systems and, where manipulations are included, for inferring cause (Fisher 1926). Yet, the replication of LUNES is typically challenging for ethical and practical reasons (how does one mimic a tsunami or build a life-size replica of the Panama Canal?). In the deepest sense, ‘fully controlled’ replication of any natural experiment is impossible, because the initial conditions have not been fully described. Furthermore, the “treatments” that create LUNES are often unique to a single location (Johnson 2002), and even if true replicates exist, the number required to preclude a Type II error may be prohibitively high due to the variability inherent in ecosystems (Carpenter 1989).

LUNES may nonetheless contribute a modern solution to the current “replication crisis” affecting science (Stroebe and Strack 2014; Nosek and others 2015). Relatively few attempts are typically made to replicate research findings and a disconcerting proportion of studies have turned out to be irreproducible (McNutt 2014). In response, some critics have sought to redefine the role of replication in

science, arguing that exact replications should no longer be treated as all-powerful, “one-off” confirmations or rejections of a theory (Stroebe and Strack 2014). Instead, scientists should prioritise “conceptual” replications that test the real-world relevance of theories in diverse field contexts (Hüffmeiera and others 2016). LUNEs offer a way to conduct such conceptual replications, as exemplified recently by Atwood and others (2015).

THE POWERS OF LUNES

LUNEs have led to new insights into a range of ecological issues, including the role of apex predators and prey within ecosystems, the effectiveness of ocean “geo-engineering” and the impact of climate change on marine communities, in addition to habitat fragmentation and eutrophication (Carpenter 1990; Vanni and others 1990; Rowan and others 1997; Kessler 2011; Gabric and others 2015). Yet, LUNEs remain an under-utilised minority in the scientific literature relative to classical experiments, that is, small-scale, controlled and replicated studies (Hillerislanders and others 2013). Ecologists and journal referees largely continue to view the latter as the scientific “gold standard” (Grossman and Mackenzie 2005), even though some of the most influential, manipulative field experiments in the history of ecology lacked replication (Schindler and others 1971, 2008; Raffaelli and Moller 1999). As a result, few LUNEs are conducted in the first place and fewer still are accepted for publication. However, in the context of emerging environmental challenges, LUNEs are uniquely positioned to answer challenging, time-sensitive ecological questions (Worm and Paine 2016).

Power 1: Testing Logistically and/or Ethically Challenging Hypotheses

Under rare circumstances, scientists have conducted controlled and randomised experiments at large spatial scales (Naeem and others 1994; Ewers and others 2011). However, in most cases, conducting classical experiments at large spatial scales is prohibitively expensive, time consuming, labour intensive or unethical. As a result, LUNEs provide one of the very few means available to scientists to test hypotheses at ecologically relevant scales (Carpenter 1990; Hargrove and Pickering 1992). For example, the construction of the Panama Canal created a unique opportunity to investigate competitive exclusion in previously isolated communities of freshwater fishes (Smith and others 2004). The Deepwater Horizon oil spill provided Kessler

(2011) with a unique and otherwise unobtainable simulation of the effects of climate change on subsea deposits of methane hydrates. Similarly, the invasions of Guam by brown tree snakes *Boiga irregularis* and Christmas Island by yellow crazy ants *Anoplolepis gracilipes* yielded insights into the ecological roles of keystone prey species that were unreplicable due to the ethical challenges around reducing the abundance of protected, endemic species (O’Dowd and others 2003; Rogers and others 2012).

