How we (mis)conceptualise p-values (and what we can do about it)

PRICEKaren
MCLEANKatrina
KHANAbeer
KNIGHTAndrew
RHEEJoel
TAMMichael
Academic Primary and Integrated Care Unit
m.tam@unsw.edu.au

preface

p

> **... surely the most bone-headedly misguided procedure ever institutionalised in the rote training of science students.**
>
> William Rozeboom

“Let me tell you why you’re here. You’re here because you know something. What you know you can’t explain, but you feel it. You’ve felt it your entire life, that there’s something wrong with the world.”

p

the choice

The scenario:

A study comparing a new antihypertensive to an older agent, with blood pressure as the primary outcome, is published in a medical journal. In the article's conclusion, the authors claim that, the new drug was superior to the old drug at lowering blood pressure (p = 0.05).

A reader makes the following interpretation:

"This means that there is a 5% probability that this result is due to chance alone, or, there is a 95% probability that the conclusion is true".

Please select the option that BEST MATCHES your understanding of p-values:

- ◯ The reader's interpretation is mostly FALSE
- ◯ The reader's interpretation is mostly TRUE

GP Unit

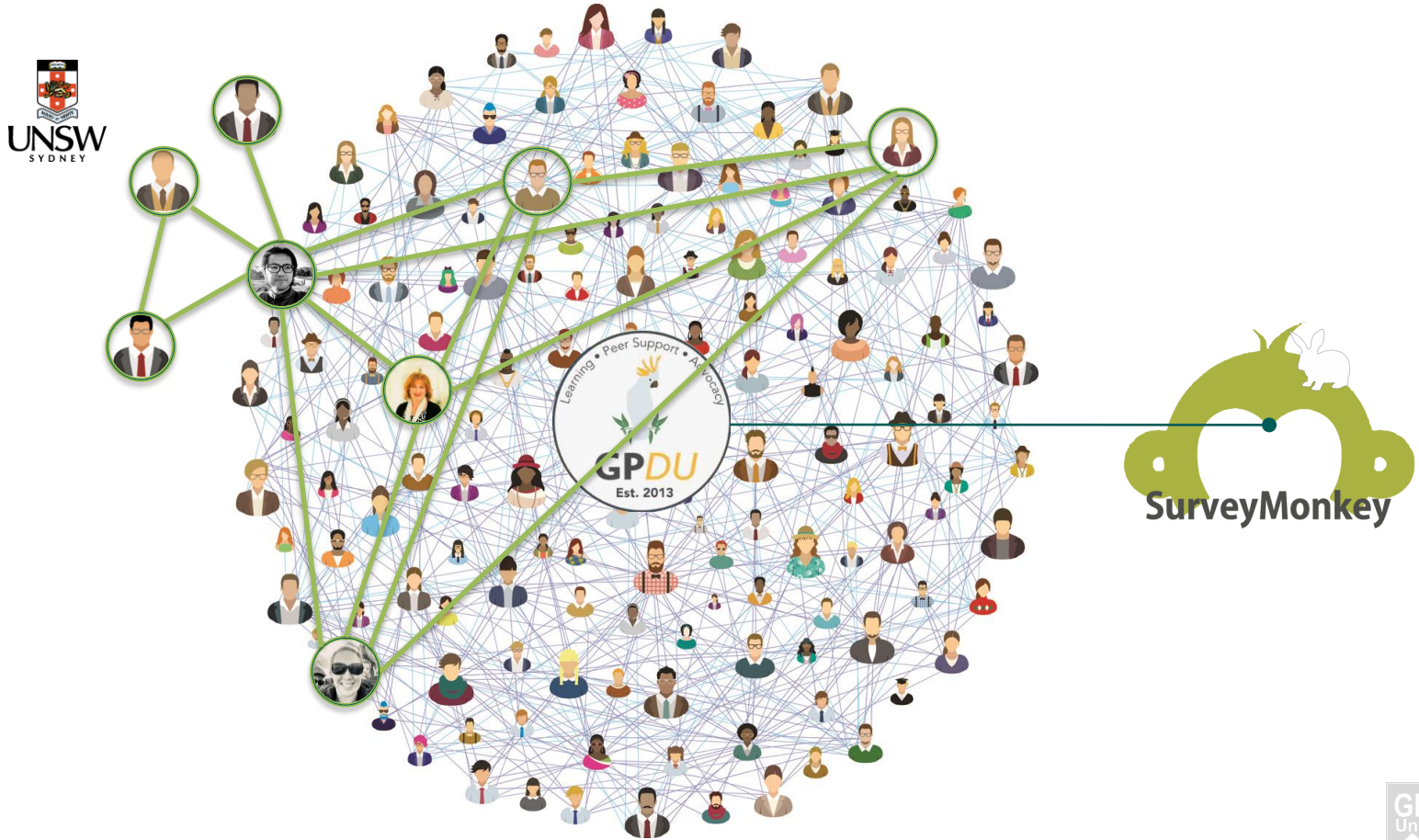
p = 0.05

This means that there is a 5% probability that this result is due to chance alone, or, there is a 95% probability that the conclusion is true.

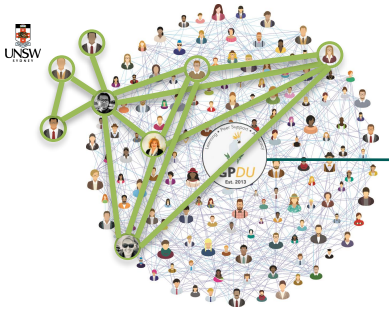
aim

describe and
