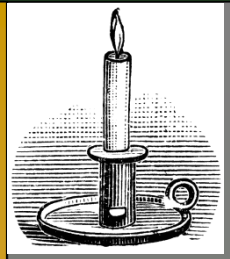
THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Office of Energetics



David B. Allison, Ph.D.

With Special Thanks to Drs. Cynthia Kroeger, Kathryn Kaiser, and Andrew Brown

Nutrition Obesity Research Center

Breaches in Research Reproducibility: Contributing Factors, Examples, & Plausible Prophylaxis

Email: Dallison@UAB.edu

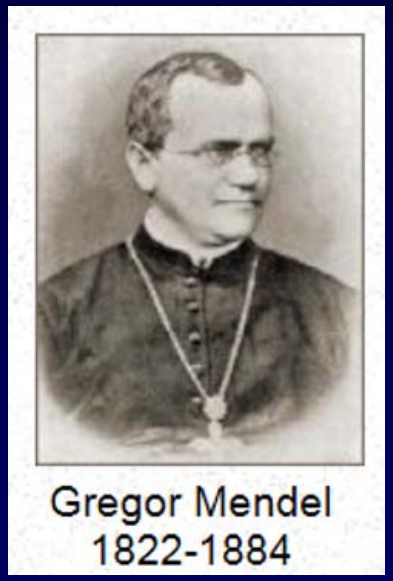
Slides Available on Request

Outline

- Ways science can go wrong
- Factors that lead to problems in science
- Examples
- Solutions
- Macro-level issues

A Taxonomy of How Science Can Go Wrong

Fabrication/ Falsification



Gregor Mendel
1822-1884

Distortion/ Obfuscation?

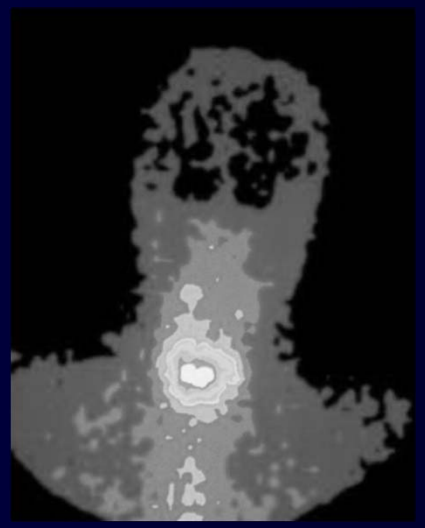
"Eddington had needed to make significant corrections to some of the measurements, for various technical reasons, and in the end decided to leave some of the Sobral data out of the calculation entirely." [1]

Error

René Blondlot
(1849-1930)

"The rays were detected by a calcium sulfide thread that glowed slightly in the dark when the rays were refracted through a 60-degree angle prism of aluminum." [2]

Substandard Methods



"By the time the radiation treatments were stopped, over 10,000 babies had died of thyroid cancer as a result of the treatments." [3]

Outline

- Ways science can go wrong
- Factors that lead to problems in science
- Examples
- Solutions
- Macro-level issues

Factors that can Lead to Breaches in Reproducibility or Rigor

- Motivational
  - Insufficient Scientific Motivation *(defined as the desire to pursue and communicate truth through scientific methods)*
  - Competing Extra-Scientific Motivations
    - Individual & Institutional
- Ignorance
- Limited Resources
- Cognitive Biases

Competing Extra-Scientific Motivational Factors

Personal

- Previous public statements
  - Willingness to revise opinions
  - Support of scientific community
- Financial
  - Influence on statements of holder
  - Fear of ad hominem attacks
- Career advancements
- Personal aggrandizement
- Unreasonable belief in validation by peer review process

Institutional

- Instead of quality of conduct, incentive systems often reward:
  - Quantity of publications
  - Exciting topics
  - First discoveries
- Institutional circumstances require grant acquisition for career continuation or advancement and not just to conduct research