Power 2: Systems That have no Replicates

In many cases, replicating a LUNE is not only ethically unconscionable but also impossible because the system is unique (Power and others 1998; Schindler 1998). For example, it is impossible to replicate natural experiments that explore the effect of invasive species on fauna endemic to specific islands or locations (O’Dowd and others 2003; Rogers and others 2012). Similarly, true replicates for LUNEs that examine lake and river processes remain elusive due to differences between water bodies in species composition, water chemistry and other factors (but see Schindler and others (1978) for statistical/partial mitigation measures). To circumvent this problem, studies may compare upstream, “control” stretches of river to downstream, “treatment” stretches that have been experimentally manipulated (Fraser and Gilliam 1987; Hildrew and others 2004). The task of finding true replicates for natural experiments conducted on coral reefs can also be daunting. Ruppert and others (2013) compared two reef systems in northwestern Australia that were similar with respect to disturbance history, productivity, habitat structure and a number of other characteristics, the primary difference being that one of the systems had a history of targeted shark exploitation. This study generated valuable evidence consistent with the hypothesis that depletion of sharks may lead to trophic cascades on coral reefs. However, it is difficult to imagine an exact “replicate” pair of reef systems due to the unique disturbance histories of coral reefs, the scarcity of pristine sites and the nature of fishing, which typically leads to the removal of not just sharks but also their prey (Pauly and others 1998).

Power 3: Meta-Analyses

Given the challenges of replicating LUNEs, the gradual accumulation of independent, unreplicated studies may be the only way to gather evidence at the ecosystem level (Carpenter and others 1995; Johnson 2002, 2006). Like lawyers, ecologists can

then “build a case” in favour of or against hypotheses (McArdle 1996). Ultimately, such studies can fuel meta-analyses in which drivers in common are reinforced when seen in combination (Cottenie and Meester 2003; Worm and Paine 2016). Although meta-analyses based on unreplicated studies have been criticised (Hurlbert 2004), they have provided insights into a range of important ecological issues including trophic cascades and overfishing (Micheli 1999; Shurin and others 2002; Prevedello and others 2013).

Meta-analyses of LUNEs have played a particularly important role in demonstrating the effectiveness of marine reserves (Edgar and others 2014). Many reserves are established precisely because a site is unique or contains endemic fauna (Allison and others 1998), precluding replication. Before-after-control-impact (BACI) analyses, in which samples are collected from “control” and “treatment” sites before and after a reserve is established, have played a key role in circumventing this problem (Bence and others 1996). However, relatively few reserves are subject to long-term monitoring programmes. Moreover, BACIs may constitute temporal autocorrelation (Stewart-Oaten and others 1986). Meta-analyses of unreplicated studies can provide a solution, but considerable time will be required to amass a sufficient number of studies to “fuel” the analysis, particularly if there is a publishing bias towards replicated studies (Lester and Halpern 2008).

Power 4: Ecological Gradients

LUNEs provide a particularly powerful tool when they operate across an ecological gradient. This approach allows a regression model to be fitted to the data, predictions to be made about how a variable changes in response to another variable and may identify “tipping points” within ecosystem function (Lennon 2011). Indeed, Kreyling and others (2014) recommend that such regression analyses be used to develop simulation models, which, in turn, can be validated by further experiments and ultimately generate data that are more informative about ecosystem processes than those produced by small-scale, manipulative experiments. Regression analyses based on LUNEs have proven to be particularly useful in exploring the role of large piscivores on coral reef systems. For example, a LUNE created by a gradient in fishing pressure along the Northern Line islands in Hawaii has been used to study the relationships between predator biomass and prey biomass, behaviour and condition, providing important validations of

hypotheses generated at smaller scales (Madin and others 2010; Walsh and others 2012).

Power 5: Qualitative Arguments

Some have argued that most robust LUNEs use evidence-based arguments rather than inferential statistics to describe ecological patterns. For example, even though Likens and others’ (1970) large-scale deforestation experiment at Hubbard Brook was unreplicated, it met with approval from critics because it “convincingly demonstrated the effects of the experimental variables, without resorting to inferential statistical tests that would have been inappropriate” (Hurlbert 1984). Similarly, of the more than nine large-scale, unreplicated iron and phosphorus enrichment ocean-based experiments that have been conducted globally, only two used inferential statistics, drawing criticism from Hale and Rivkin (2007). Moreover, LUNEs provide a powerful test of hypotheses when the effect is both consistent with *a priori* hypotheses and the magnitude of the effect is ecologically meaningful (Stewart-Oaten and others 1992; Moss and others 1996). For instance, when there is a single perturbation in an otherwise unmanipulated, well-described lake, it is possible to infer causality (Schindler, D.W., pers. comm.).