categorise
how
clinicians
conceptualise
p-values

method


UNSW
SYDNEY
Learning • Peer Support • Advocacy
GPDU
Est. 2013
SurveyMonkey

method



SurveyMonkey

mixed methods

qualitative question

dichotomous choice question

confidence question

Please consider the following scenario.

The scenario:

A study comparing a new antihypertensive to an older agent, with blood pressure as the primary outcome, is published in a medical journal. In the article's conclusion, the authors claim that, the new drug was superior to the old drug at lowering blood pressure (p = 0.05).

qualitative question

In no more than two to three sentences, describe what "p = 0.05" means in the above statement.

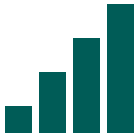
The scenario:
A study comparing a new antihypertensive to an older agent, with blood pressure as the primary outcome, is published in a medical journal. In the article's conclusion, the authors claim that, the new drug was superior to the old drug at lowering blood pressure (p = 0.05).

A reader makes the following interpretation:
"This means that there is a 5% probability that this result is due to chance alone, or, there is a 95% probability that the conclusion is true".

dichotomous choice question

Please select the option that BEST MATCHES your understanding of p-values:

- ◯ The reader's interpretation is mostly FALSE
- ◯ The reader's interpretation is mostly TRUE



confidence question

Please indicate how CONFIDENT you are of your answer:

Not at all	Slightly	Somewhat	Very	Entirely
◯	◯	◯	◯	◯

results

247

mean

25 — 40 — 67

age range

21%
registrars

mean

0 — 10 — 40

yrs in GP

55%
post-grad degree

RACGP

71%
Fellows

5%
Fellows

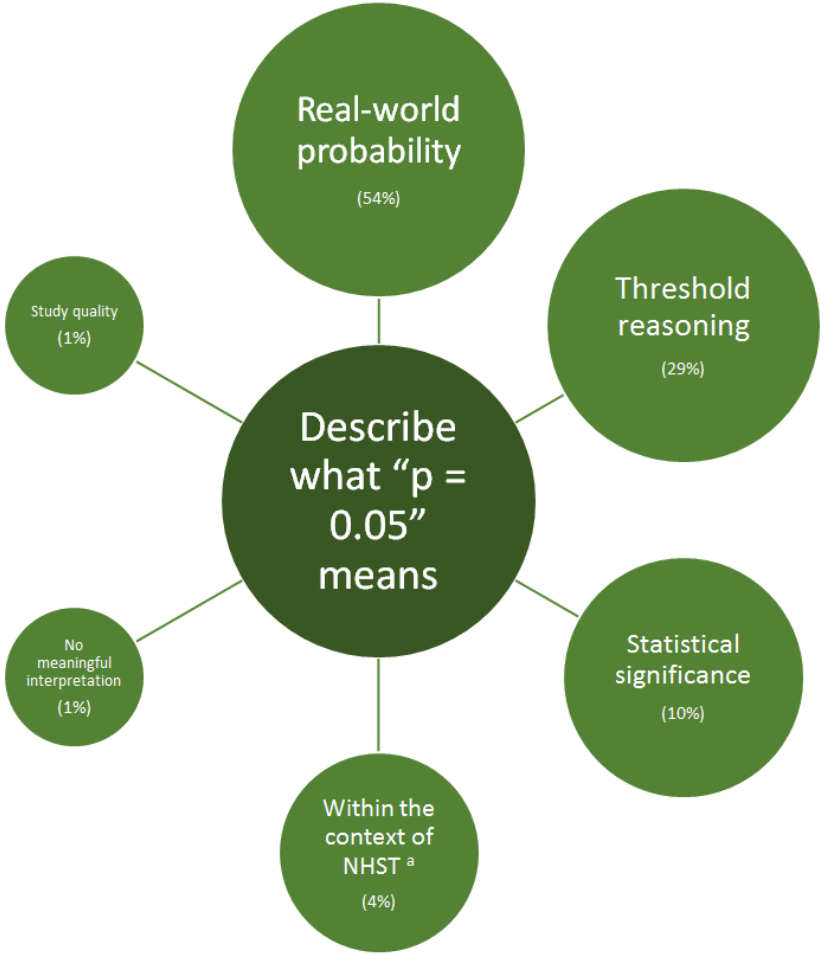
4% FRNZCGP
3% FARGP

72%
teacher

37%
researcher

82%
