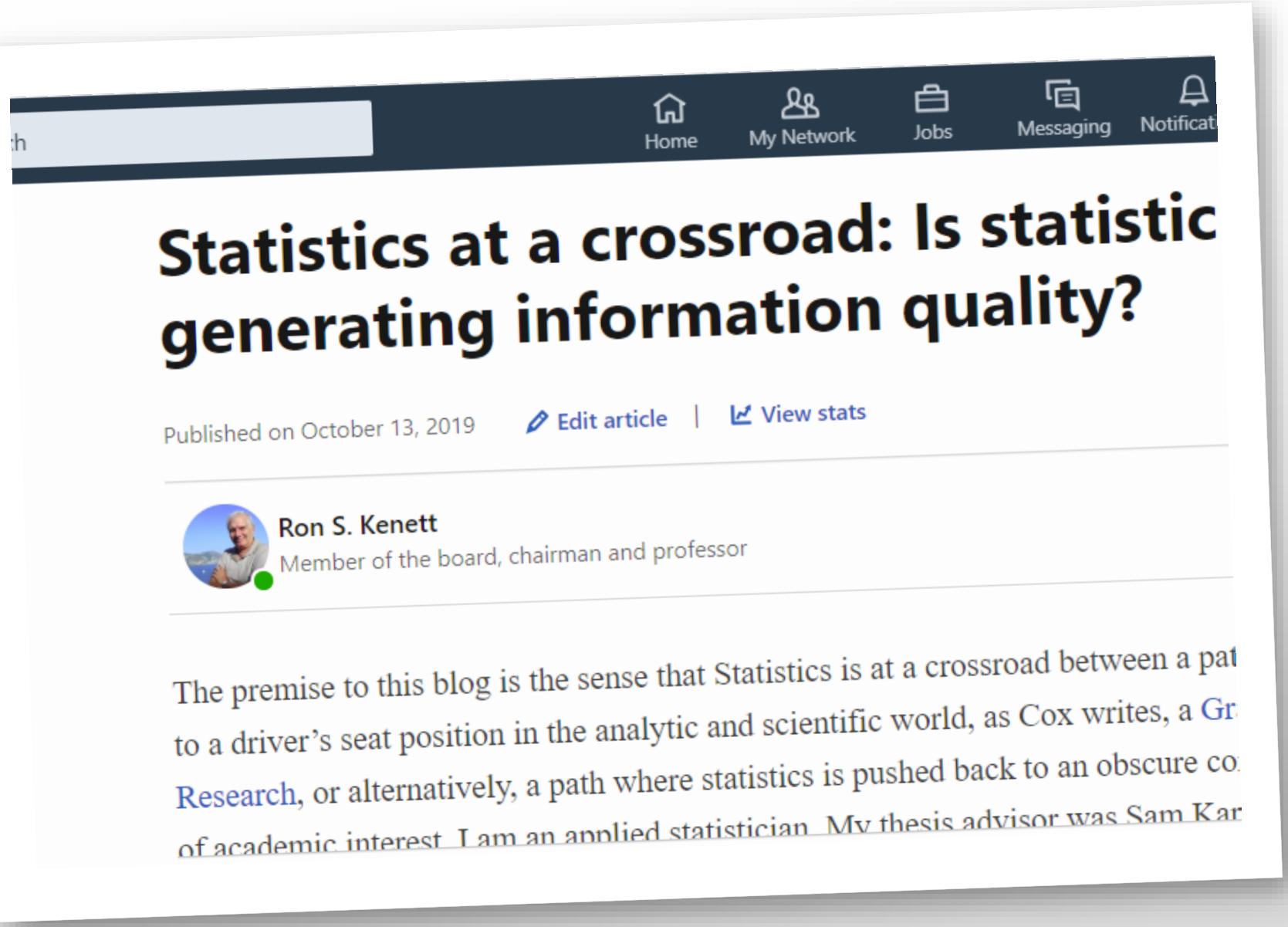
Ron S. Kenett

Statistics at a crossroad

KPA
Insights through analytics

Samuel Neaman Institute
for National Policy Research

TECHNION
Israel Institute
of Technology


Home
My Network
Jobs
Messaging
Statistics at a crossroad: Is statistic generating information quality?
Published on October 13, 2019
Edit article
View stats
Ron S. Kenett
Member of the board, chairman and professor
The premise to this blog is the sense that Statistics is at a crossroad between a pat
to a driver's seat position in the analytic and scientific world, as Cox writes, a Gr
Research, or alternatively, a path where statistics is pushed back to an obscure co
of academic interest. I am an applied statistician. My thesis advisor was Sam Kar

A pragmatic view on the role of statistics and statisticians in modern data analytics

Ron S. Kenett (KPA Ltd., Raanana, Samuel Neaman Institute, Technion, Haifa and Institute for Drug Development, The Hebrew

Search ...

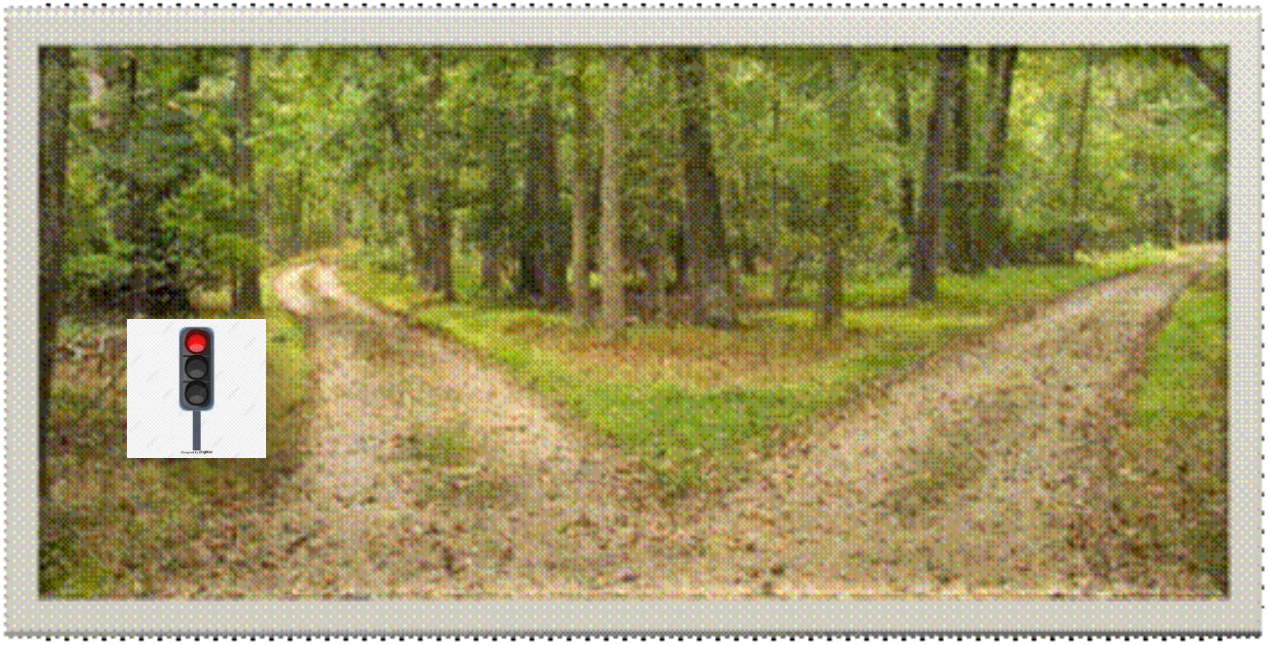
Top Posts

Statistical Models for Credit I Banking

Data Science at Alibaba

https://blogisbis.wordpress.com/2019/11/12/a-pragmatic-view-on-the-role-of-statistics-and-statisticians-in-modern-data-analytics/

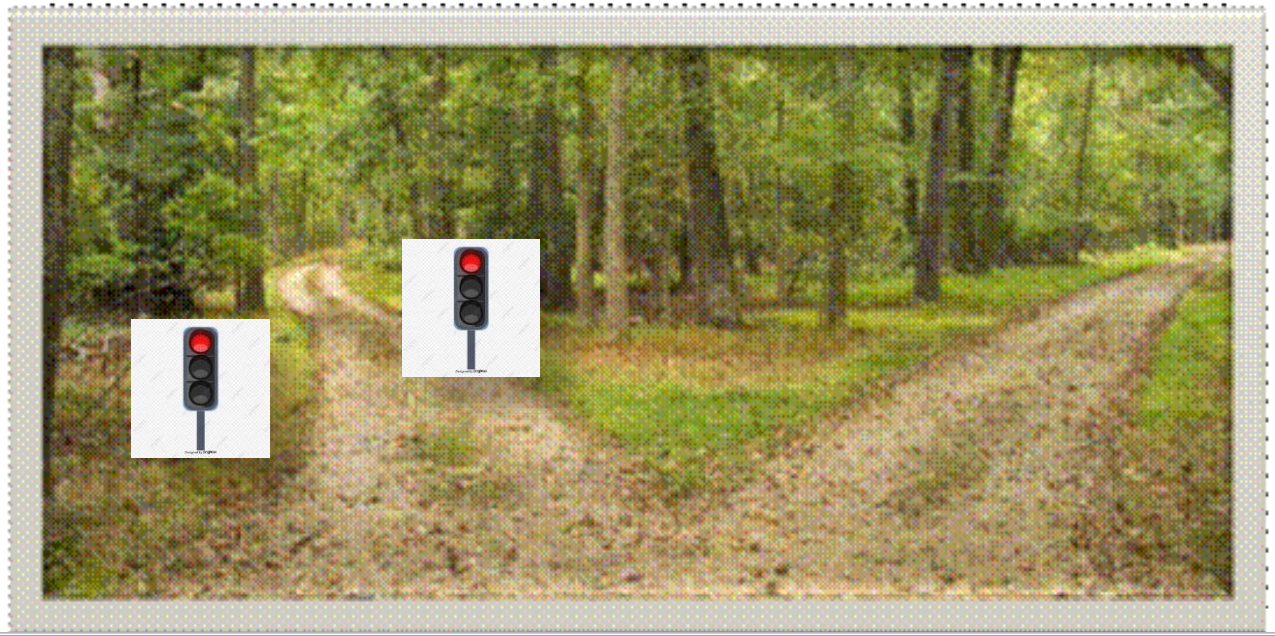
KPA Insights through analytics



https://tandfonline.com/doi/full/10.1080/01973533.2015.1012991

Basic and Applied Social Psychology

The *Basic and Applied Social Psychology* (BASP) 2014 Editorial emphasized that the null hypothesis significance testing procedure (NHSTP) is invalid, and thus authors would be not required to perform it (Trafimow, 2014). However, to allow authors a grace period, the Editorial stopped short of actually banning the NHSTP. The purpose of the present Editorial is to announce that the grace period is over. From now on, BASP is banning the NHSTP.

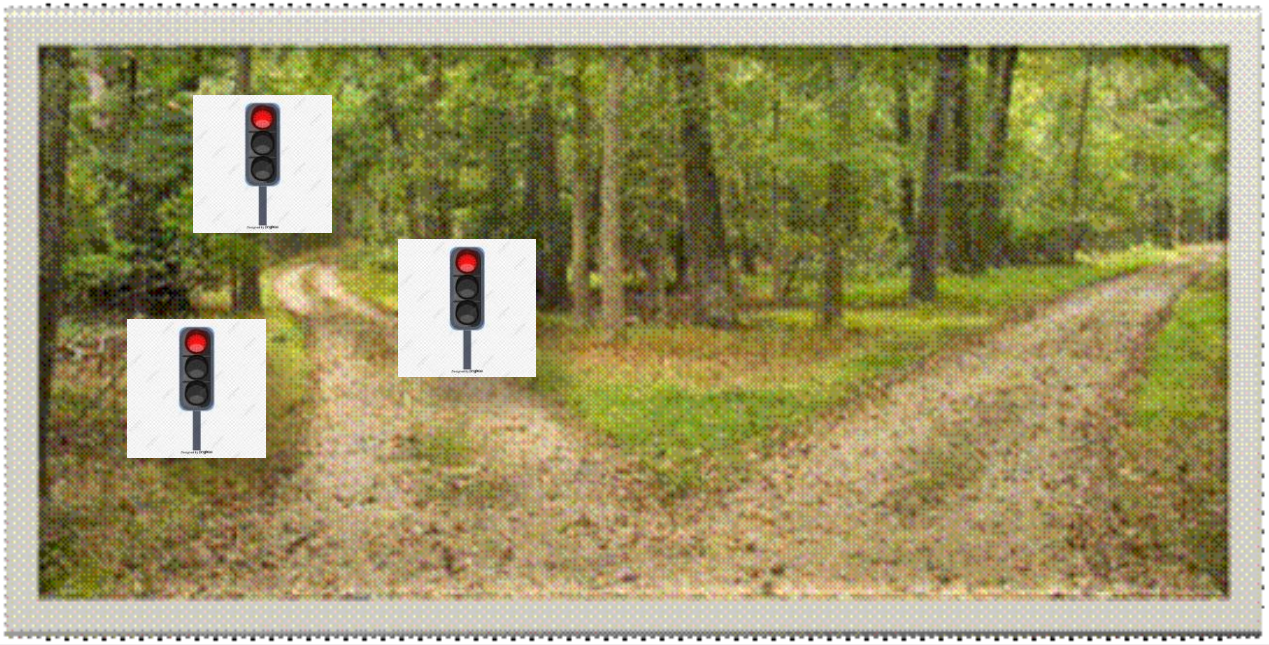


SOC ARXIV SocArXiv Papers Submit a Paper Search Donate Sign Up

New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS Research [Paper accepted at the Journal of the Association for Information Systems]