THE PITFALLS OF LUNES

Although LUNEs offer unique insights into ecological processes, they are fallible and must be replicated in a range of systems before becoming the basis for changes in policy. However, it is increasingly common for the results of unreplicated studies to be treated as definitive (an “ill-informed” strategy; Ioannidis 2005). In contrast, we would argue that ecologists who conduct LUNEs should present their results as either “consistent” or “inconsistent” with the tested hypothesis, while thoroughly exploring and ranking alternative explanations. In addition, it is important that the other scientists and the public be aware of the limitations of LUNEs and view the results in the context of other studies.

A cautionary example of the pitfalls associated with the interpretation of LUNEs is provided by the recent critique of Myers and others (2007), a study that linked the collapse of a scallop fishery in the Atlantic Ocean to declines in large sharks due to overfishing. This LUNE not only led to changes in fishing policy but also a vigorous campaign to reduce numbers of the cownose ray *Rhinoptera bonasus*, a slow-growing species with poor resi-

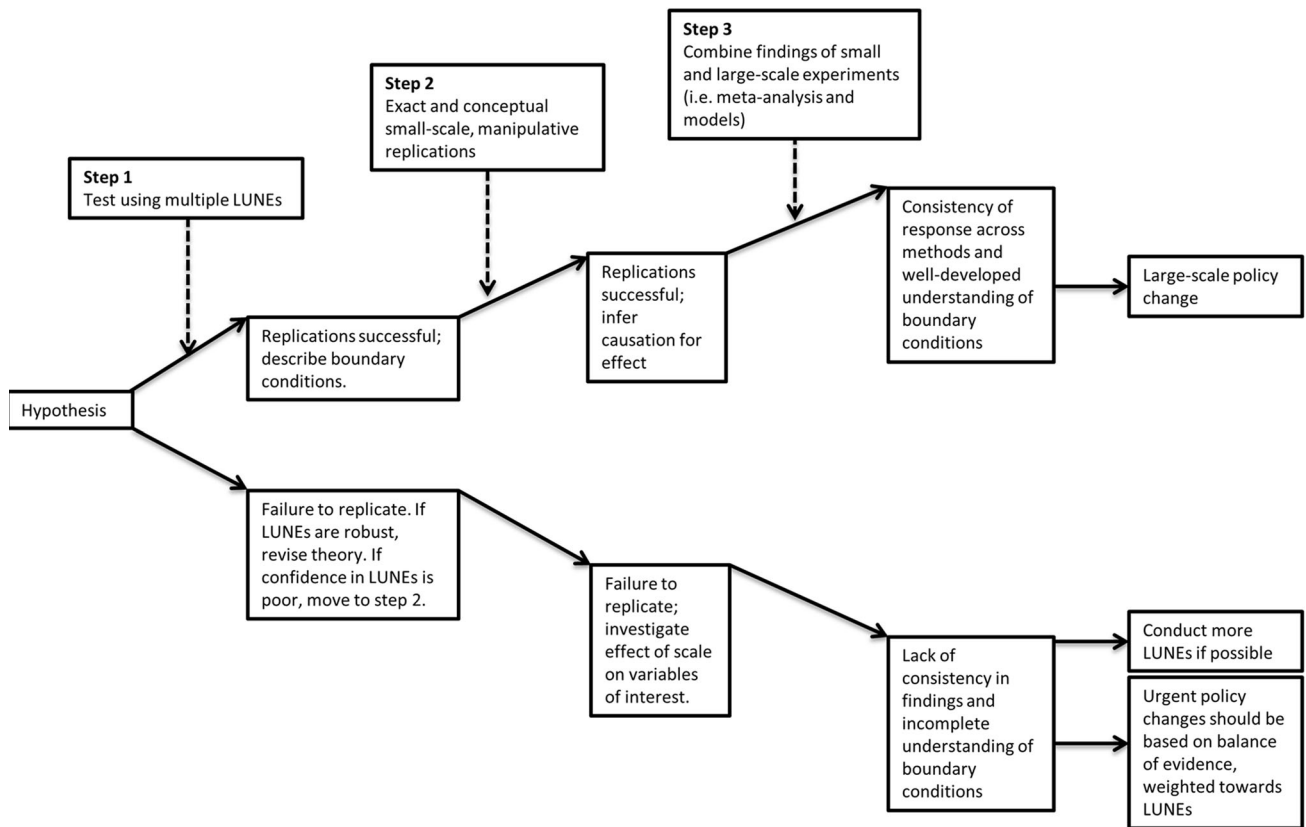


Figure 1. A new approach to replication (adapted from Hüffmeiera and others 2016).

lience to fishing. Recently, however, the findings were questioned (Grubbs and others 2016). We would argue that both studies, despite their seemingly contradictory results, are crucial steps in the accumulation of the large-scale evidence needed to test hypotheses about ecological processes. Indeed, Worm and Paine (2016) recently argued that the findings of Myers and others (2007) and Grubbs and others (2016) may be reconciled by taking into account non-trophic and/or behavioural factors.

CONCLUDING REMARKS

Despite their lack of replication, LUNEs have a unique power, not attainable in any other way, to test hypotheses at large scales and in complex systems. For example, cosmologists get by with just the one Big Bang (Gadgil and Bossert 1970), seismologists rely on sound waves generated by unpredictable earthquakes to constrain deep Earth structures (Ritsema and Van Heijst 2000) and the single submarine slump that entombed unique Cambrian fauna allows the Burgess Shale to be a key calibrant of evolutionary radiation models (Morris 1989). Of course, there are limitations to what conclusions can be drawn from LUNEs and