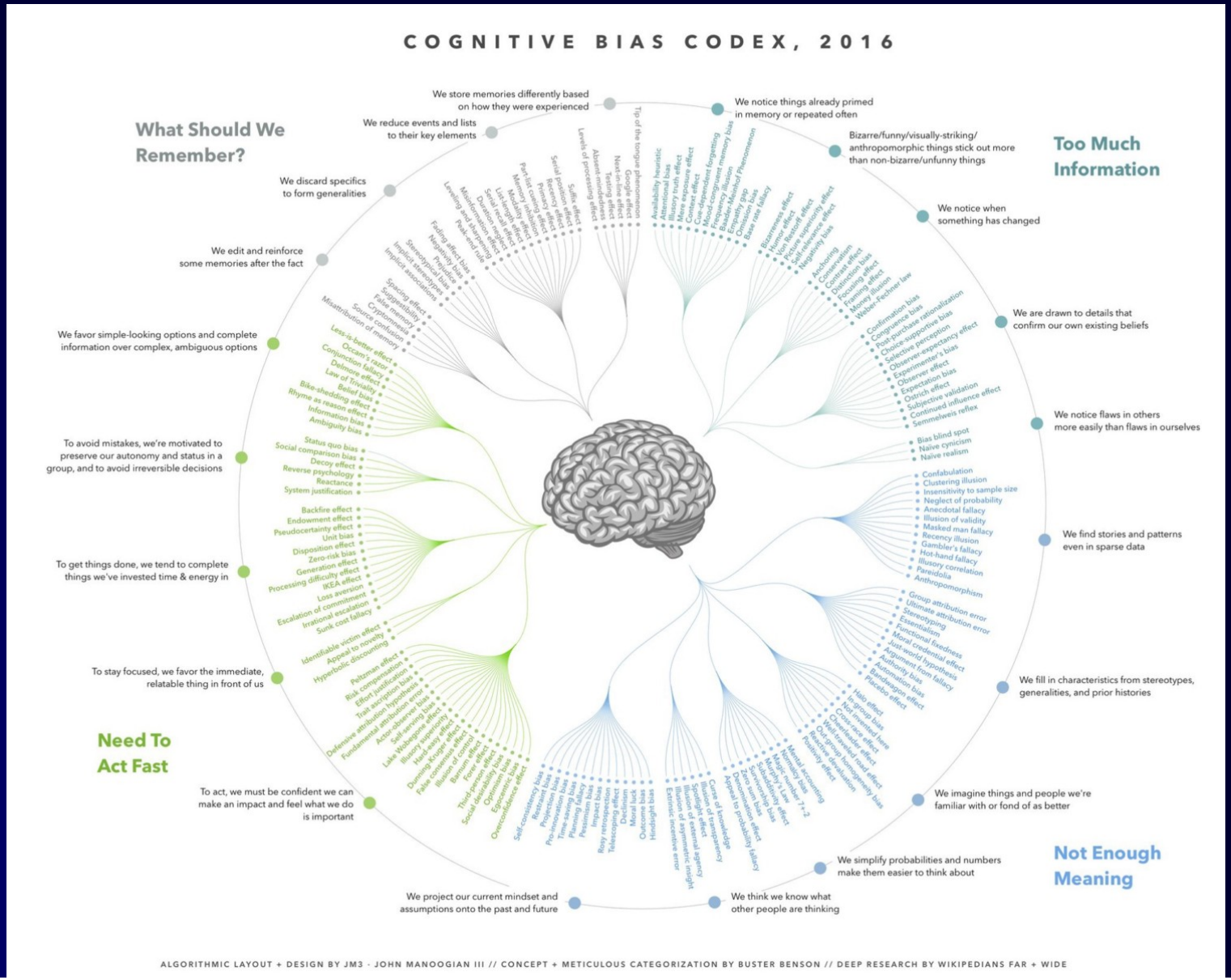
Ignorance

- Lack of or ineffective training
- Errors of interdisciplinarity
  - We cannot ask about what we do not even know that we do not know

Limited Resources

- Time (e.g., for training, scrupulosity, collaboration)
- Money (e.g., for statistician, gold-standard methods)
- Opportunity (e.g., platform, infrastructure)
  - Difficult to communicate critical information to all stakeholders (e.g., to those who do not know to look)
  - Limits when people do look (e.g., expanding literature, inconsistent indexing, limited access to literature)
  - Even when knowledgeable and willing, infrastructure needed to implement changes may be lacking or defective
  - Platforms to overcome ‘market failures’ in labor capital

Cognitive Biases



Outline

- Ways science can go wrong
- Factors that lead to problems in science
- Examples
- Solutions
- Macro-level issues

A Working Taxonomy

Errors of Measurement

- Self-reported energy intake
- Self-reported weights
- Weights of unknown origin

Errors of Design

- Gratuitous replication
- Cluster randomized trials with no *df*
- Lack of blinding
- Lack of control for non-specific factors (failure to isolate the independent variable)

Errors of Analysis

- Inappropriate baseline testing in parallel groups RCTs
- Failure to appropriately manage missing data
- Cluster randomized trials without clustering taken into account
- Investigator *df*, data fiddling, and undisclosed multiple testing

Errors of Reporting

- Publication Bias
- Reporting Bias
- Citation Bias

Errors of Interpretation

- Causal language inappropriately used in observational studies
- Extrapolation error
- 3500 kcal rule
- Conflating surrogate markers with outcomes of interest
- Not control-correcting in RCTs
- Conflating empirical evidence with tastes and moral values
- Ignoring statistical significance

We use the word ‘errors’ without implication as to intentionality or lack thereof.

Nostra Culpa: We have committed some of these errors, too.

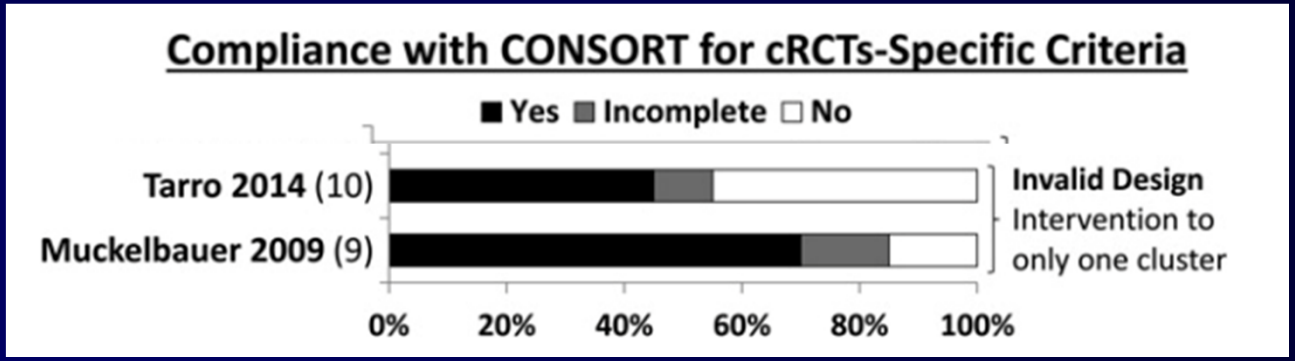
Detailed references to examples and explanations are provided upon request.