AUTHORS
Willem Mertens, Jan Recker

https://osf.io/preprints/socarxiv/5qr7v/



Journal
The American Statistician ›
Volume 73, 2019 - Issue sup1: Statistical Inference in the 21st Century: A World Beyond p < 0.05

Enter keywords, authors, DOI, ORC

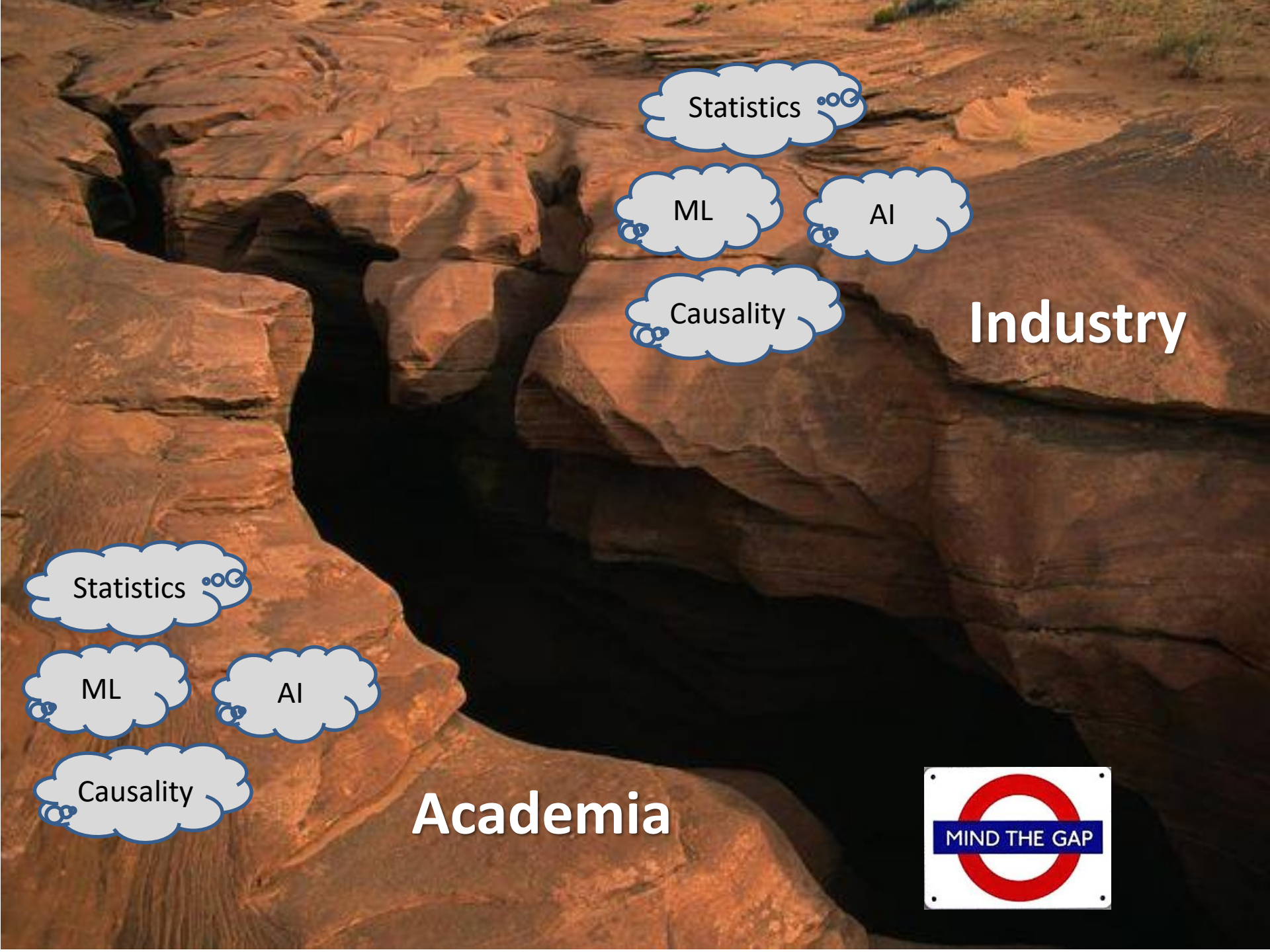
https://www.tandfonline.com/doi/full/10.1080/00031305.2019.1583913

Listen
Editorial

Moving to a World Beyond "$p < 0.05$"

Ronald L. Wasserstein, Allen L. Schirm & Nicole A. Lazar

Pages 1-19 | Published online: 20 Mar 2019


Statistics
ML
AI
Causality
Industry
Statistics
ML
AI
Causality
Academia
MIND THE GAP

Reproducible Research

Marcia McNutt is Editor-in-Chief of *Science*.

Reproducibility

SCIENCE ADVANCES ON A FOUNDATION OF TRUSTED DISCOVERIES. REPRODUCING AN EXPERIMENT is one important approach that scientists use to gain confidence in their conclusions. Recently, the scientific community was shaken by reports that a troubling proportion of peer-reviewed preclinical studies are not reproducible. Because confidence in results is of paramount importance to the broad scientific community, we are announcing new initiatives

www.sciencemag.org **SCIENCE** VOL 343 17 JANUARY 2014

Published by AAAS

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands

KPA Insights through analytics

Clarifying the terminology that describes scientific reproducibility

To the Editor: There has recently been a growing interest in discussions of reproducible/repeatable scientific research[1,2]. The scientific press appears to be witnessing a confusion of terms: reproducibility, repeatability and replicability are referred to with different and sometimes conflicting meanings, both between and within fields. We suggest that these terms can be clarified by considering the intended generalization of the study at hand.

In industrial systems, for example, reproducibility and repeatability are used in the context of 'gauge repeatability and reproducibility' (GR&R) testing for evaluating measurement error of equipment. In these experiments, several testers are asked to retest a set of items. Differences between testers under different conditions are used to estimate reproducibility, whereas repeat evaluations under identical conditions are used for estimating repeatability[3]. In contrast, the term replicability is used in genome-wide association studies to describe a repetition of a study by the same lab or researchers but with a different technology or a different data set (typically a follow-up subpopulation but possibly a different human population)[4], whereas in GR&R, such a case would be called repeatability. In machine learning and computational mathematics, experiments are used to evaluate algorithms. The common terms in these fields are reproducibility and replicability, but different researchers have different definitions[5]. One distinction is whether the exact numerical results are recreated—for instance, by rerunning the code (repeatability)—versus whether the overall result can be rederived (replicability). Finally, in preclinical studies, reproducibility often relates to recreating the same numbers by different labs, whereas in GR&R and machine learning, the same term is used to describe changing experimental conditions beyond the researchers or lab. We see that the same terms are used with different meanings in different contexts. Our goal here is to provide conceptual clarification to this situation.

different lab technicians or test environments (scientific generalization), and therefore both test conditions and testers are varied. Poor reproducibility calls for considering the overall measurement process, including operating procedures and provided training.

As an example from biological studies, we consider the recent criticism of standardization in animal behavior experiments[7]. The authors show that, in contrast to standardization being beneficial, introducing systematic variation of experimental conditions (which they call "heterogenization") may attenuate spurious results and improve reproducibility[8]. Considering this from the standpoint of generalization clarifies the issue. Standardized animal behavior experiments are differently generalizable than experiments with induced systematic variation of experimental conditions. In particular, standardization intends statistical generalization, whereas heterogenization intends scientific generalization.

In summary, although terminology can remain domain specific, we propose that researchers should clearly state the intended generalization of their study. Such an approach will clarify the implications of a study within and across fields.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Ron S Kenett[1,2] & Galit Shmueli[3]

[1]University of Turin, Turin, Italy. [2]The KPA Group, Raanana, Israel. [3]Institute of Service Science, National Tsing Hua University, Hsinchu, Taiwan.
e-mail: ron@kpa-group.com

1. McNutt, M. *Science* **343**, 229 (2014).
2. Banks, D. *Stat. Politics Policy* **2**, doi:10.2202/2151-7509.1023 (2011).
3. Kenett, R.S., Zacks, S. & Amberti, D. *Modern Industrial Statistics: With Applications in R, MINITAB and JMP* 2nd edn. (Wiley, 2014).
4. Ioannides, J.P. *et al. Nat. Genet.* **41**, 149–155 (2009).
5. Drummond, C., Japkowicz, N., Klement, W. & Macskassy, S.A. in *Proc. 26th. Int. Conf. Mach. Learn.* doi:10.1145/1553374.1553546 (ACM, 2009).
6. Kenett, R.S. & Shmueli, G. *J. R. Stat. Soc. Ser. A Stat. Soc.* **177**, 3–38 (2014).
7. Richter, S.H., Garner, J.P. & Würbel, H. *Nat. Methods* **6**, 257–261 (2009).
8. Richter, S.H., Garner, J.P., Auer, C., Kunert, J. & Würbel, H. *Nat. Methods* **7**, 167–168 (2010).

Kenett, R.S. (2015). On generalization of scientific reproducibility, *Nature Methods*, Vol. 12(8), p 699.

Reproducibility versus Replicability

Replicability is not Reproducibility: Nor is it Good Science

Chris Drummond
Institute for Information Technology
National Research Council Canada
Ottawa, Ontario, Canada, K1A 0R6

CHRIS.DRUMMOND@NRC-CNRC.GC.CA

Proc. of the Evaluation Methods for Machine Learning Workshop at the 26 th ICML, Montreal, Canada, 2009.

"Reproducibility requires changes; replicability avoids them. A critical point of reproducing an experimental result is that irrelevant things are intentionally not replicated. One might say, one should replicate the result not the experiment."

A highly standardized experiment supplies direct information only in respect of the narrow range of conditions achieved by standardization. Standardization, therefore, weakens rather than strengthens our ground for inferring a result, when, as is the case in practice, these conditions are somewhat varied.

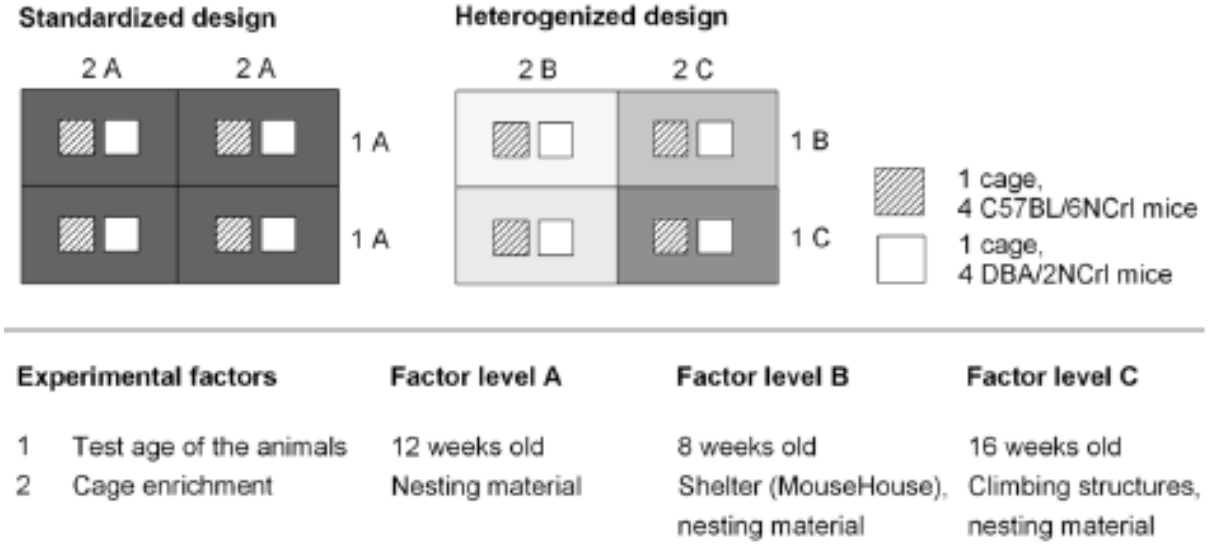
Ronald A. Fisher 1935

KPA Insights through analytics

Reproducibility in Animal Behavior

- Standardization is the attempt to increase reproducibility at the expense of external validity
- Standardization **reduces** external validity and thus also reproducibility
- Heterogenization **increases** external validity and thus also reproducibility

Würbel et al. 2000 Nature Genetics
Richter et al. 2010 Nature Methods
Richter et al. 2011 PLoS ONE

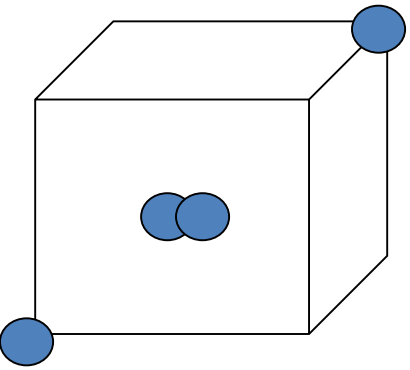


Experimental factors		Factor level A	Factor level B	Factor level C
1	Test age of the animals	12 weeks old	8 weeks old	16 weeks old
2	Cage enrichment	Nesting material	Shelter (MouseHouse), nesting material	Climbing structures, nesting material

KPA
Insights through analytics

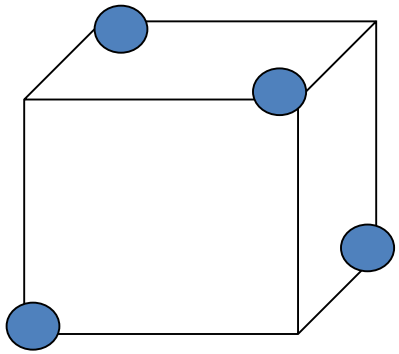
Design of Experiments Strategy

Are Results Reproducible?