any research experience

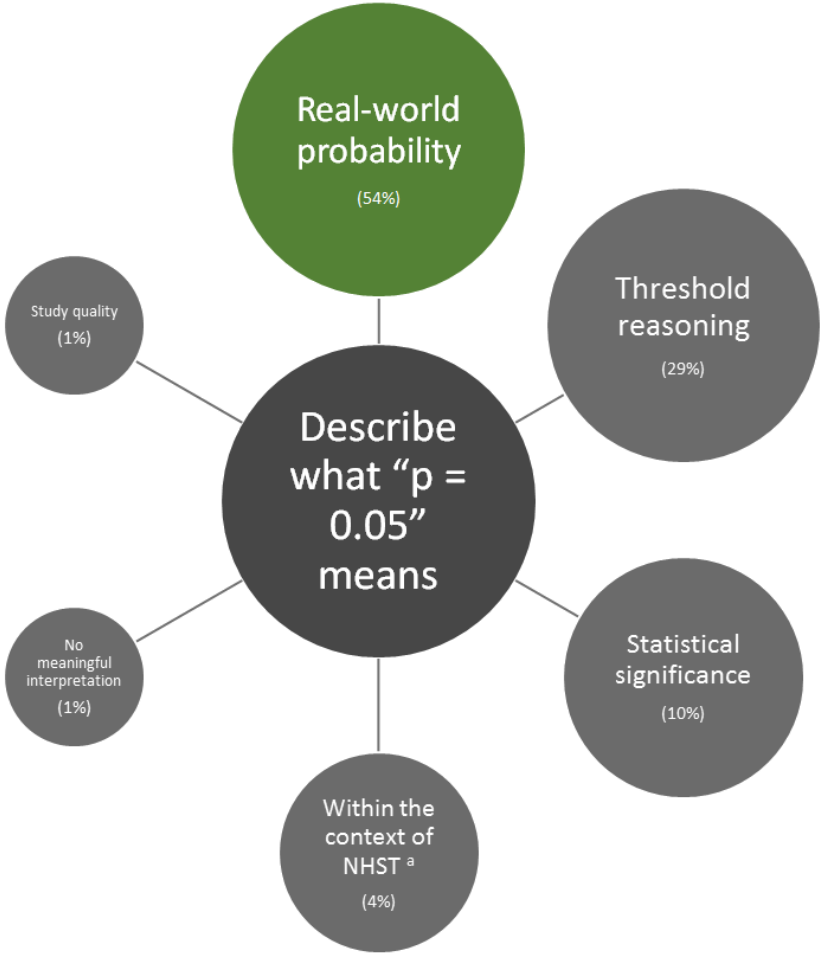
results



qualitative question

6 conceptual categories

results

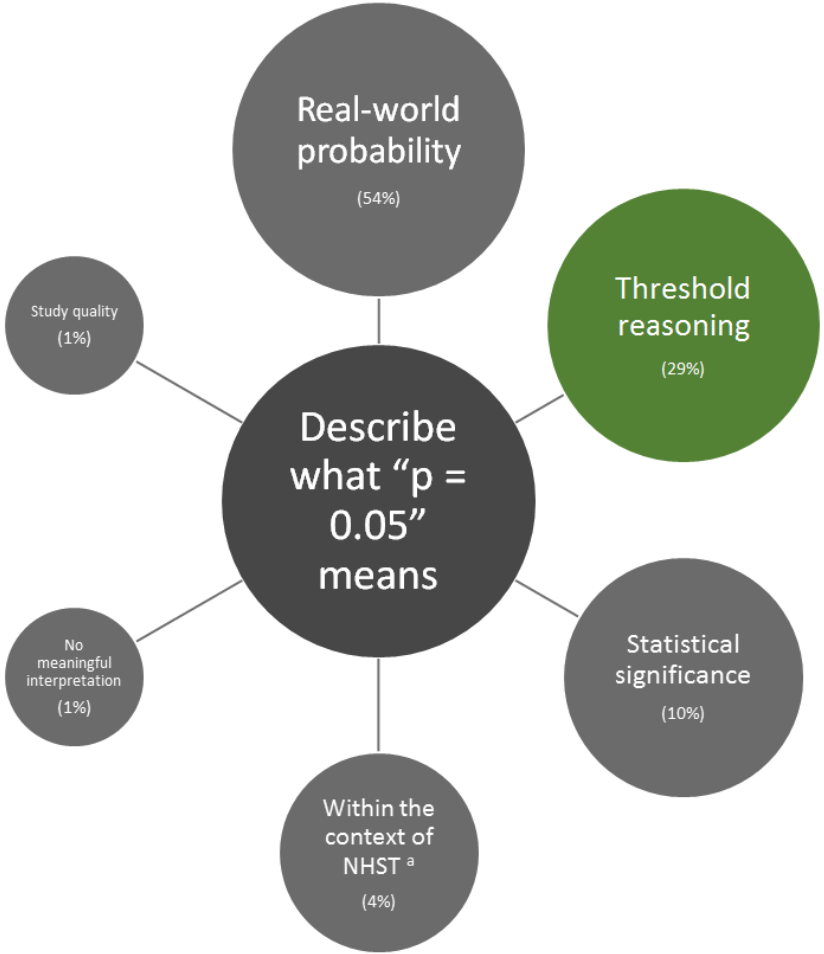


qualitative question

“We are 95% sure that the new drug is superior to the old drug. Or there is 5% chance the drugs perform equally well”