Errors of Design

Cluster Randomized Trials
Less than Two Clusters Per Condition

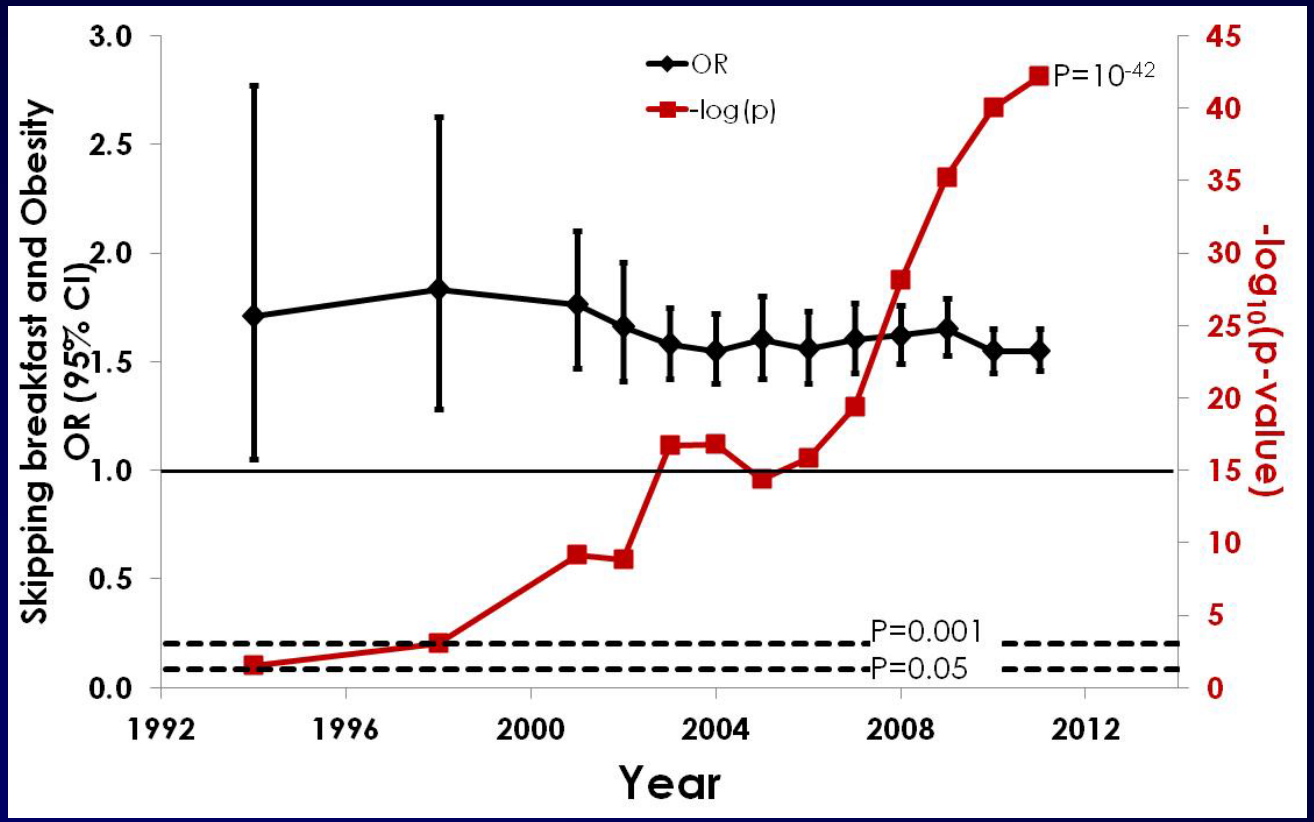
Best (but oft-forgotten) practices: designing, analyzing, and reporting cluster randomized controlled trials[1,2]

Andrew W Brown,[3,4] Peng Li,[3] Michelle M Bohan Brown,[6] Kathryn A Kaiser,[3,4] Scott W Keith,[7] J Michael Oakes,[8] and David B Allison[3–5]**



Gratuitous Replication

Over-reliance on Observational Data



Brown AW, Bohan Brown MM, Allison DB. AJCN. 2013.

We often devote our journal pages, time, and resources to research that increases belief, instead of to research that increases knowledge.

Facilitates Mere Exposure Effect (a cognitive bias)

Does Eating Breakfast Promote Weight Loss?
6 RCTs say No

Effect of breakfast eating versus breakfast skipping on obesity related anthropometry: a systematic review

Jillian E Milanes[1], David B Allison[2], Andrew W Brown[2] and Michelle M Bohan Brown[1]

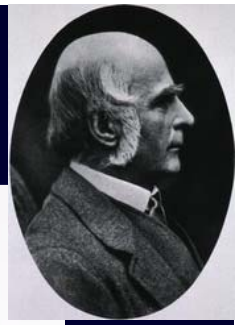
This Article

April 2016
The FASEB Journal
vol. 30 no. 1 Supplement
lb394

» Abstract

Results The search resulted in 4,047 results across databases; 3,255 after duplicates removed; and 3,182 after animal studies were removed using an automated approach. Titles and abstracts from 47 entries were reevaluated for inclusion, with 19 evaluated at the full text level. Two additional articles were identified from alternative sources. Only 6 articles met all criteria, with another 6 papers (possibly only 3 unique studies) seeming to meet all criteria except for reporting of weight. In the 6 papers that met all the criteria, 4 studies were conducted in the US, 2 in the UK; 5 studies in adults, 1 in adolescents (age 18+); study length ranged from 2–16 weeks; 4 parallel arm RCTs, 2 cross-over RCTs; 3 provided meals and 3 gave recommendations/meal plans; weight-related outcomes included fat mass, fat mass index, percent fat mass, BMI, fat free mass, waist circumference, waist:hip ratio, and body mass; and the n randomized ranged from 5 to 56 per treatment arm. Of 38 comparisons among breakfast vs skipping on weight-related outcomes, 2 were significant in favor of a breakfast, and two were significant in favor of skipping. The rest were non-significant.