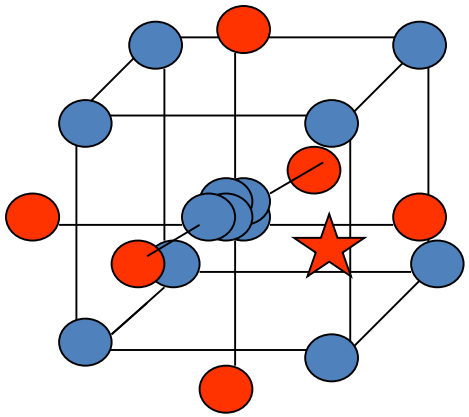
Scoping

Initial assessment



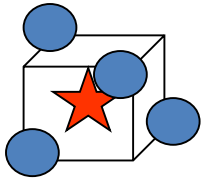
Screening

Fractional designs



Optimizing

Response surfaces



Robustness

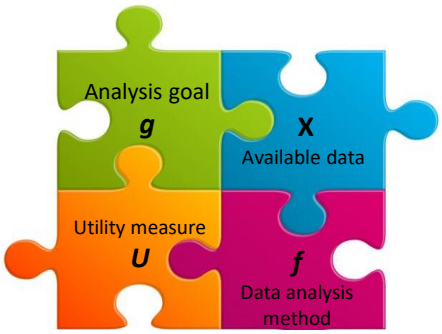
Robust designs

Gain Knowledge

Build Confidence

Information Quality

The Potential of Data and Analytics to Generate Knowledge



What

WILEY https://www.wiley.com//legacy/wileychi/kenett/presentation.html?type=SupplementaryMaterial Contact Us Tech Support

Kenett, Shmueli:
Information Quality: The Potential of Data and Analytics to Generate Knowledge

Home | Resources | More Information

Presentations on InfoQ

requires Adobe Acrobat Reader

Title	Location	Date
Do you want to make an impact with quantitative methods? Make sure you generate high InfoQ	Toulon-Verona Conference, Israel	September 3, 2012
A Workshop on Modern Analysis of Customer Satisfaction Surveys	22nd Colombian Statistics Symposium, The National University of Colombia Bucaramanga, Colombia	July 17, 2012
Quantitative and Qualitative Aspects of Bayesian Networks: A General Approach for Integrating Expert Opinions and Structured Data	Séminaire Parisien de Statistique, Institut Henri Poincare, Paris	April 7, 2014
ENBIS Management Day Round Table Discussion	ENBIS 2011, Coimbra, Portugal	September 7, 2011

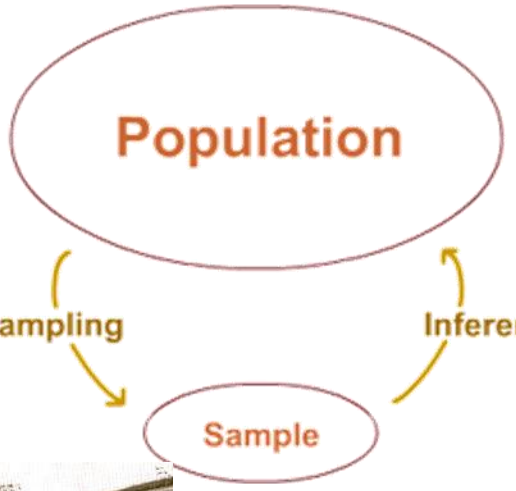
Get Help With:

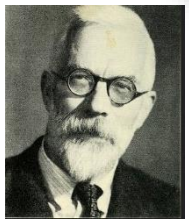
Adobe PDF and Acrobat Reader

* These links will open a new window

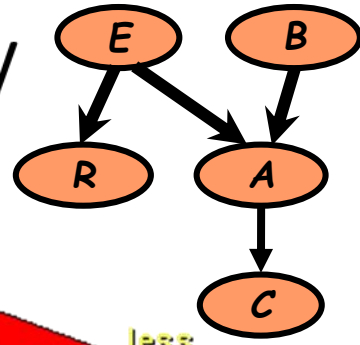
Generalizability

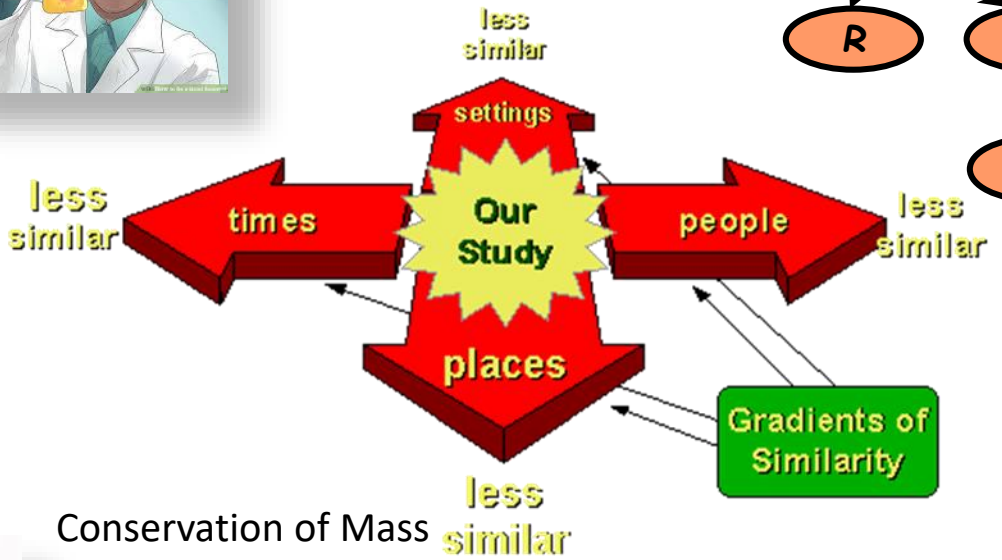
Statistical generalizability





Scientific generalizability





Conservation of Mass
Conservation of Energy
Conservation of Momentum
Newton Laws
PK/PD
Laws of thermodynamics
Maxwell's equations

Intuition

KPA
Insights through analytics

Generalizability

DE GRUYTER

Causal, Casual and Curious

Judea Pearl*

Generalizing Experimental Findings

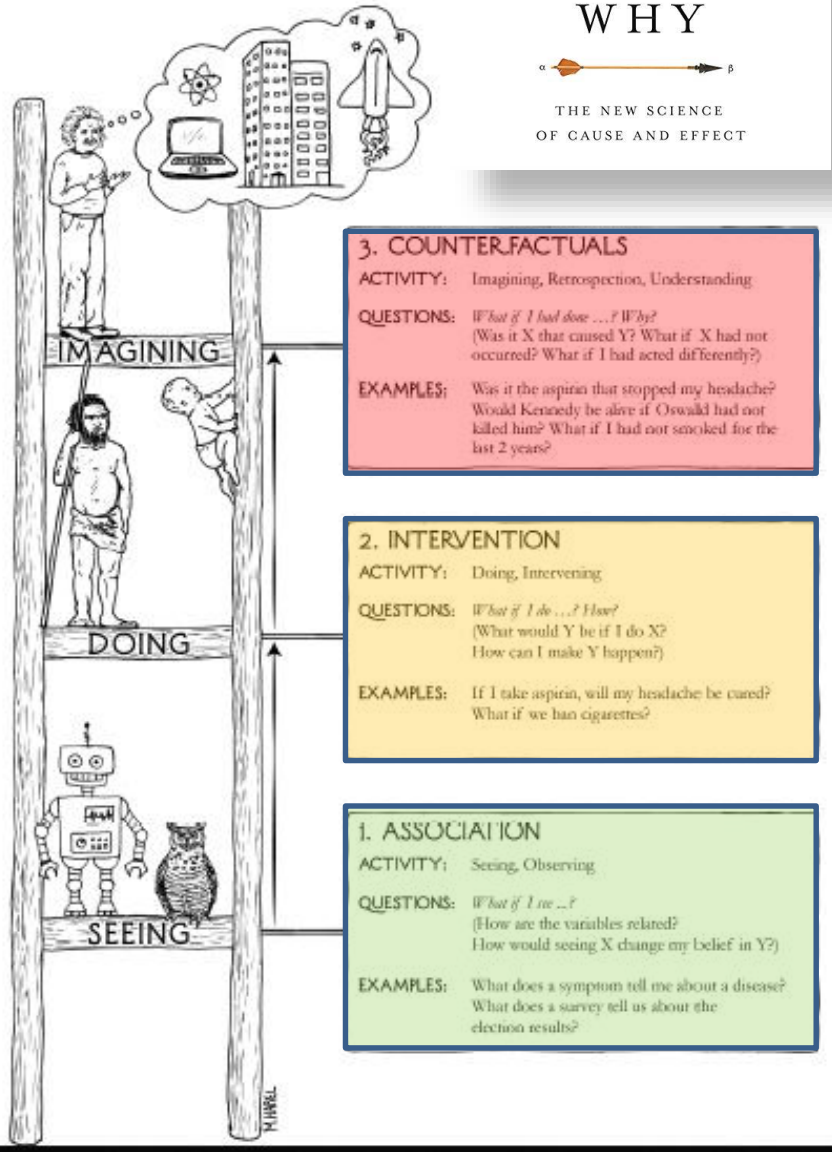
DOI 10.1515/jci-2015-0025

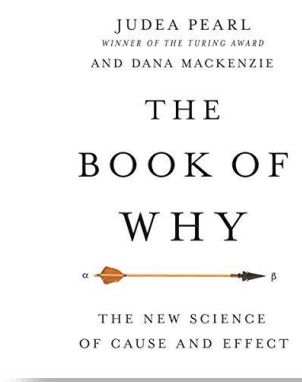
Abstract: This note examines one of the most crucial questions in causal inf
randomized clinical trials?" The question has received a formal treatment r
setting, and has led to a simple and general solution. I will describe t
ramifications, and compare it to the way researchers have attempted to
language of ignorability. We will see that ignorability-type assumptions nee
assumptions in order to capture the full spectrum of conditions that permit
judge their plausibility in specific applications.