the findings must be viewed in the context of classical experiments, models and long-term monitoring programmes. However, as noted by Lennon (2011), “unreplicated results do not equal lies, just as replicated results do not equal truth”. The potential pitfalls introduced by diminished control can be mitigated against by a willingness to explore and rank alternative interpretations, in addition to the appropriate use of statistical tests. In line with the recent “replication crisis”, LUNEs also provide a way to conceptually replicate the findings of smaller-scale, controlled experiments.

Demonstration that human activities are altering ecological processes is often necessary for policy changes to occur (Morrisette 1989). In cases where rapid action is required to avert environmental disaster, we would therefore argue that LUNEs and other large-scale studies should be prioritised (Figure 1). An illustration of this point is the dramatic decline in water quality in rivers and lakes in North America in the 1960s. Whereas small-scale experiments implicated nitrogen and carbon, David Schindler’s large-scale, unreplicated and manipulative experiments on lakes eventually persuaded the Canadian government and several US states to ban phosphorus (Schindler 1998). Similarly, the

discovery of a hole in the ozone layer precipitated a global consensus to ban chlorofluorocarbons, even though evidence based on small-scale lab experiments and modelling had existed for almost 20 years (Morrisette 1989). These examples suggest that when an environmental threat is immediate and significant, policy-makers will act on large-scale evidence and that these actions are largely appropriate. Our review suggests that we must urgently fast-track such LUNES if we are to anticipate ecosystem-level feedbacks to contemporary perturbations.

ACKNOWLEDGMENTS

We would like to thank the University of Western Australia, Perth, which partly funded this research via an International Postgraduate Research Scholarship to SB.