Regression to the Mean



Regression to the Mean: A Commonly Overlooked and Misunderstood Factor Leading to Unjustified Conclusions in Pediatric Obesity Research

To cite this article:
Cockrell Skinner Asheley, Goldsby TaShauna U., and Allison David B.. Childhood Obesity. March 2016, 12(2): 155-158. doi:10.1089/chi.2015.0222.

Published in Volume: 12 Issue 2: March 23, 2016
Online Ahead of Print: March 14, 2016

Table 1. BMI Z-score and Z-score Change from ECLS-K

	From age 5 to age 6 (n = 17,888)		From age 11 to age 14 (n = 8278)	
	Baseline z-score	Z-score change	Baseline z-score	Z-score change
	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)
Total sample	0.387 (0.008)	−0.004 (0.005)	0.573 (0.011)	0.003 (0.007)
Healthy weight only	−0.070 (0.007)	0.058 (0.007)	−0.121 (0.010)	0.112 (0.009)
Overweight only	1.308 (0.003)	−0.174 (0.011)	1.344 (0.004)	−0.206 (0.014)
Obesity only	2.174 (0.008)	−0.183 (0.010)	2.057 (0.006)	−0.159 (0.011)

ECLS-K, Early Childhood Longitudinal Study, Kindergarten cohort; SE, standard errror.

- The observed one-year change in z-score of −0.183 for children with obesity is similar to purported "effect" sizes reported in some studies.
- Without a proper control group, changes thought to be due to an intervention may be entirely due to RTM.

Ignoring regression to the mean leads to unsupported conclusion about obesity

Asheley Cockrell Skinner ✉, Steven B Heymsfield, Angelo Pietrobelli, Myles S Faith and David B Allison

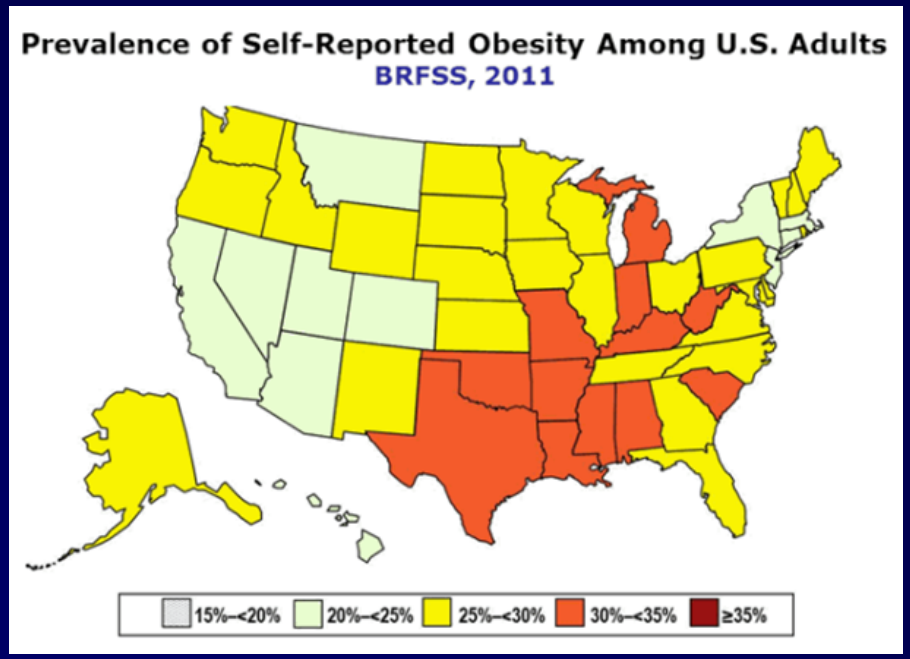
International Journal of Behavioral Nutrition and Physical Activity 2015 **12**:56 | **DOI:** 10.1186/s12966-015-0212-6