– participant 98

results

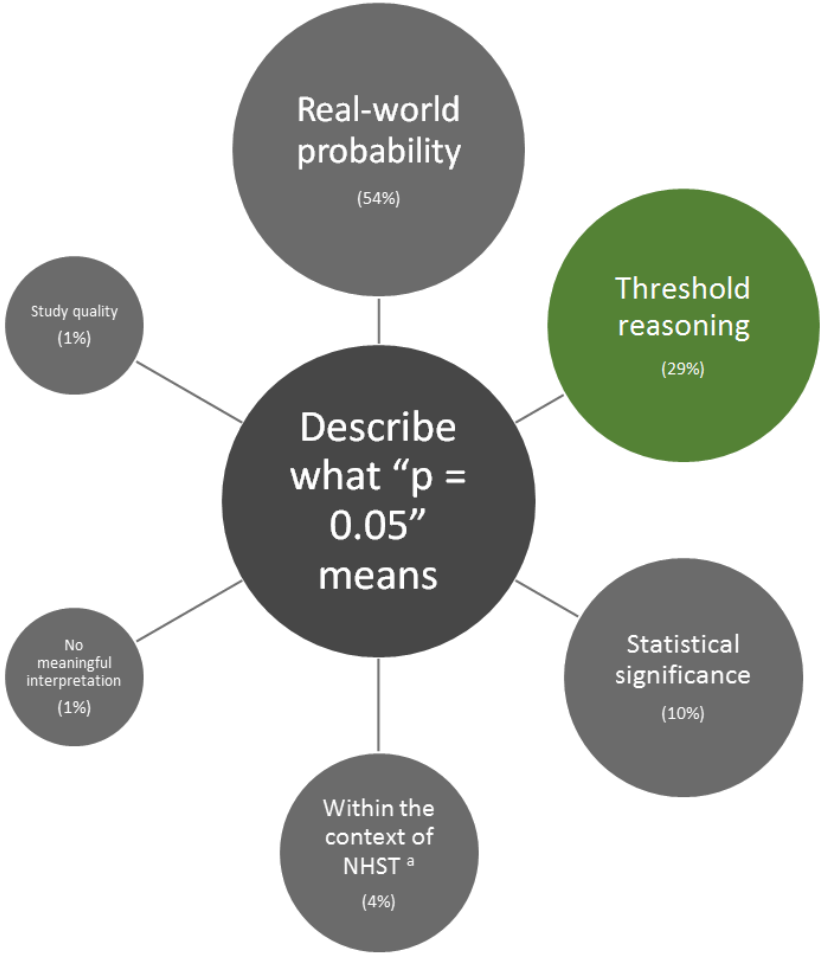


qualitative question

“That the new drug was statistically significant to show superiority over the old drug. The new drug is better than the old drug

– participant 138

results

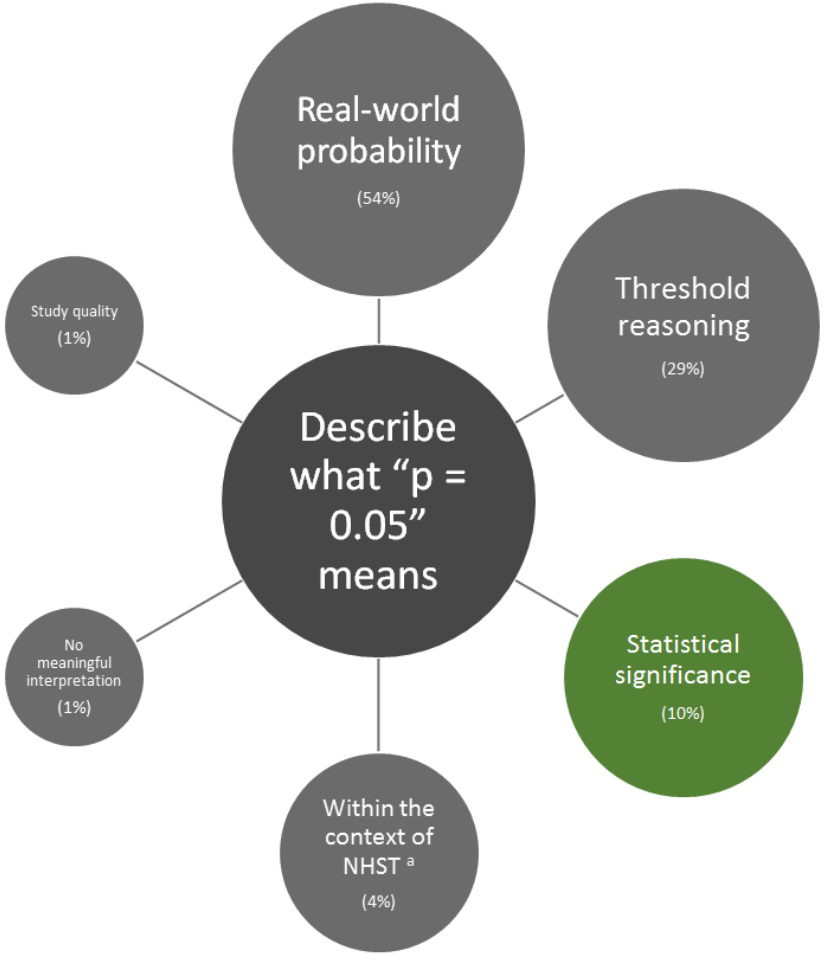


qualitative question

“As it is not less than 0.05, then there is not a statistically significant difference. Therefore the new drug can not be considered superior to the old drug based on this study.”

– participant 46

results



qualitative question

“cant remember a single thing, other than it means it is statistically significant in a research approved way.”