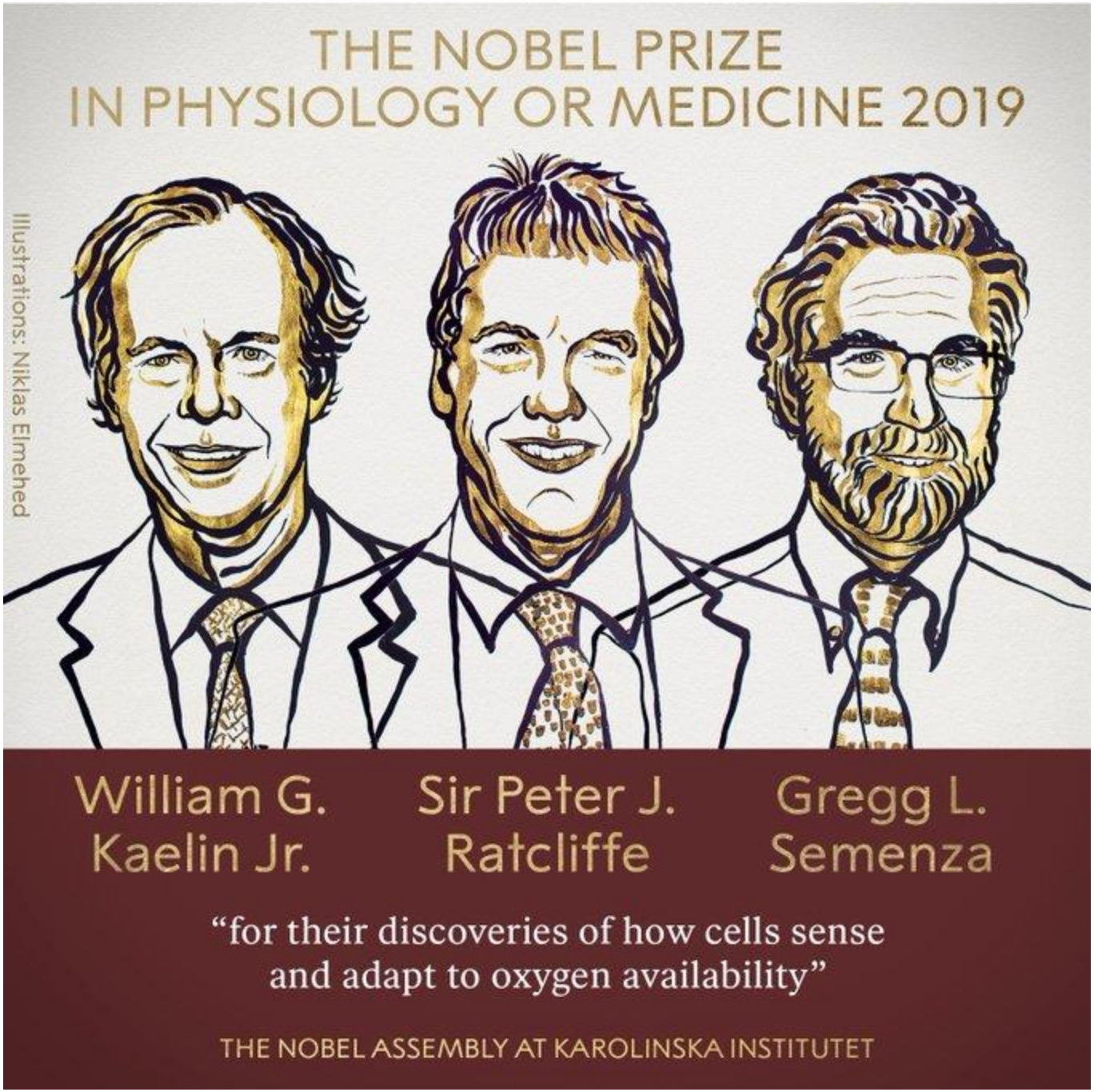
Keywords: generalizability, transportability, selection bias, admissibility, ig

1 Transportability and selection bias

The long-standing problem of generalizing experimental findings from the t
a whole, also known as the problem of "sample selection-bias" [1, 2], has re
past decade, as more researchers come to recognize this bias as a major
mental findings in both the health sciences [3] and social policy makin
randomized trial cannot be mandated, we cannot guarantee that the study
as the population of interest. For example, the study population may consis
financial and medical incentives offered by pharmaceutical firms or experi
tion of outcomes in the study may differ substantially from the distribution
interest.







The Journal of Physiology

PHD2 inactivation in Type I cells drives HIF-2α-dependent multilineage hyperplasia and the formation of paraganglioma-like carotid bodies

James W. Fielding[1,2,*], Emma J. Hodson[1,*], Xiaotong Cheng[1,2], David J. P. Ferguson[3], Luise Eckardt[1], Julie Adam[1,2], Philomena Lip[1], Matthew Maton-Howarth[1], Indrika Ratnayaka[2], Christopher W. Pugh[1], Keith J. Buckler[4], Peter J. Ratcliffe[1,2,5] and Tammie Bishop[1]

[1]*Target Discovery Institute, University of Oxford, Oxford, UK*
[2]*Ludwig Institute for Cancer Research, University of Oxford, Oxford, UK*
[3]*John Radcliffe Hospital, University of Oxford, Oxford, UK*
[4]*Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK*
[5]*The Francis Crick Institute, London, UK*

Edited by: Harold Schultz & Benedito Machado

Key points

- The carotid body is a peripheral arterial chemoreceptor that regulates ventilation in response to both acute and sustained hypoxia.
- Type I cells in this organ respond to low oxygen both acutely by depolarization and dense core vesicle secretion and, over the longer term, via cellular proliferation and enhanced ventilatory responses.
- Using lineage analysis, the present study shows that the Type I cell lineage itself proliferates and expands in response to sustained hypoxia.
- Inactivation of HIF-2α in Type I cells impairs the ventilatory, proliferative and cell intrinsic (dense core vesicle) responses to hypoxia.
- Inactivation of PHD2 in Type I cells induces multilineage hyperplasia and ultrastructural changes in dense core vesicles to form paraganglioma-like carotid bodies.
- These changes, similar to those observed in hypoxia, are dependent on HIF-2α.
- Taken together, these findings demonstrate a key role for the PHD2–HIF-2α couple in Type I cells with respect to the oxygen sensing functions of the carotid body.

Abstract The carotid body is a peripheral chemoreceptor that plays a central role in mammalian oxygen homeostasis. In response to sustained hypoxia, it manifests a rapid cellular proliferation

The statistical analysis section states: “Data are shown as the mean ± SEM. Statistical analyses were performed using unpaired Student’s t tests. For repeated measures, data were analysed by ANOVA followed by Tukey’s multiple comparison test or t test with Holm–Sidak correction for multiple comparisons as appropriate and as described in Hodson et al. (2016). $P < 0.05$ was considered statistically significant.”

In communicating their findings, they list Key Points. The first three being:

- The carotid body is a peripheral arterial chemoreceptor that regulates ventilation in response to both acute and sustained hypoxia.
- Type I cells in this organ respond to low oxygen both acutely by depolarization and dense core vesicle secretion and, over the longer term, via cellular proliferation and enhanced ventilatory responses.
- **Using lineage analysis, the present study shows that the Type I cell lineage itself proliferates and expands in response to sustained hypoxia.**

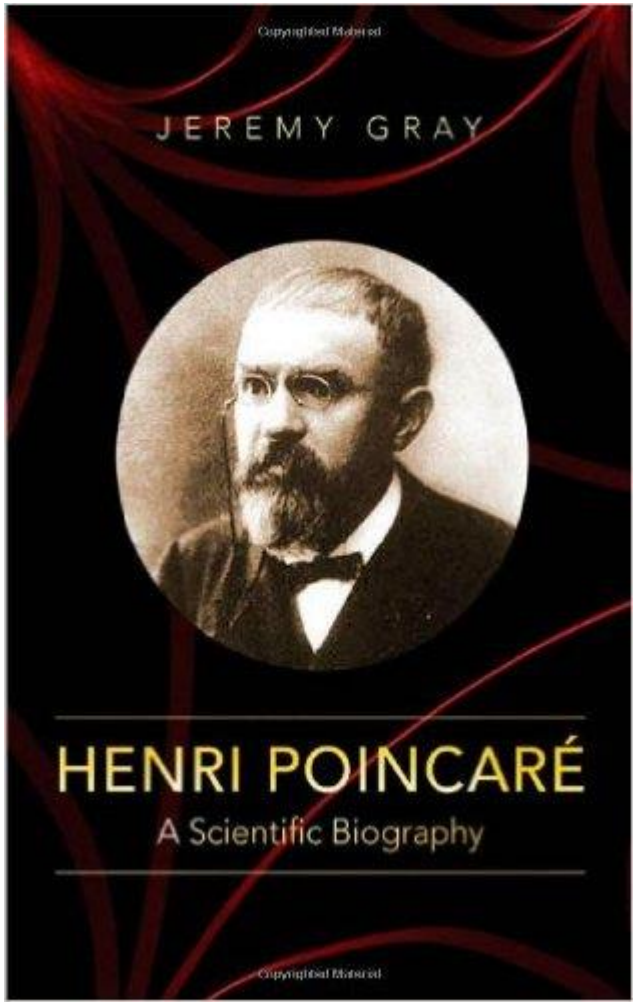
These are statements supported by statistical analysis but formulated in plain language so that they can be communicated.

Two questions come to mind in reviewing this list:

Question 1. What did they not find?

Question 2. What is the probability that they got it wrong? For example, that the Type I cell lineage itself *shrinks* in response to sustained hypoxia.

https://psyarxiv.com/jqw35



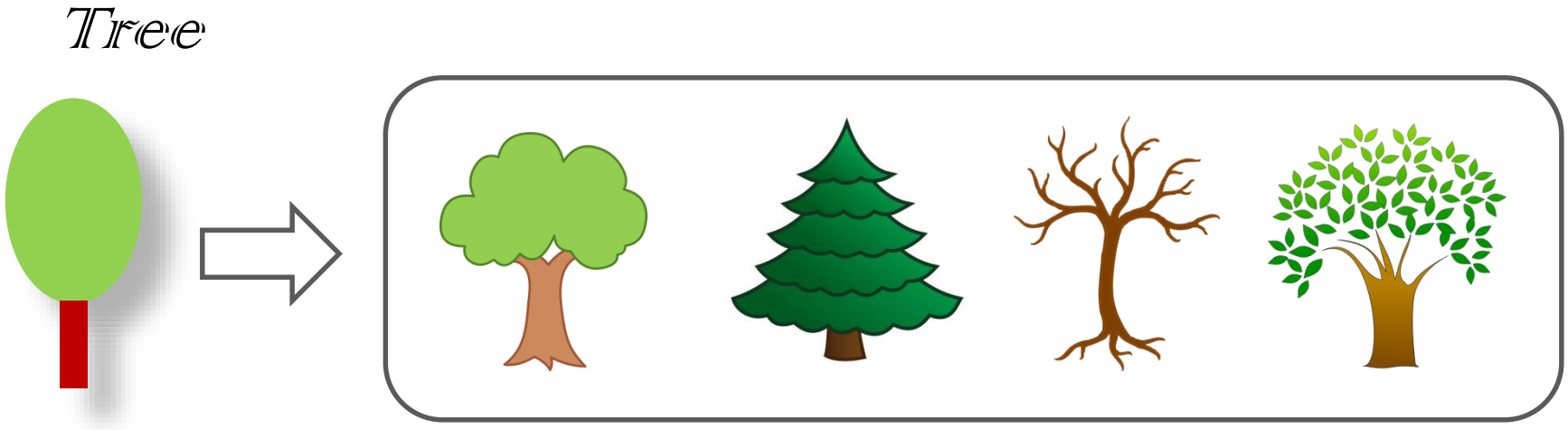
Princeton University Press, 2012

*"What he emphasized above all was the act of human **understanding**. His preferred means of attaining the understanding of a problem was to find the right **generalization** of its **core concepts**, often in the form of an **analogy**."*

J. Gray, preface to Henri Poincarre, a scientific biography

"A **concept** is an abstraction or generalization from experience or the result of a transformation of existing concepts."

Wikipedia

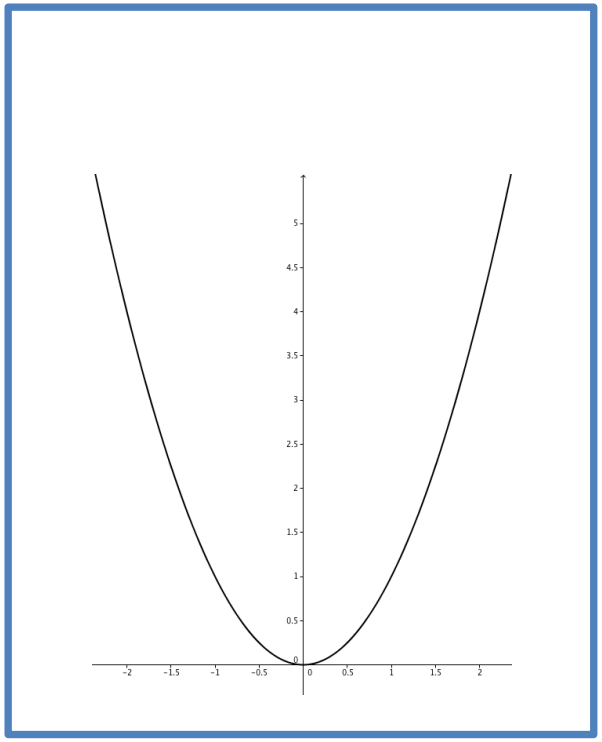


A concept can be represented in alternative forms

How do we communicate research outcomes?