Photo from: http://resource.nlm.nih.gov/101415748

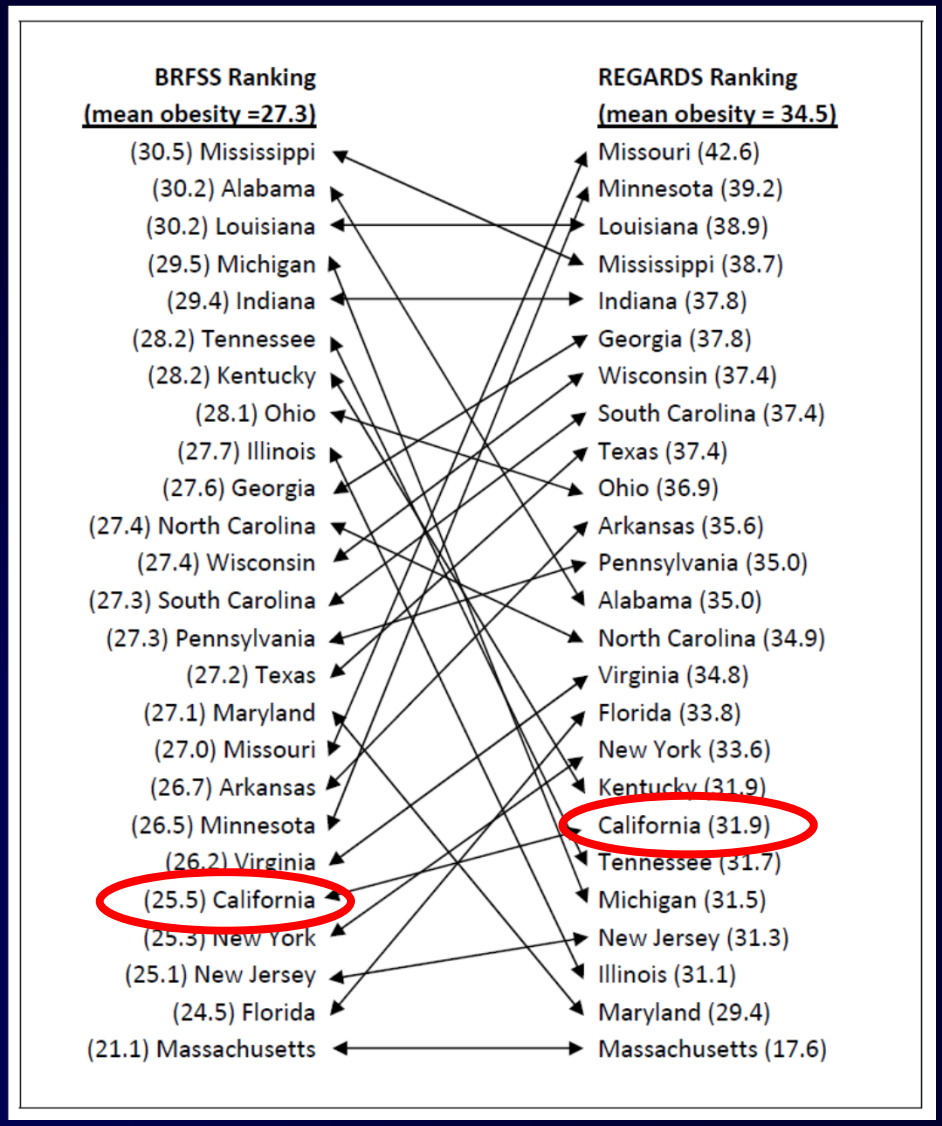
Error in Measurement

Self-Reported Body Weight

Failure to take measurement as seriously as we do in other domains



http://www.cdc.gov/obesity/data/adult.html



Le, A., Judd, S. E., Allison, D. B., Oza-Frank, R., Affuso, O., Safford, M. M., Howard, V. G., & Howard, G. (2014). The Geographic Distribution of Obesity in the US and the Potential Regional Differences in Misreporting of Obesity. *Obesity, 22,(1), 300–306.*

Self-Reported Energy Intake

Just because a measurement method one has at hand is the best available, does not make it adequate.

Energy balance measurement: when something is not better than nothing

N V Dhurandhar[1], D Schoeller[2], A W Brown[3], S B Heymsfield[4], D Thomas[5], T I A Sørensen[6], J R Speakman[7], M Jeansonne[8], D B Allison[8] and the Energy Balance Measurement Working Group[9]

International Journal of Obesity (2014), 1–5

Addressing Current Criticism Regarding the Value of Self-Report Dietary Data[1,2]

Amy F Subar,[3]* Laurence S Freedman,[5] Janet A Tooze,[6] Sharon I Kirkpatrick,[7] Carol Boushey,[8] Marian L Neuhouser,[9] Frances E Thompson,[3] Nancy Potischman,[4] Patricia M Guenther,[10] Valerie Tarasuk,[11] Jill Reedy,[3] and Susan M Krebs-Smith[3]

- do not use self-reported energy intake as a measure of true energy intake
- do use self-reported energy intake for energy adjustment of other self-reported dietary constituents to improve risk estimation in studies of diet-health associations

Error in Analysis

Inappropriate Baseline Testing in Parallel-Arm RCTs

LETTER TO THE EDITOR | OPEN ACCESS

Inappropriate statistical method in a parallel-group randomized controlled trial results in unsubstantiated conclusions

Rositsa B. Dimova and David B. Allison

Nutrition Journal 2016 15:58 | DOI: 10.1186/s12937-016-0163-z

Received: 27 February 2015 | Accepted: 18 April 2016 | Published: 6 June 2016

Impact of weight loss diet associated with flaxseed on inflammatory markers in men with cardiovascular risk factors: a clinical study

Roberta Soares Lara Cassani, Priscila Giacomo Fassini , Jose Henrique Silvah, Cristiane Maria Mártires Lima and Júlio Sérgio Marchini

Nutrition Journal 2015 **14**:5 | **DOI:** 10.1186/1475-2891-14-5

Received: 9 September 2014 | **Accepted:** 29 December 2014 | **Published:** 10 January 2015

The Retraction Note to this article has been published in *Nutrition Journal* 2016 **15**:59

Incorrect Data or Calculations of Effect Sizes in Meta-Analysis

Correction of data errors and reanalysis of "The effect of glucomannan on body weight in overweight or obese children and adults: A systematic review of randomized controlled trials"

Bartłomiej M. Zalewski, M.D., Anna Chmielewska, M.D., Ph.D., Hania Szajewska, M.D., Ph.D.
Department of Pediatrics, The Medical University of Warsaw, Warsaw, Poland
Joyce K. Keithley, D.N.Sc.
College of Nursing, Rush University Medical Center, Chicago, IL
Peng Li, Ph.D., TaShauna U. Goldsby, Ph.D., David B. Allison, Ph.D.
Office of Energetics and Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL

"The effect of treatment after 2 wk that was originally reported as statistically significant in the meta-analysis published turned out to be insignificant after the correction."