– participant 236

results

p = 0.05

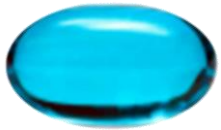
This means that there is a 5% probability that this result is due to chance alone, or, there is a 95% probability that the conclusion is true.

dichotomous choice question

“mostly FALSE”
(CORRECT response)

29%
95%CI: 24–35%

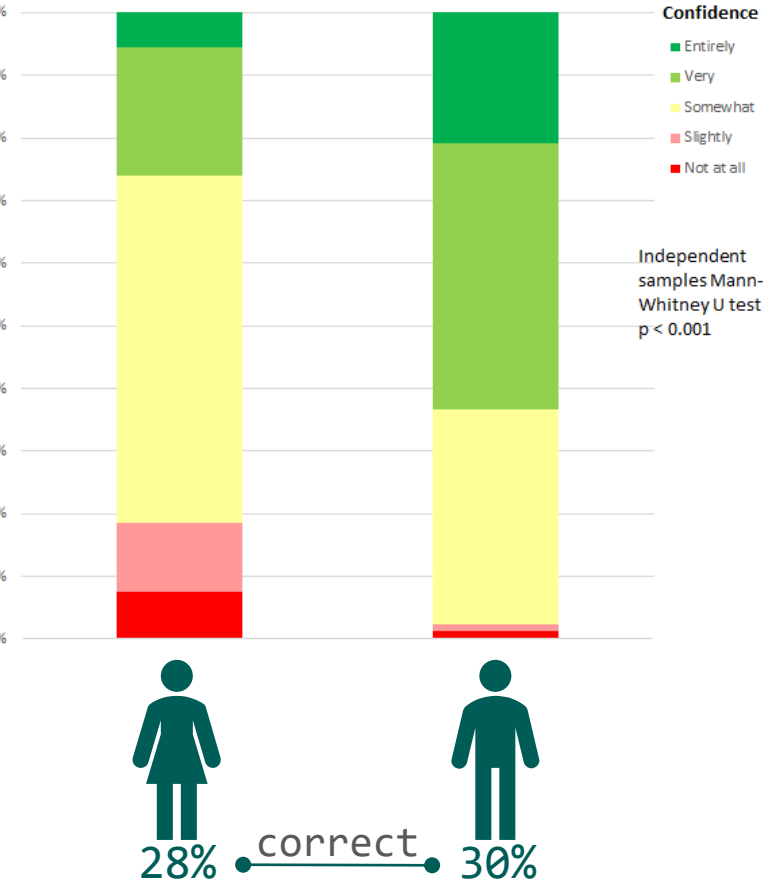
“mostly TRUE”
(INCORRECT response)



71%

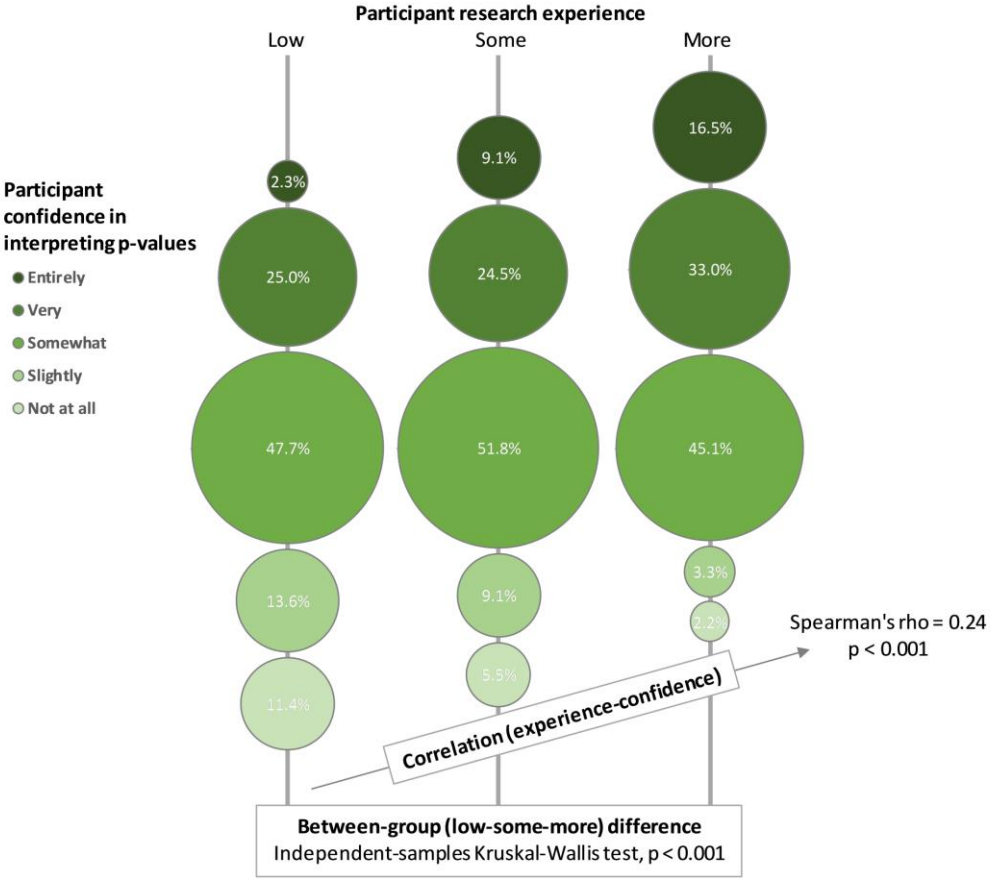
no demographic factors associated with better performance