REFERENCES

- Allison GW, Lubchenco J, Carr MH. 1998. Marine reserves are necessary but not sufficient for marine conservation. *Ecol Appl* 8:79–92.
- Atwood TB, Connolly RM, Ritchie EG, Lovelock CE, Heithaus MR, Hays GC, Fourqurean JW, Macreadie PI. 2015. Predators help protect carbon stocks in blue carbon ecosystems. *Nat Clim Chang* 5:1038–45.
- Bell JD, Westoby M. 1986. Variation in seagrass height and density over a wide spatial scale: effects on common fish and decapods. *J Exp Mar Biol Ecol* 104:275–95.
- Bence JR, Stewart-Oaten A, Schroeter SC. 1996. Estimating the size of an effect from a before-after-control impact paired series design. In: Russell J, Schmitt CWO, Eds. *Detecting ecological impacts: concepts and applications in coastal habitats*. San Diego: Academic Press.
- Bennett S, Bellwood DR. 2011. Latitudinal variation in macroalgal consumption by fishes on the Great Barrier Reef. *Mar Ecol Prog Ser* 426:241–52.
- Brown GP, Phillips BL, Shine R. 2011. The ecological impact of invasive cane toads on tropical snakes: field data do not support laboratory-based predictions. *Ecology* 92:422–31.
- Carpenter RC. 1990. Mass mortality of *Diadema antillarum*. I. Long-term effects on sea urchin population-dynamics and coral reef algal communities. *Mar Biol* 104:67–77.
- Carpenter SR, Chisholm SW, Krebs CJ, Schindler DW, Wright RF. 1995. Ecosystem experiments. *Science* 269:324–7.
- Carpenter SR. 1989. Replication and treatment strength in whole-lake experiments. *Ecology* 70:453–63.
- Chave J. 2013. The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecol Lett* 16:4–16.
- Chown SL, Gaston KJ, Robinson D. 2004. Macrophysiology: large-scale patterns in physiological traits and their ecological implications. *Funct Ecol* 18:159–67.
- Clark JS, Gelfand AE. 2006. A future for models and data in environmental science. *Trends Ecol Evol* 21:375–80.
- Cottenie K, De Meester L. 2003. Comment to Oksanen (2001): reconciling Oksanen (2001) and Hurlbert (1984). *Oikos* 100:394–6.
- Darimont CT, Fox CH, Bryan HM, Reimchen TE. 2015. The unique ecology of human predators. *Science* 349:858–60.
- Edgar G, Stuart-Smith R, Willis T, Kininmonth S, Baker S, Banks S, Barrett N, Becerro M, Bernard A, Berkhout J, Buxton C, Campbell S, Cooper A, Davey M, Edgar S, Försterra G, Galván D, Irigoyen A, Kushner D, Moura R, Parnell P, Shears N, Soler G, Strain E, Thomson R. 2014. Global conservation outcomes depend on marine protected areas with five key features. *Nature* 506:216–20.
- Ehrlich PR, Ehrlich AH. 2013. Can a collapse of global civilization be avoided? *Proc Biol Sci* 280:1–9.
- Estes JA, Palmisano JF. 1974. Sea otters: their role in structuring nearshore communities. *Science* 185:1058–60.
- Estes JA, Terborgh J, Brashares JS, Power ME, Berger J, Bond WJ, Carpenter SR, Essington TE, Holt RD, Jackson JBC, Marquis RJ, Oksanen L, Oksanen T, Paine RT, Pikitch EK, Ripple WJ, Sandin SA, Scheffer M, Schoener TW, Shurin JB, Sinclair ARE, Soulé ME, Virtanen R, Wardle DA. 2011. Trophic downgrading of planet Earth. *Science* 333:301–6.
- Ewers R, Didham R, Fahrig L, Ferraz G, Hector A, Holt R, Kapos V, Reynolds G, Sinun W, Snaddon J, Turner E. 2011. A large-scale forest fragmentation experiment: the stability of altered forest ecosystems project. *Philos Trans Royal Soc B: Biol Sci* 366:3292–302.
- Fisher RA. 1926. The arrangement of field experiments. *J Minist Agri* 33:503–13.
- Folt CL, Nislow KH, Power ME. 1998. Implications of temporal and spatial scale for Atlantic salmon (*Salmo salar*) research. *Can J Fish Aquat Sci* 55:9–21.
- Fraser DF, Gilliam JF. 1987. Feeding under predation hazard: response of the guppy and Hart's rivulus from sites with contrasting predation hazard. *Behav Ecol Sociobiol* 21:203–9.
- Gabric AJ, Cropp R, McTainsh G, Butler H, Johnston BM, O'Loingsigh T, Van Tran D. 2015. Tasman Sea biological response to dust storm events during the austral spring of 2009. *Marine and Freshwater Research* 9999.
- Gadgil M, Bossert WH. 1970. Life historical consequences of natural selection. *Am Naturalist* 104:1–24.
- Grossman J, Mackenzie FJ. 2005. The randomized controlled trial: gold standard, or merely standard? *Perspect Biol Med* 48:516–34.
- Grubbs RD, Carlson JK, Romine JG, Curtis TH, McElroy WD, McCandless CT, Musick CFC & JA. 2016. Critical assessment and ramifications of a purported marine trophic cascade. *Sci Rep*:20970.
- Hale M, Rivkin RB. 2007. Interpreting the results of oceanic mesoscale enrichment experiments: caveats and lessons from limnology and coastal ecology. *Limnol Oceanogr* 52:912–16.
- Hargrove WW, Pickering J. 1992. Pseudoreplication: a sine qua non for regional ecology. *Landsc Ecol* 6:251–8.
- Hildrew AG, Woodward G, Winterbottom JH, Orton S. 2004. Strong density dependence in a predatory insect: large-scale experiments in a stream. *J Anim Ecol* 73:448–58.
- Hillerislambers J, Ettinger AK, Ford KR, Haak DC, Horwith M, Miner BE, Rogers HS, Sheldon KS, Tewksbury JJ, Waters SM, Yang S. 2013. Accidental experiments: ecological and evolutionary insights and opportunities derived from global change. *Oikos* 122:1649–61.
- Hixon MA, Anderson TW, Buch KL, Johnson DW, Mcleod JB, Stallings CD. 2012. Density dependence and population regulation in marine fish: a large-scale, long-term field manipulation. *Ecol Monogr* 82:467–89.