KPA
Insights through analytics

Alternative representations with Meaning Equivalence



Q2

$Y = X^2$

Q2: Looks different but carries same meaning

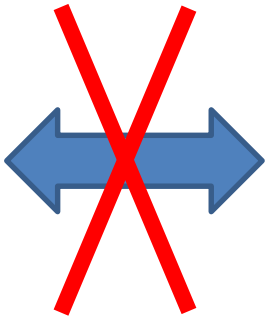
KPA
Insights through analytics

Alternative representations with Surface Similarity

Q3

$$y = \frac{k}{x^2}$$

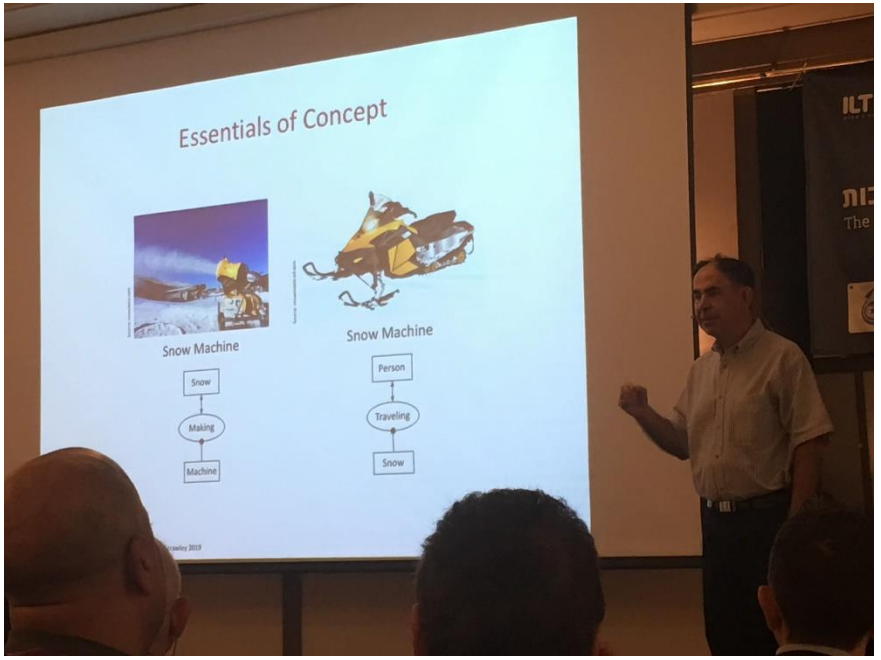
k constant



$$y = \frac{k}{x}$$

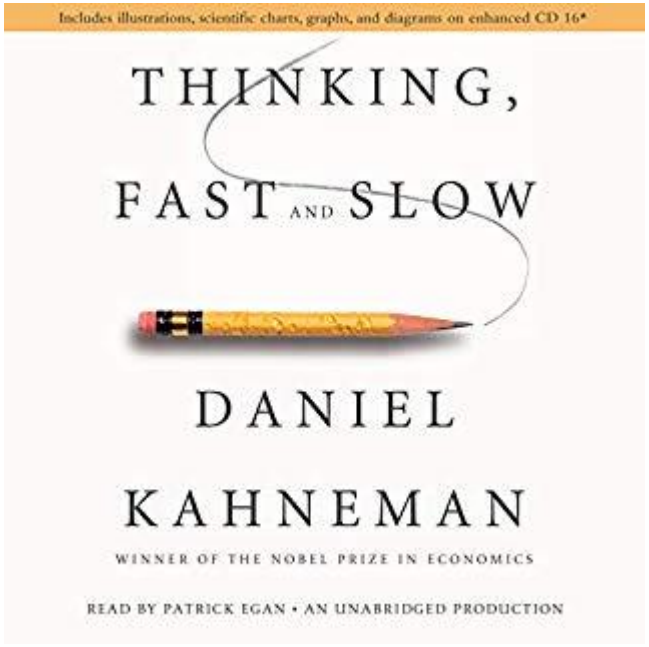
k constant

Q3: Looks similar but carries a different meaning

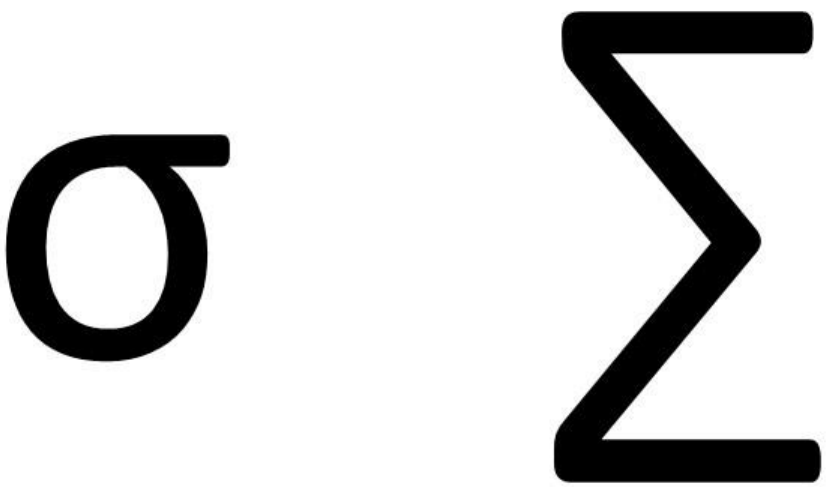


2+2

27+15



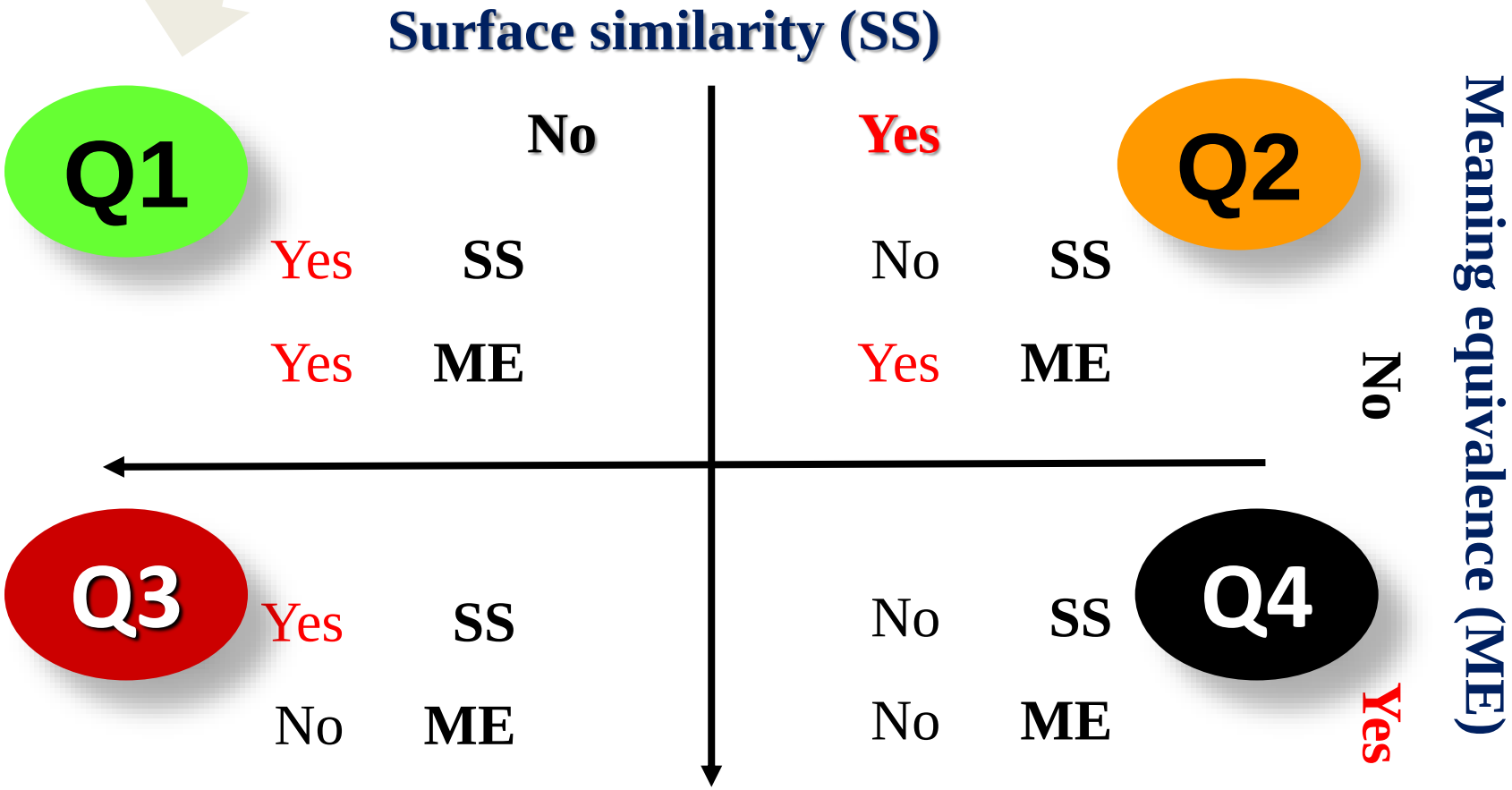
KPA
Insights through analytics



https://www.linkedin.com/pulse/little-sigma-big-sounds-same-has-totally-different-meaning-kenett/

Research findings

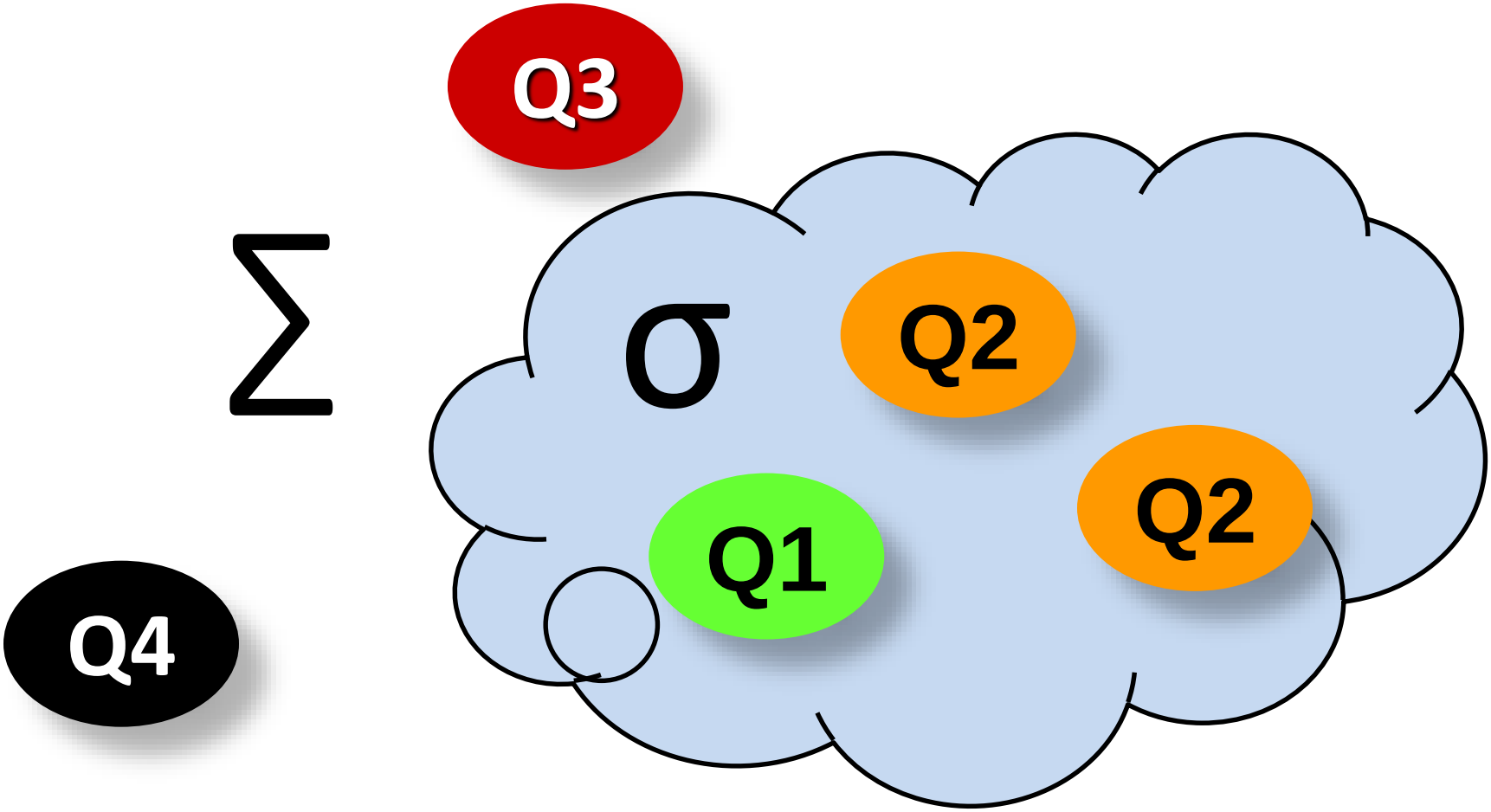
Generalize with Alternative Representations



Shafrir, U. and Kenett, R.S. (2015), Concept Science Evidence-Based MERLO Learning Analytics, in Handbook of Applied Learning Theory and Design in Modern Education, IGI Global

KPA Insights through analytics

Boundary of Meaning (BOM)



An example

THE JOURNAL OF PEDIATRICS • www.jpeds.com

ORIGINAL ARTICLES

A Structured Gradual Exposure Protocol to Baked and Heated Milk in the Treatment of Milk Allergy

Adi Efron, MD[1], Yuri Zeldin, MD[2,3], Leora Gotesdyner, MSc[1], Tali Stauber, MD[1,4,5], Ramit Maoz Segal, MD[5], Inga Binson, MD[3], Mira Dinkin, MD[3], Larisa Dinkowitz, MD[3], Danit Shahar, BSN[4,5], Michal Deutch, BSN[5], Mazal Yaron, BSN[3], Ayelet Nevet, BSN[3], Avner Reshef, MD[5], Nancy Agmon-Levin, MD[1,5], Ron S. Kenett, PhD[6], and Mona I. Kidon, MD[1,3,4,5]

Objective To evaluate the efficacy and safety of a structured gradual exposure protocol (SGEP) with extensively heated and baked milk in promoting allergy resolution in children with cow milk allergy (CMA).