results



confidence question

results

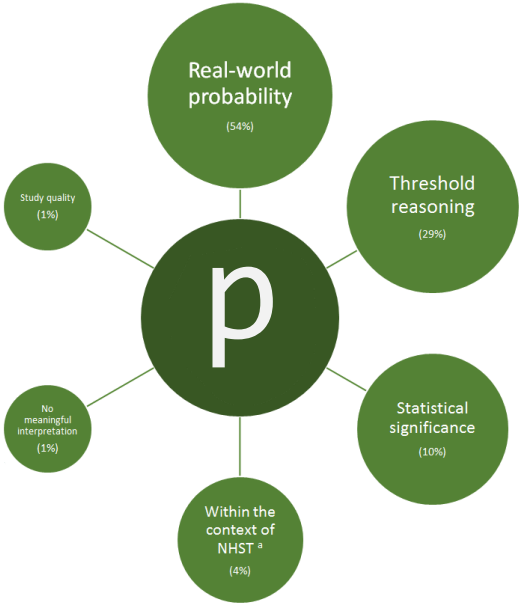


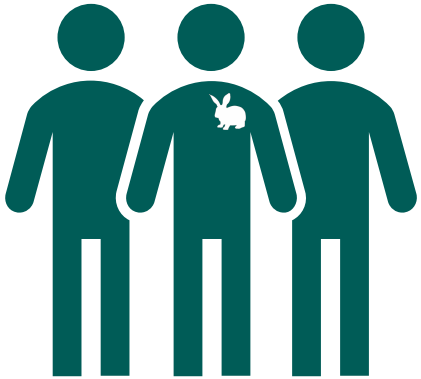
confidence question

correct

23% 33% 29%

synthesis





"… the probability under a specified statistical model that a statistical summary of the data (e.g. the sample mean difference between two compared groups) would be equal to or more extreme than its observed value."

Wasserstein & Lazar (2016) *for* ASA

p-values indicate the incompatibility of the data with the model – *within the system*

Numerically, they tell us nothing directly about the "real world".

"… the probability under a specified statistical model that a statistical summary of the data (e.g. the sample mean difference between two compared groups) would be equal to or more extreme than its observed value."

Wasserstein & Lazar (2016) *for* ASA

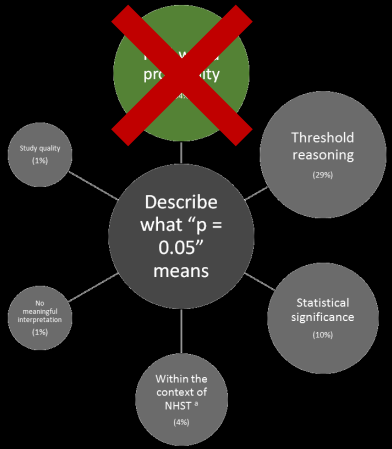
p-values indicate the incompatibility of the data with the model – *within the system*

Numerically, they tell us nothing directly about the “real world”.

p = 0.05
on the real world

prior	after
5%	11%
50%	71%
90%	96%

Nuzzo (2014)



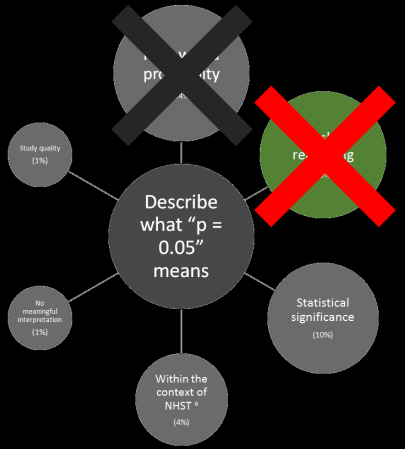
p-values indicate the incompatibility of the data with the model – *within the system*