- Högberg P, Nordgren A, Buchmann N, Taylor AF, Ekblad A, Högberg MN, Nyberg G, Ottosson-Löfvenius M, Read DJ. 2001. Large-scale forest girdling shows that current photosynthesis drives soil respiration. *Nature* 411:789–92.
- Hüffmeiera J, Mazeia J, Schultze T. 2016. Reconceptualizing replication as a sequence of different studies: a replication typology. *J Exp Soc Psychol*. doi:10.1016/j.jesp.2015.09.009.
- Hurlbert SH. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 54:187.
- Hurlbert SH. 2004. On misinterpretations of pseudoreplication and related matters: a reply to Oksanen. *Oikos* 104:591–7.
- Ioannidis JPA. 2005. Why most published research findings are false. *PLoS Med* 2:0696–701.
- Johnson DH. 2002. The importance of replication in wildlife research. *J Wildl Manag* 66:919–32.
- Johnson DH. 2006. The many faces of replication. In: *Crop science*. Vol. 46. pp 2486–91.
- Kates RW, Clark WC. 1996. Environmental Surprise: expecting the Unexpected? *Environ: Sci Policy Sustain Dev* 38:6–34.
- Kessler JD. 2011. A persistent oxygen anomaly reveals the fate of spilled methane in the deep gulf of Mexico. *Science* 331:312.
- Knowlton N. 1992. Thresholds and multiple stable states in coral reef community dynamics. *Am Zool* 32:674–82. <http://www.jstor.org/stable/pdfplus/3883648.pdf>.
- Kreyling J, Jentsch A, Beier C. 2014. Beyond realism in climate change experiments: gradient approaches identify thresholds and tipping points. *Ecol Lett* 17:125–8. <http://doi.wiley.com/10.1111/ele.12193>.
- Lennon JT. 2011. Replication, lies and lesser-known truths regarding experimental design in environmental microbiology. *Environ Microbiol* 13:1383–6.
- Lester SE, Halpern BS. 2008. Biological responses in marine no-take reserves versus partially protected areas. *Mar Ecol Prog Ser* 367:49–56.
- Leuzinger S, Luo Y, Beier C, Dieleman W, Vicca S, Körner C. 2011. Do global change experiments overestimate impacts on terrestrial ecosystems? *Trends Ecol Evol* 26:236–41.
- Lezama F, Baeza S, Altesor A, Cesa A, Chaneton EJ, Paruelo JM. 2014. Variation of grazing-induced vegetation changes across a large-scale productivity gradient. *J Veg Sci* 25:8–21.
- Likens GE, Bormann FH, Johnson NM, Fisher DW, Pierce RS. 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. *Ecol Monogr* 40:23–47.
- Madin EMP, Gaines SD, Warner RR. 2010. Field evidence for pervasive indirect effects of fishing on prey foraging behavior. *Ecology* 91:3563–71.
- McArdle BH. 1996. Levels of evidence in studies of competition, predation, and disease. *N. Z. J Ecol* 20:7–15.
- McNutt M. 2014. Reproducibility. *Science (New York)* 343:229. <http://www.sciencemag.org/content/343/6168/229.short>.
- Miao S, Carstenn S. 2006. A new direction for large-scale experimental design and analysis. *Front Ecol Environ* 4:227.
- Micheli F. 1999. Eutrophication, fisheries, and consumer-resource dynamics in marine pelagic ecosystems. *Science (New York, NY)* 285:1396–8.
- Morris SC. 1989. Burgess Shale faunas and the Cambrian explosion. *Science* 246:339–46.
- Morrisette PM. 1989. The Evolution of Policy Responses to Stratospheric Ozone Depletion. *Nat Res J* 29:793–820.