Study design In a case control study, children with CMA aged 1-4 years who were treated with SGEP including extensively heated and baked milk, were compared with children treated with strict avoidance. Data were collected from medical records and from validated telephone questionnaires. Data analysis was performed using a nonparametric Kaplan-Meier and proportional hazard Cox regression model, after evaluation of the adequacy of the case control matching.

Results There were 43 children with milk allergy—26 (62%) males with a mean age at intervention of 21 months (range, 12-47 months)—who were treated with SGEP and followed to a mean age of 40 months (range, 20-82 months). The median age at resolution of CMA was compared with a matched group of 67 children treated with strict avoidance at least until 4 years of age or followed until earlier resolution, with a mean age at follow-up of 71 months (range, 11-176 months). The median estimated age at CMA resolution in the SGEP group was 36 months (95% CI, 34.5-49.7) compared with 98 months (95% CI, 82.4-114.1) in controls ($P < .001$). At last follow-up, 86% of treated children were tolerant to unheated milk proteins vs 52% of controls ($P = .003$).

Conclusion A structured protocol with extensively heated and baked milk seems to promote faster resolution of CMA. *(J Pediatr 2018;■■:■■-■■).*

“The quality of life of patients and families affected with a food allergy to staple foods (milk, egg, sesame, peanut) is impaired”
is **equivalent in meaning** to: “Food allergy in children impacts negatively on day to day activities of the whole family “

“Food allergy in children impacts negatively on day to day activities of the whole family “ has **surface similarity** to: “Educating patients on strict avoidance and carrying an epinephrine autoinjector, is completely effective in avoiding accidental exposures in preschool children”.

Table III. Boundary of meaning statements **BOM**

Target statement	Meaning equivalence findings included in BOM	Surface similarity findings not included in BOM
Finding 1: The quality of life of patients and families affected with a food allergy to staple foods (milk, egg, sesame, peanut) is impaired	Food allergy in children impacts negatively on the day-to-day activities of the whole family	Educating patients on strict avoidance and carrying an epinephrine autoinjector is completely effective in avoiding accidental exposures in preschool children
	The incidence of accidental exposures to allergenic foods in preschool children is high	
	The currently recommended management of food allergy in children is patient education, strict avoidance, and carrying an epinephrine autoinjector	
Finding 2: All children suspected of an allergic reaction to foods should be referred to a center that includes appropriate facilities, medical, and support staff experienced in the diagnosis and treatment of children with food allergies as early as possible	The diagnosis of food allergy in children should be performed soon after the suspected event	Recommending strict avoidance of suspected allergenic foods is the best treatment for all young food allergic children
	There are no age limitations on the performance of diagnostic allergy tests, such as SPTs or observed food challenges, provided these are performed by well trained and experienced medical teams	Laboratory test such as sIgE to food can accurately diagnose food allergy in children
Finding 3: The natural history of CM allergy in children is still favorable as in most—it seems to resolve with time	The median age at resolution of CMA (by which time 50% of children have resolved their allergies) is between 6 and 8 years	Food allergy in children resolves in the first years of life
	Children with CMA and a positive family history of atopy, an initial anaphylactic reaction, recurrent wheezing or moderate/severe atopic dermatitis are less likely to resolve their CMA	Avoidance of allergenic foods is beneficial in preventing food allergy in children
Finding 4: A majority of children with IgE mediated CMA are capable of consuming certain amounts of EHBM proteins	Some children with CMA can develop immediate, life-threatening reactions to the ingestion of EHBM	Families of children with IgE-mediated CMA should be encouraged to try baked milk at home
	A minority of children with CMA are allergic also to heat denatured milk products. These are the most severely affected and least likely to resolve their allergies	All forms of heated and baked milk are similarly safe
Finding 5: In preschool children with CMA capable of ingesting EHBM safely, SGEP seems to promote earlier resolution	The median age at CMA resolution of preschool children, capable of ingesting EHBM safely and treated with SGEP including EHBM, seems to be significantly lower than in children treated with avoidance	Preschool children capable of ingesting EHBM safely and treated with SGEP including EHBM are developing true long-term tolerance to milk
	Most preschool children capable of ingesting EHBM safely and treated with SGEP including EHBM will be able to tolerate milk in their regular diet before entering school	EHBM is not a form of oral immunotherapy in food allergic children and therefore the follow- up recommended for these children is similar to patients with natural resolution of CMA (none)
Finding 6: A protocol of SGEP including EHBM, seems safe in children <4 years of age	A protocol of SGEP, including EHBM, performed by medical teams trained and experienced in the treatment of food allergy in children is safe	All children with IgE-mediated CMA should be treated with an SGEP with EHBM

Special issue: Research article

polymers advanced technologies

Received: 13 February 2015, Accepted: 22 March 2015, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/pat.3531

A multifactorial analysis of complex pharmaceutical platforms: an application of design of experiments to targetable polyacrylamide and ultrasound contrast agents

Meital Bloch[a], Ron Kenett[a]*, Lauren Jablonowski[b], Margaret Wheatley[b], Eylon Yavin[a] and Abraham Rubinstein[a]*

To improve visualizati
cently suggested a m
near infrared dye deri
to the recognition pep
jugate (Flu-CPAA-Pep)
tect it from pre-matur
directed ultrasound in
the MBs rupture into s
vasculature and allow

Contents lists available at ScienceDirect

Journal of Controlled Release

journal homepage: www.elsevier.com/locate/jconrel

ELSEVIER

The effect of linker type and recognition peptide conjugation chemistry on tissue affinity and cytotoxicity of charged polyacrylamide

Meital B.D. Bloch [a], Eylon Yavin [a], Aviram Nissan [b], Ilana Ariel [c], Ron Kenett [a,d], Dovrat Brass [e], Abraham Rubinstein [a,*]

Another example

KPA Insights through analytics

The medical problem

Colorectal cancer (CRC):

- The 3rd most common cancer diagnosed in USA.
- The 2nd leading cause of cancer-related death.

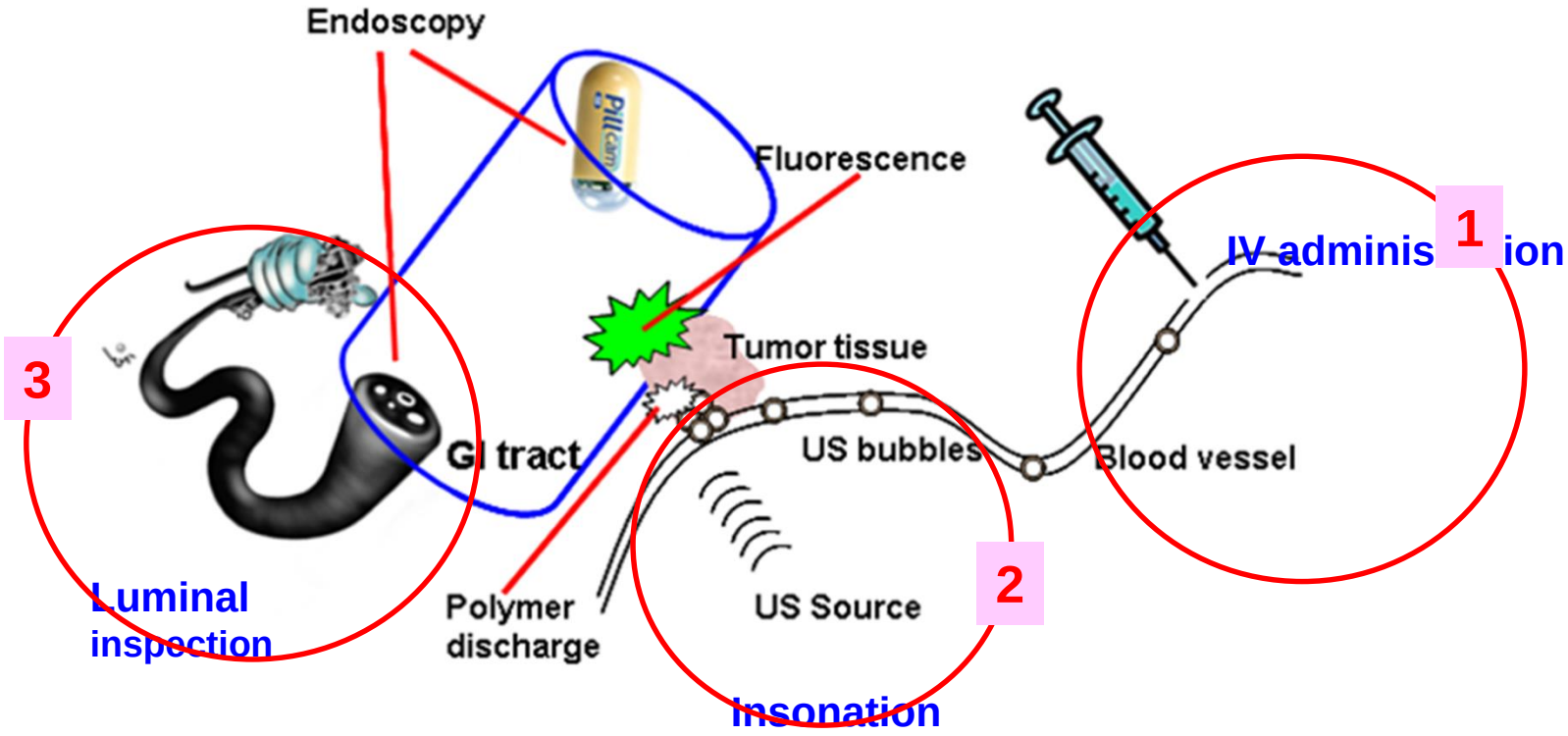
CRC treatment:

- Surgery
- Chemo/radio adjuvant therapy – depending on the CRC stage

- Overall incidence of CRC decline due to an advance in:
  - early diagnosis
  - improved medical treatments.
- This decline could even accelerate if efficient screening system is available.

Rex, D.K., *et al*. *Gastroenterology*, 112: 24, 1997.
Levin, B., *et al. Gastroenterology*, 134: 1570, 2008.
Mayer R.J. *et al*. *N. Engl. J Med,* 352: 476, 2005.
Vogelstein B. *et al*. *N. Engl. J Med*, 319: 525, 1998.
Edwards BK. *et al. Cancer, 116*: 544, 2010.

The concept



Hypotheses:

1. Targetability of Flu-CPAA towards dysplastic colon tissues is improved by adding a recognition peptide (Flu-CPAA-Pep).
2. Microbubbles protect Flu-CPAA and Flu-CPAA-Pep from premature affinity in the blood stream.

Bloch M, *et al. Int. J. Pharm*, 478: 504, 2015.