Numerically, they tell us nothing directly about the "real world".

p as thresholds on evidence

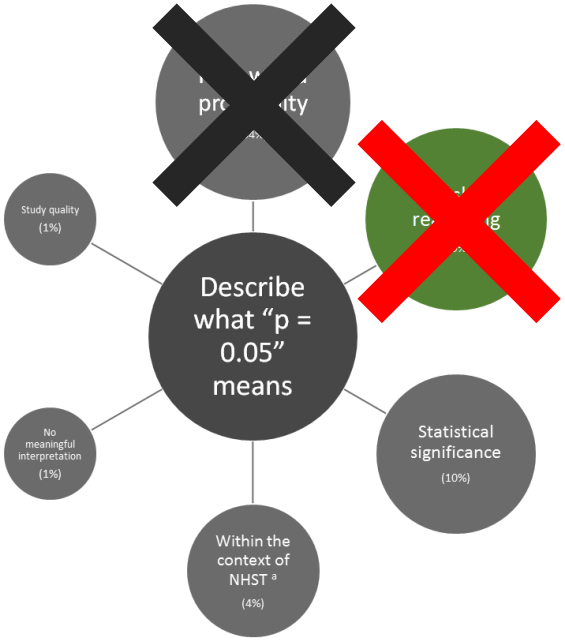
0.04 ≈ 0.06

∴ similar conclusion



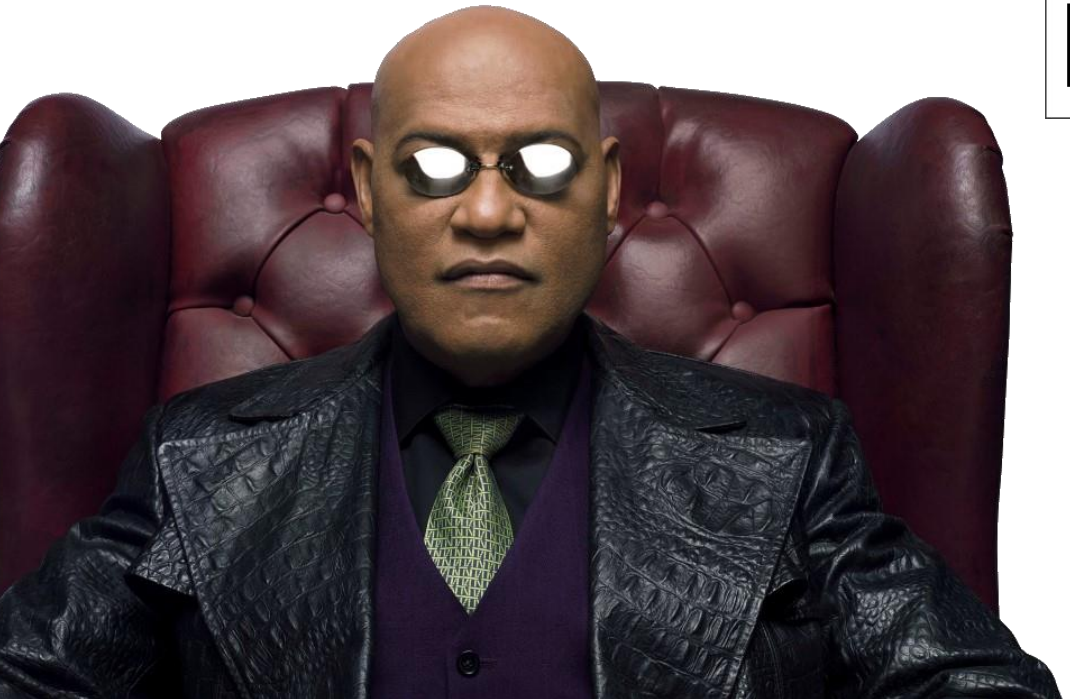
implications

Most clinicians get it wrong and don't know it.



Focus on the main misconceptions in education.

“…sooner or later you will realise, just as I did, there is a difference between knowing the path, and walking it.”



Tam CWM, Khan A, Knight A, Rhee J, McLean K, Price K. **How doctors conceptualise P values**. *Australian Journal of General Practice* 2018; 47: 705-10

tiny.cc/pvalues

TAMMichael
m.tam@unsw.edu.au