Errors of Reporting and Interpretation

Math Problems

CHILDHOOD OBESITY
February 2014 | Volume 10, Number 1

DOI: 10.1089/chi.2013.0082

Modeling Potential Effects of Reduced Calories in Kids' Meals with Toy Giveaways

Maysoun Y. Freij, PhD, MPH,[1] Randall L. Sell, ScD,[2] Anne K. Bozack, MPH,[1] Linda J. Weiss, PhD,[1] and Ana C. Garcia, MPA[3]

calculation is based on the fact that a pound of body fat equals approximately 3500 calories. Therefore, a child who eats a kids' meal with a toy twice per week would avoid approximately 132 calories per week if he or she consumed an NPLAN-like regulated kids' meal in place of an unregulated kids' meal. This equals approximately 6864 calories per year (132 calories × 52 weeks) or approximately two pounds per year averted. A lower calorie limit

nance). Calculations in the model include children who are estimated to eat fast food four or more times per day. Though rare, such children could theoretically expect to avert weight gain of 27 pounds per year if an NPLAN-like toy ordinance were fully implemented.

From Kevin Hall:

"I used the 'rule of thumb' equations relating excess body weight to excess energy intake …

I came up with about 0.5 kg of excess weight …

by increasing daily intake by 19 kcal/d from 7 to 12 years [7]."

The original investigators were off by a factor of about 9

Original article retracted [6]

A

Obesity Facts
The European Journal of Obesity

Original Article

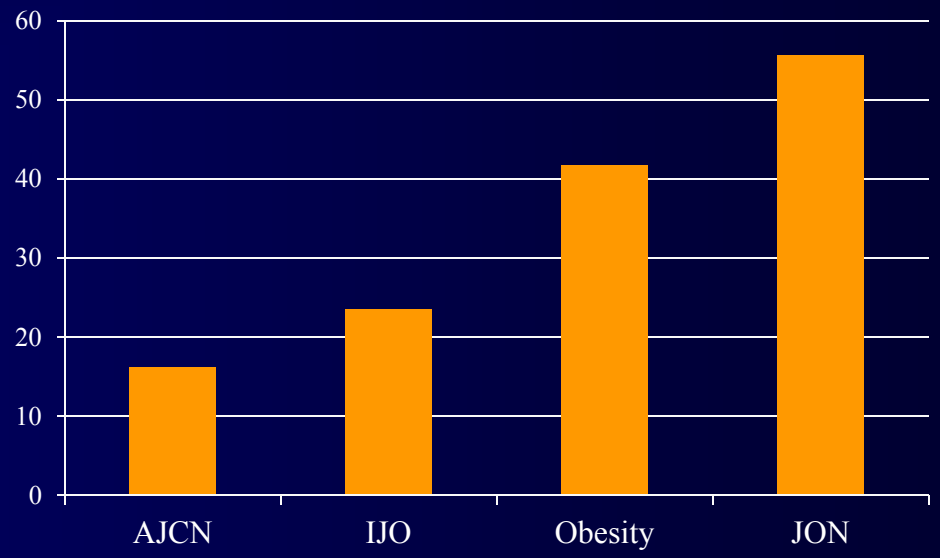
Obes Facts 2010;3:353–356
DOI: 10.1159/000322940

Published online: December 10, 2010

Use of Causal Language in Observational Studies of Obesity and Nutrition

Stacey S. Cofield[a] Rachel V. Corona[b] David B. Allison[a,c]

Percentage of Articles using Unjustified Causal Language



Unreasonable Extrapolation

Headline

Study Abstract

"The health benefits of apple bioactive compounds have been extensively reported. However, only few studies have focused on bioactive compounds that are not absorbed and metabolised during gastrointestinal digestion and can induce changes in microbial populations of faeces. We have characterised Braeburn, Fuji, Gala, Golden Delicious, Granny Smith, McIntosh and Red Delicious cultivars and found significant differences for extractable phenolics (1.08-9.2mg/g) non-extractable proanthocyanidins (3.28 5.7mg/g), and dietary fibre (20.6 32.2%) among cultivars with Granny Smith having the highest contents.

Granny Smith was used after in vitro digestion for fermentation of faeces from diet-induced obese mice.

Results showed that relative abundances of Firmicutes, Bacteroidetes, Enterococcus, Enterobacteriaceae, Escherichia coli, and Bifidobacterium in apple cultured faeces tended to resemble the abundance in faeces from lean mice with increased trend in the production of butyric acid. These results suggest that apple non-digestible compounds might help to re-establish a disturbed microbiota balance in obesity."

Assessing non-digestible compounds in apple cultivars and their potential as modulators of obese faecal microbiota in vitro. Food Chem. 2014 Oct 15;161:208-15.

Ignoring Primary Results

Headline

"A new peer-reviewed study confirms we are delivering on our mission of reducing the prevalence of childhood obesity." "an important means of supporting schools in reducing obesity." [4]

Study Abstract

"Analyses showed no difference between Healthy School Program schools and control schools in overweight and obesity prevalence."