- Moss R, Watson A, Parr R. 1996. Experimental prevention of a population cycle in Red Grouse. *Ecology* 77:1512–30.
- Myers RA, Baum JK, Shepherd TD, Powers SP, Peterson CH. 2007. Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* 315:1846–50.
- Naeem S, Thompson LJ, Lawler SP, Lawton JH, Woodfin RM. 1994. Declining biodiversity can alter the performance of ecosystems. *Nature* 368:734–7.
- Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G, Contestabile M, Dafoe A, Eich E, Freese J, Glennerster R, Goroff D, Green DP, Hesse B, Humphreys M, Ishiyama J, Karlan D, Kraut A, Lupia A, Mabry P, Madon TA, Malhotra N, Mayo-Wilson E, McNutt M, Miguel E, Paluck EL, Simonsohn U, Soderberg C, Spellman BA, Turitto J, VandenBos G, Vazire S, Wagenmakers EJ, Wilson R, Yarkoni T. 2015. Promoting an open research culture. *Science* 348:1422–5.
- O'Dowd DJ, Green PT, Lake PS. 2003. Invasional 'meltdown' on an oceanic island. *Ecol Lett* 6:812–17.
- Pace ML, Cole JJ, Carpenter SR. 1998. Trophic cascades and compensation: differential responses of microzooplankton in whole-lake experiments. *Ecology* 79:138–52.
- Pauly D, Christensen V, Dalsgaard J, Froese R, Torres F. 1998. Fishing down marine food webs. *Science (New York, NY)* 279:860–3. <http://www.ncbi.nlm.nih.gov/pubmed/9452385>.
- Power ME, Dietrich WE, Sullivan KO. 1998. Experiment, observation, and inference in river and watershed investigations. In: Bernardo J, Reserits WJ, Eds. *Experimental ecology: issues and perspectives*. Oxford: Oxford University Press. p 113–32.
- Prevedello J, Dickman C, Vieira M, Vieira E. 2013. Population responses of small mammals to food supply and predators: a global meta-analysis. *J Anim Ecol* 82:927–36.
- Raffaelli D, Moller H. 1999. Manipulative field experiments in animal ecology: do they promise more than they can deliver? *Adv Ecol Res* 30:299–338.
- Ritsema J, Van Heijst HJ. 2000. Seismic imaging of structural heterogeneity in Earth's mantle: evidence for large-scale mantle flow. *Science progress* 83.
- Rogers H, HilleRisLambers J, Miller R, Tewksbury J. 2012. 'Natural experiment' demonstrates top-down control of spiders by birds on a landscape level. *PLoS One* 7:1–8.
- Rose GA, Leggett WC. 1990. The importance of scale to predator-prey spatial correlations: an example of Atlantic fishes. *Ecology* 71:33–43.
- Rowan R, Knowlton N, Baker A, Jara J. 1997. Landscape ecology of algal symbionts creates variation in episodes of coral bleaching. *Nature* 388:265–9.
- Ruppert JLW, Travers MJ, Smith LL, Fortin MJ, Meekan MG. 2013. Caught in the middle: combined impacts of shark removal and coral loss on the fish communities of coral reefs. *PLoS One* 8:e74648.
- Sarnelle O. 1997. Daphnia effects on microzooplankton: comparisons of enclosure and whole-lake responses. *Ecology* 78:913–28.
- Schindler DW, Armstrong FAJ, Holmgren SK, Brunskill GJ. 1971. Eutrophication of lake 227, experimental lakes area, northwestern Ontario, by addition of phosphate and nitrate. *J Fish Res Board Can* 28:1763–82.
- Schindler DW, Fee EJ, Rusczyński T. 1978. Phosphorus input and its consequences for phytoplankton standing crop and production in the experimental lakes area and in similar lakes. *J Fish Res Board Can* 35:190–6.