Power of the *in vitro* studies

Power Analysis

Significance Level	0.05
Anticipated RMSE	1

Term	Anticipated Coefficient	Power
Intercept	1	1
Mol% cat	1	1
Peptide	1	1
Presenting platform 1	1	0.988
Presenting platform 2	-1	0.917
Metastatic stage	1	0.993
Mol% cat*Peptide	1	1
Mol% cat*Presenting platform 1	-1	0.988
Mol% cat*Presenting platform 2	1	0.917
Mol% cat*Metastatic stage	-1	0.993
Peptide*Presenting platform 1	1	0.988
Peptide*Presenting platform 2	-1	0.917
Peptide*Metastatic stage	1	0.993
Presenting platform*Metastatic stage 1	-1	0.899
Presenting platform*Metastatic stage 2	1	0.84

Effect	Power
Presenting platform	0.974
Mol% cat*Presenting platform	0.974
Peptide*Presenting platform	0.974
Presenting platform*Metastatic stage	0.883

Power of the *in vivo* studies

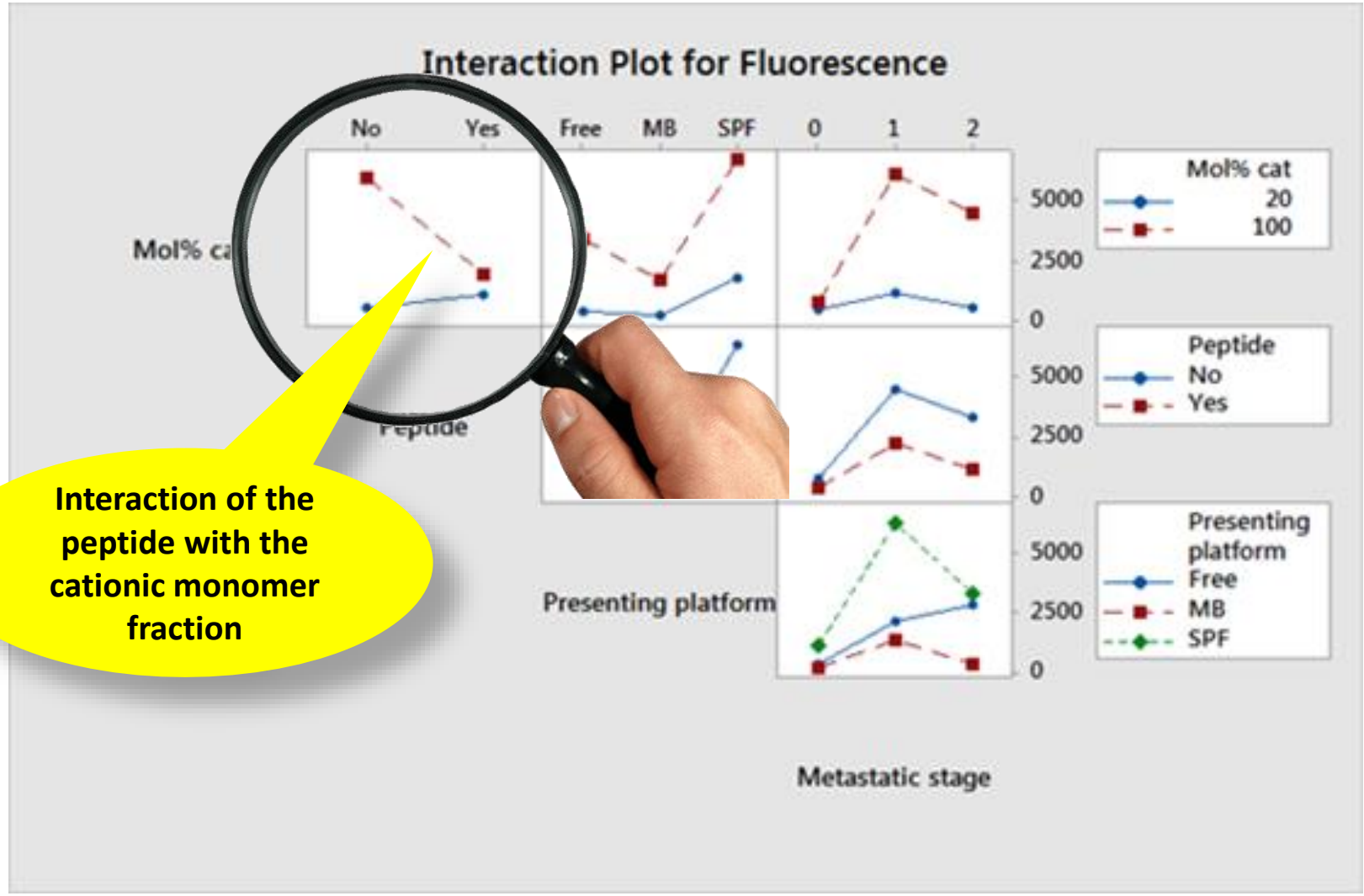
Design Evaluation

Power Analysis

Significance Level	0.05
Anticipated RMSE	1

Term	Anticipated Coefficient	Power
Intercept	1	0.864
Peptide	1	0.864
Mode of administration	1	0.877
SPF	1	0.864
Peptide*Mode of administration	1	0.877
Peptide*SPF	-1	0.864
Mode of administration*SPF	1	0.877

Interaction plot for the *in vitro* studies



Bloch M., *et al*., *Pol. Adv. Tech.* , 26: 898, 2015

KPA
Insights through analytics

Main Findings

1. **Increasing the charge density of Flu- CPAA-Pep leads to cross-reaction with the recognition peptide, VRPMPLQ .**

2. **Apart of Flu- CPAA-100, incorporation of the polymers into MBs did not significantly affect the MBs echogenic properties.**

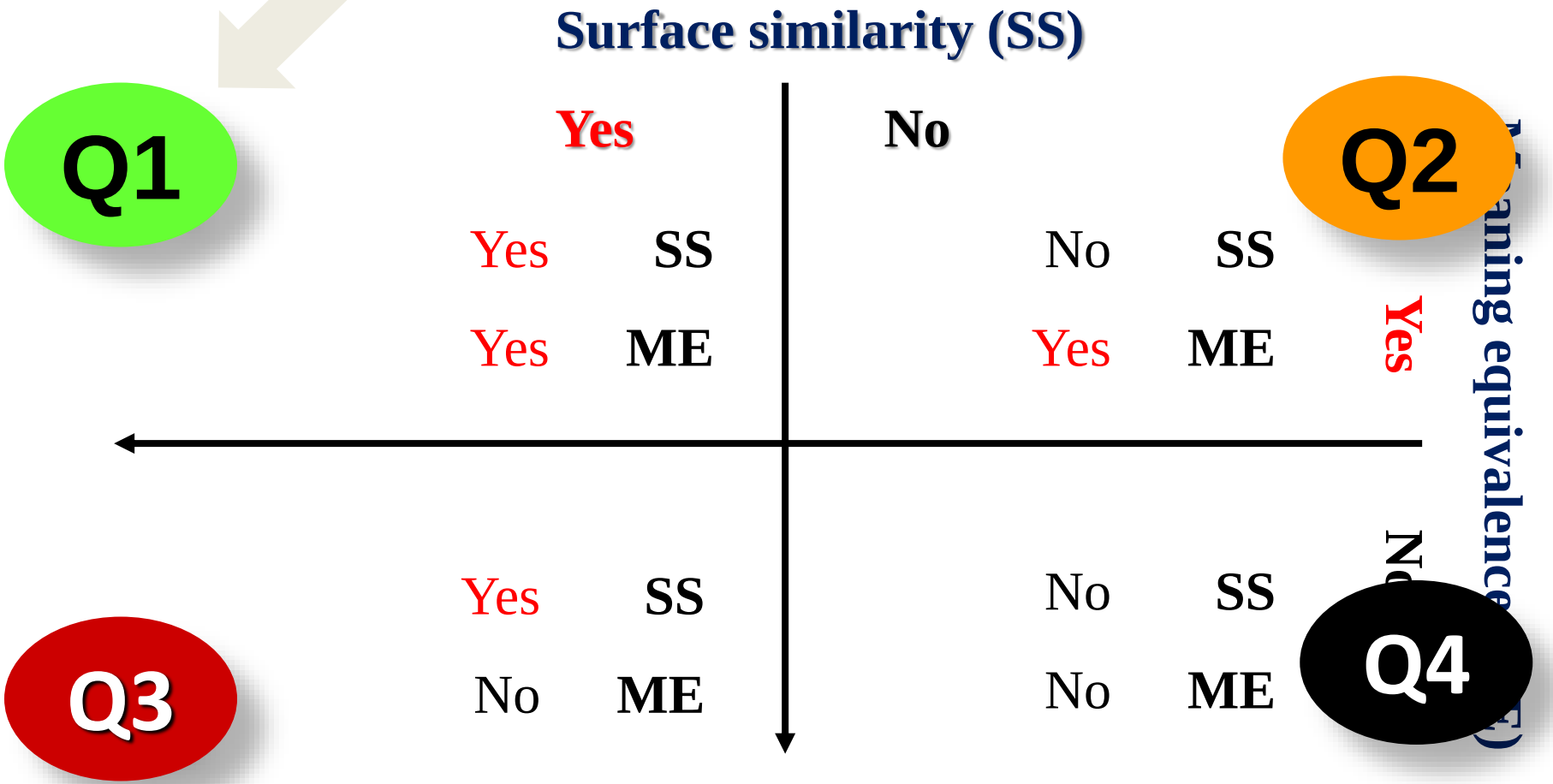
3. **Flu-CPAA-Pep binds to dysplasia regions, after both IV and rectal administrations in the rat model.**

4. **Fragmenting MBs into SPF does not interfere with the affinity of Flu-CPAA and Flu-CPAA-Pep to malignant colon tissues after IV or rectal administrations in the rat.**

5. **SPF protected their Flu-CPAA-Pep cargo from non-specific interaction with serum proteins.**

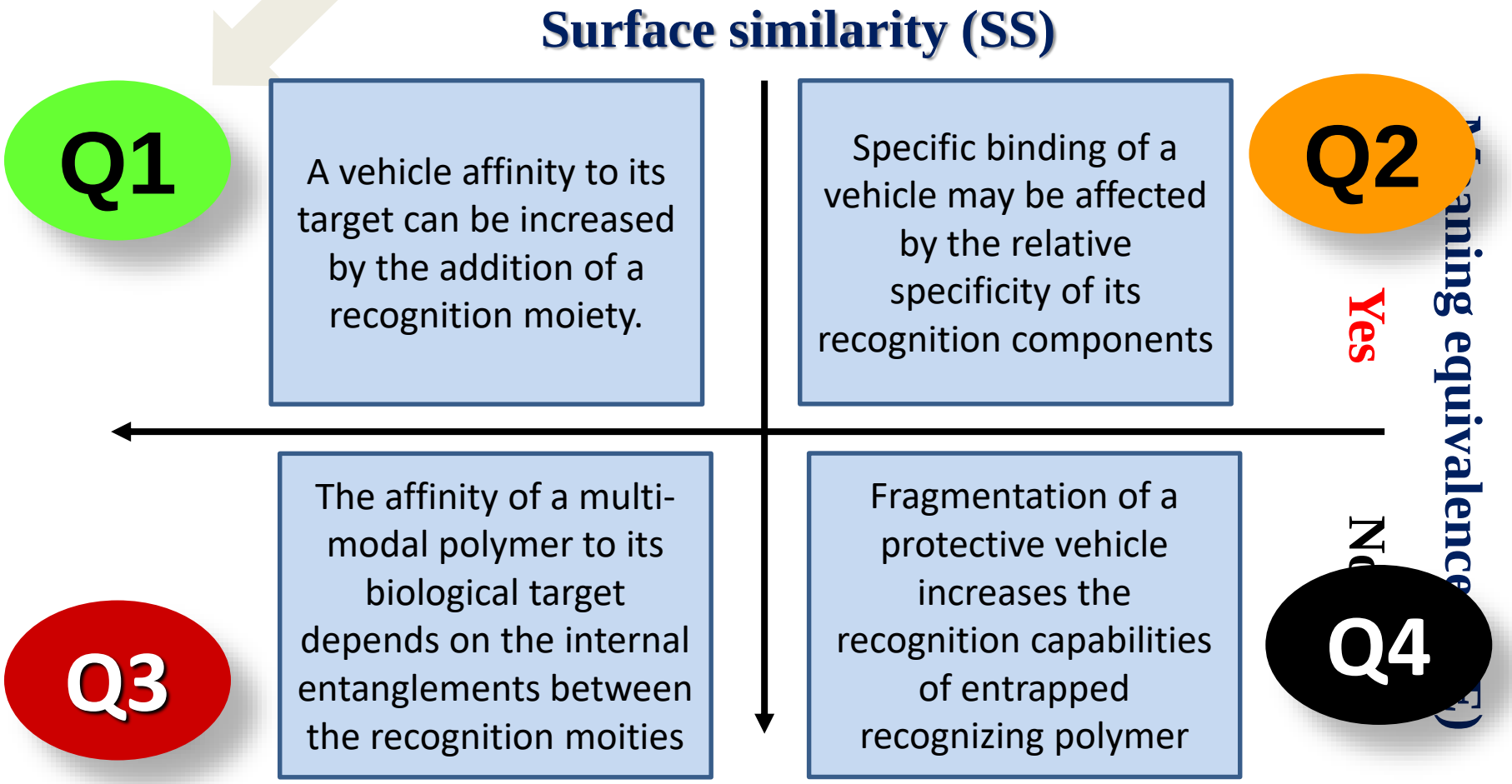
Bloch M., *et al*., *Pol. Adv. Tech.* , 26: 898, 2015