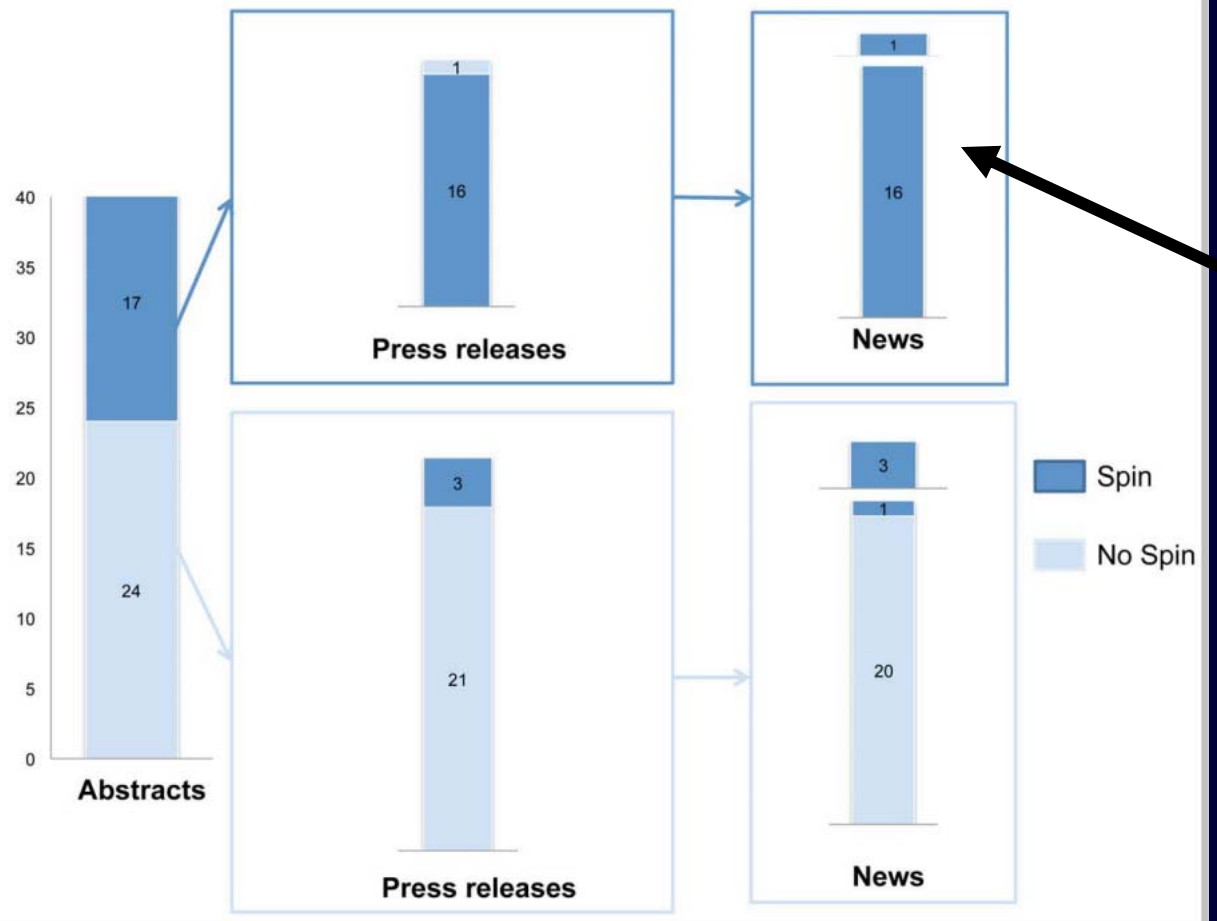
"Healthy School Program appears to be an important means of supporting schools in reducing obesity." [5]

Perpetuation of Spin

Spin: reporting strategies emphasizing the beneficial effect of the experimental treatment

Misrepresentation of Randomized Controlled Trials in Press Releases and News Coverage: A Cohort Study

Amélie Yavchitz[1,2,3], Isabelle Boutron[1,2,3]*, Aida Bafeta[1,2,3], Ibrahim Marroun[4], Pierre Charles[4], Jean Mantz[5], Philippe Ravaud[1,2,3]



Spin perpetuates throughout reporting

Idiopathic Errors

Growth in Height?

World Journal of Acupuncture-Moxibustion (*WJAM*)
Vol. 25, No.1, 30th Mar. 2015

ELSEVIER

Letters

Letter to the Editor: Exceptional Data in Paper on "The effect of meridian massage on BM, BMI, WC and HC in simple obesity patients: a randomized controlled trial"

14, November, 2014

Dear Editor:

Given the tremendous need for effective weight loss treatments, we read with interest the paper by Yan et al. [1] and were intrigued by the reported finding that a form of massage produced weight loss in a randomized controlled trial.

Upon closer inspection, we were struck by the magnitude of the results. Specifically, in an 8-week period, the treatment (massage) group lost over 7 kg, more than 9% of their baseline body weight, and 3.7 kg more than the control group lost. Such results rival those of all available obesity treatment procedures except surgery and very

Using approximations involving geometric means, we were able to show this would have required a growth in height of ~6 cm.

Journal of Paramedical Sciences (JPS) *Summer 2015 Vol.6, No.3 ISSN 2008-4978*

Errors in statistical analysis and questionable randomization lead to unreliable conclusions

Brandon J George[1], Andrew W Brown[1,2], David B Allison[1,2,3]

We read with interest the paper, "The effect of food service system modifications on staff body mass index in an industrial organization"[1]. We noticed several substantial issues with data and calculations, calling into question the randomized nature of the study and validity of analyses.

The distribution of baseline weight was significantly differentbetween groups (p-value = "0.00"). We replicated the test using reported means and standard deviations (SDs) andobtained a p-value of approximately $1.9*10^{-17}$. It is

If every one of the roughly 7 billion persons on the planet each did their own RCT and in each of these RCTs tested 1,000 separate variables for baseline differences, and we used a Bonferroni correction to adjust for the (1000*7 billion) tests done, a p-value of 10^{-17} would still be significant

Considering that the reported findings are essentially impossible given the stated study design, we encourage the authors to explain the treatment allocationand make the raw data available, or the journal to act according to the Committee on Publication Ethics[5] in situations where findings are unreliable.