- Schindler DW, Hecky RE, Findlay DL, Stainton MP, Parker BR, Paterson MJ, Beaty KG, Lyng M, Kasian SEM. 2008. Eutrophication of lakes cannot be controlled by reducing nitrogen input: results of a 37-year whole-ecosystem experiment. *Proc Natl Acad Sci U. S. A.* 105:11254–8.
- Schindler DW. 1998. Replication versus realism: the need for ecosystem-scale experiments. *Ecosystems* 1:323–34.
- Schindler DW. 2012. The dilemma of controlling cultural eutrophication of lakes. *Proc Royal Soc B: Biol Sci* 279:4322–33.
- Shurin JB, Borer ET, Seabloom EW, Anderson K, Blanchette CA, Broitman B, Cooper SD, Halpern BS. 2002. A cross-ecosystem comparison of the strength of trophic cascades. *Ecol Lett* 5:785–91.
- Smith SA, Bell G, Bermingham E. 2004. Cross-cordillera exchange mediated by the panama canal increased the species richness of local freshwater fish assemblages. *Proc Biol Sci* 271:1889–96.
- Stewart-Oaten A, Bence JR, Osenberg CW. 1992. Assessing effects of unreplicated perturbations: no simple solutions. *Ecology* 73:1396–404.
- Stewart-Oaten A, Murdoch WW, Parker KR. 1986. Environmental impact assessment: 'pseudoreplication' in time? *Ecology* 67:929–40.
- Stroebe W, Strack F. 2014. The alleged crisis and the illusion of exact replication. *Perspect Psychol Sci* 9:59–71.
- Vanni MJ, Luecke C, Kitchell JF, Allen Y, Temte J, Magnuson JJ. 1990. Effects on lower trophic levels of massive fish mortality. *Nature* 344:333–5.
- Walsh SM, Hamilton SL, Ruttenberg BI, Donovan MK, Sandin SA. 2012. Fishing top predators indirectly affects condition and reproduction in a reef-fish community. *J Fish Biol* 80:519–37.
- Wardle DA, Barker GM. 1997. Competition and herbivory in establishing grassland communities: implications for plant biomass, species diversity and soil microbial activity. *Oikos* 80:470–80.
- Weimin X, Peet RK. 2011. The complexity of catastrophic wind impacts on temperate forest. In: Lupo A, Ed. Recent hurricane research - climate, dynamics, and societal impacts. InTech. pp 503–34.
- Worm B, Paine RT. 2016. Humans as a hyperkeystone species. *Trends in Ecology and Evolution* 2.
- Zoback MD, Gorelick SM. 2012. Earthquake triggering and large-scale geologic storage of carbon dioxide. *Proc Natl Acad Sci* 109:10164–8.