Increasing the charge density leads to cross-reaction with the recognition peptide



KPA
Insights through analytics

Increasing the charge density leads to cross-reaction with the recognition peptide



The boundary of meaning (BOM)

Boundary of meaning

	Phrased Finding	Meaning Equivalence of the Finding (MEF) [1]	Surface Similarity Finding (SSF) [2]
1	The addition of VRPMPLQ to the Flu-CPAA backbone increased the specific binding of the polymer to their biological target.	MEF1-1: A vehicle affinity to its target can be increased by the addition of a recognition moiety.	SSF1-1: The affinity of a multi-modal polymer to it biological target depends on the internal entanglements between recognition moieties.
		MEF1-2: Specific binding of a vehicle may be affected by the relative specificity of its recognition compon	SSF1-2: When one recognition moiety depends on its charge, the higher the charge density the higher the affinity obtained.
2	Loading the Flu-CPAA into MBs, significantly reduced the ability of the Flu-CPAA polymers to interact with their biological targets.	MEF2-1: Loading a targeted polymer into a protective vehicle interferes with the affinity properties of the polymer.	SSF2-1: Recognition polymers express reduced affinity to their biological targets when loaded into a degradable vehicle.
		MEF2-2: Recognition of a biological target by a targetable polymer depends on the free acquaintance of the recognition moieties.	SSF2-2: Recognition polymer mode of loading into a protective vehicle affects the affinity to the biological target.
3	Fragmenting the MBs into SPF restored the recognition properties of the Flu-CPAA polymers and even increased them.	MEF3-1: Rupturing the barrier functions of a protective vehicle regenerates the recognition properties of its polymeric cargo.	SSF3-1: Targeted nanoparticles enhance their recognition properties towards biological targets after fragmentation.
		MEF3-2: Unveiling a shield from a support carrier restores the properties of the cargo polymer.	SSF3-2: Fragmentation of a protective vehicle increases the recognition capabilities of entrapped recognizing polymer.

Q1

Q2

Q2

Q3

Q3

Type S (sign) errors

"Contrary to the common impression, retrospective design calculation may be more relevant for statistically significant findings than for nonsignificant findings: The interpretation of a statistically significant result can change drastically depending on the plausible size of the underlying effect.
Like power analysis, the design calculations we recommend require external estimates of effect sizes or population differences."

Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors

Perspectives on Psychological Science
2014, Vol. 9(6) 641–651

DOI: 10.1177/1745691614551642
pps.sagepub.com
SAGE

Andrew Gelman[1] and John Carlin[2,3]
[1]Department of Statistics and Department of Political Science, Columbia University; [2]Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Parkville, Victoria, Australia; and [3]Department of Paediatrics and School of Population and Global Health, University of Melbourne

Type S error: $\theta 1 > \theta 2$, but I claim that $\theta 1 < \theta 2$ (or vice versa)

From external information…

D : the true effect size

From the data (or model if prospective design)…

d : the observed effect

s : SE of the observed effect

p : the resulting p-value

Hypothetical replicated data

d^{rep} : the effect that would be observed in a hypothetical replication study with a design like the one used in the original study (so assumed also to have SE = s)

Use a value (or set of values) of the treatment effect considered plausible in advance of doing the study. Condition on a result being significant to calculate the Bayesian posterior probability of its being of the correct sign (S)

Design calculations:

- *Power*: the probability that the replication d^{rep} is larger (in absolute value) than the critical value that is considered to define "statistical significance" in this analysis.
- *Type S error rate*: the probability that the replicated estimate has the incorrect sign, if it is statistically significantly different from zero.

Type S error: $\theta 1 > \theta 2$, but I claim that $\theta 1 < \theta 2$ (or vice versa)

KPA
Insights through analytics

Testing a BOM

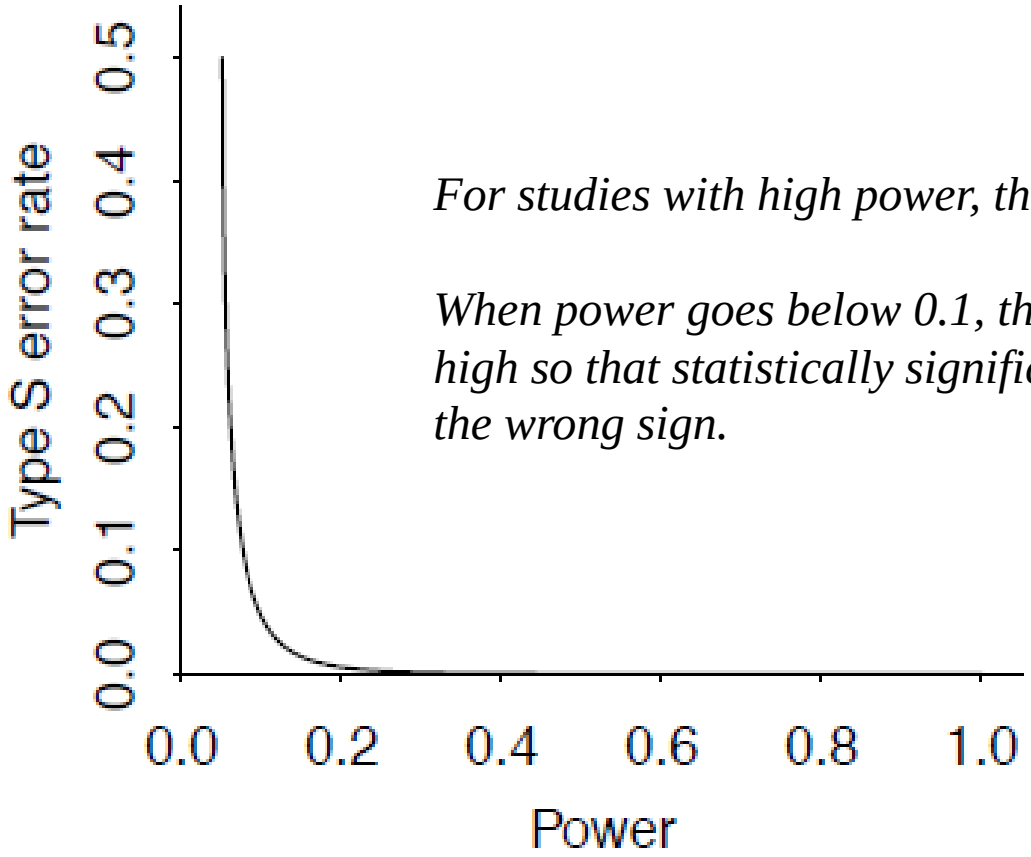
A **Type I error** consists of rejecting the "null hypothesis" (roughly speaking, the assumption of no effect, the hypothesis you typically set out to disprove) in favour of the "alternative hypothesis" when in fact the null hypothesis is true.

A **Type II error** consists of accepting the null hypothesis (technically, failing to reject the null hypothesis) when in fact the null hypothesis is false.

Identify effects

Interpret significant effects

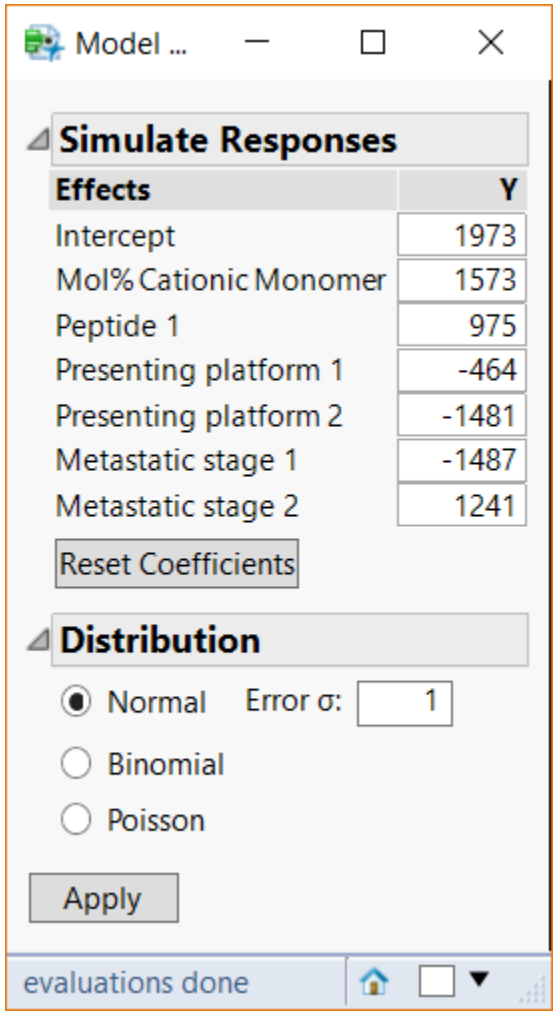
Type S (sign) errors

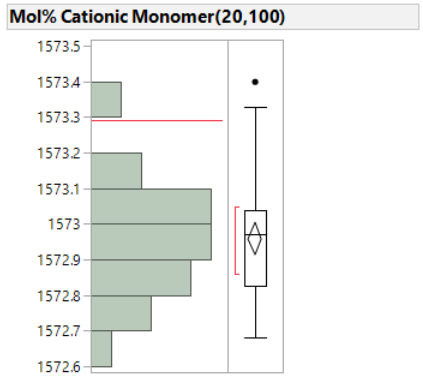


For studies with high power, the Type S error rate is low.

When power goes below 0.1, the Type S error rate becomes high so that statistically significant estimates are likely to be the wrong sign.

Type S (sign) errors





Quantiles

100.0%	maximum	1573.3981178
99.5%		1573.3981178
97.5%		1573.3784291
90.0%		1573.165101
75.0%	quartile	1573.0367376
50.0%	median	1572.9702101
25.0%	quartile	1572.825733
10.0%		1572.7420112
2.5%		1572.6854329
0.5%		1572.6803051
0.0%	minimum	1572.6803051

Summary Statistics

Mean	1572.9584
Std Dev	0.1606457
Std Err Mean	0.0227187
Upper 95% Mean	1573.004
Lower 95% Mean	1572.9127
N	50

Simulation Results

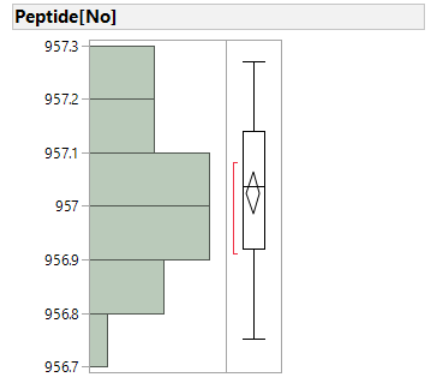
Y_0 = 1573.29 (Original Estimate)

Confidence Intervals

Alpha	Lower CI	Upper CI
0.05	1572.69	1573.38
0.10	1572.7	1573.31
0.20	1572.74	1573.17
0.50	1572.83	1573.04

Empirical p-Values

Test	p-Value
$Y \geq \lvert Y_0 \rvert$	0.0600
$Y \leq Y_0$	0.9400
$Y \geq Y_0$	0.0600



Quantiles

100.0%	maximum	957.27028523
99.5%		957.27028523
97.5%		957.26817444
90.0%		957.23832603
75.0%	quartile	957.14014185
50.0%	median	957.03812428
25.0%	quartile	956.91961203
10.0%		956.84776408
2.5%		956.76479565
0.5%		956.75239413
0.0%	minimum	956.75239413

Summary Statistics

Mean	957.02434
Std Dev	0.1355272
Std Err Mean	0.0191664
Upper 95% Mean	957.06286
Lower 95% Mean	956.98583
N	50

Simulation Results

Y_0 = 957.51 (Original Estimate)

Confidence Intervals

Alpha	Lower CI	Upper CI
0.05	956.765	957.268
0.10	956.812	957.258
0.20	956.848	957.238
0.50	956.92	957.14

Empirical p-Values

Test	p-Value
$Y \geq \lvert Y_0 \rvert$	<.0001*
$Y \leq Y_0$	1.0000
$Y \geq Y_0$	<.0001*

KPA
Insights through analytics

44

What did we cover?

- Reproducibility
- Information quality
- Generalizability
- Boundary of meaning (BOM)
- Testing a BOM

To make a reproducibility claim:

1. State your findings
2. Generalize your findings
3. Present a boundary of meaning table
4. Perform S type tests to support the BOM

Thank you for your attention