Due to lack of author cooperation to provide the data used in the article"The effect of food service system modifications on staff body mass index in an industrial organization"*, it is decided to remove the article from journal.*

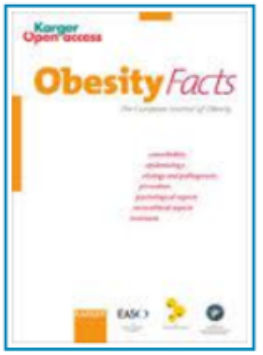
Resisting the Self-Correcting Nature of Science

Retraction Watch

Tracking retractions as

When should a paper be retracted? A tale from the obesity literature

with one comment



In our line of work, we see it all — mega-corrections that don't quite rise to the level of retraction, letters to the editor that point out seemingly fatal flaws in papers that remain untouched, and studies retracted for what seem like minor reasons. It can make you wonder what makes a paper worthy of a retraction. A recent case in an obesity journal may not provide a definitive answer, but it gives us a lot to chew on.

Here's the story: In September 2013, Rosely Sichieri and a colleague from the State University of Rio de Janeiro submitted an article to *Obesity Facts*, "Unbalanced Baseline in School-Based Interventions to Prevent Obesity: Adjustment Can Lead to Bias?" The article examined statistical issues in randomized controlled trials of school-based weight loss programs. Peer reviewers said the paper needed major revisions before it could be accepted; the authors revised the paper enough in a second draft, submitted in November 2013, that the original reviewers accepted it. The paper was published in June 2014.

Then, in September 2014, a group of authors including David Allison of the University of Alabama, Birmingham, and colleagues from Clemson, Thomas Jefferson, and the University of Minnesota, wrote a critical letter that was published in the journal in April. The letter, according to a just-published editorial:

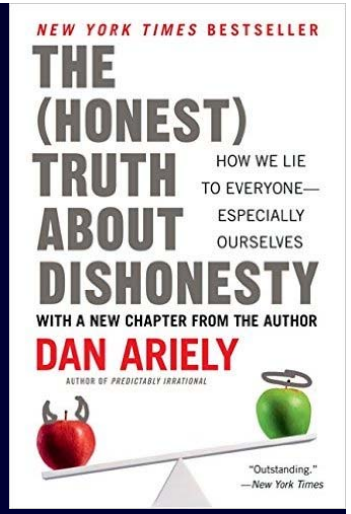
> expressed fundamental and severe criticism with regard to the above mentioned article that culminated in the conclusion that the article should be retracted.

More specifically, the letter argued that by criticizing some of the statistical tools used in these types of studies, the authors dissuade scientists from employing "legitimate power-enhancing analytic methods." Here's more from the letter itself:

Outline

- Ways science can go wrong
- Factors that lead to problems in science
- Examples
- Solutions
- Macro-level issues

Remedies for Insufficient Scientific Motivation



Articulate standards

- Current efforts
  - Trial registry
  - Reporting guidelines
  - Data-sharing
  - Professional societies
  - Transparency pledge
  - Workflow perspective

Inculcate a spirit of science

Value of a culture of Truthfulness

Proposed Remedies for Competing Extra-scientific Motivations and Cognitive Biases

- Disclosure is essential, but is an aid to trust, not an aid to trustworthiness.
- The methods of science are themselves the solutions (e.g., blinding, use of formal statistics).
- Meta-methods (e.g., study registration; raw data sharing; reporting guidelines) are essential to:
  - Ensure use of the methods of science
  - Make the methods, data, and results transparent
- Altering incentive structure to reward quality of research conception and execution over to a greater extent and specific research to a lesser extent.

Remedies for Ignorance
Education

- Current efforts
  - Training modules
  - Resource hubs
  - Workshops
- Efforts for the future
  - Emphasize philosophy of science in graduate training
  - Full course on research reproducibility, rigor, and transparency
  - Methodologic training
  - Training in interdisciplinary collaboration
  - More statisticians needed
  - Regular ‘scoring’ of the literature.

Remedies for Limited Resources

- Forego conducting an inadequate study
- Triage studies for needed degree of rigor (do all studies merit being done rigorously?)
- Crowdsourcing research
- More funding (you knew that was coming)

Outline

- Ways science can go wrong
- Factors that lead to problems in science
- Examples
- Solutions
- Macro-level issues

Macro-level Issues

- To what extent do we need evidence for proposed solutions?
- How do we assess cost/benefit analysis of implementing solutions?
  - Most suggested remedies are not currently supported by formal evidence
  - Like other realms of policy, we suggest they have sufficient face-validity to try, so long as their effectiveness and costs are adequately monitored and the state of evidence disclosed.

Come visit us in Alabama and we can talk some more on the trail

Cheaha Mountain

Photo courtesy Rohan Dhurandhar

References

1. http://ircamera.as.arizona.edu/NatSci102/NatSci102/text/lightbend.htm
2. http://skepdic.com/blondlot.html
3. http://www.orderofthegooddeath.com/poverty-the-thymus-and-sudden-infant-death-syndrome
4. http://www.wassermanfoundation.org/news/new-study-shows-that-combating-childhood-obesity-in-schools-works/
5. http://www.cdc.gov/pcd/issues/2015/15_0020.htm
6. http://www.ncbi.nlm.nih.gov/pubmed/25496036
7. https://www.ncbi.nlm.nih.gov/pubmed/24349967
8. Young & Karr. 2011, Significance; 8: 116-120

Disclosures

Dr. Allison has received financial and other benefits from the following entities: the Frontiers Foundation; The Federal Trade Commission; The FDA; The Nutrition Science Initiative (NUSI), and numerous additional government, non-profit and for-profit (including publishing, food, beverage, and pharmaceutical companies) organizations with interests in obesity, nutrition, and health.