| Rogers Kaliisa, Kamila Misiejuk,
Sonsoles López-Pernas, Mohammad Khalil, Mohammed Saqr

Have Learning Analytics Dashboards Lived Up to the Hype? A Systematic Review of Impact on Students' Achievement, Motivation, Participation and Attitude

Rogers Kaliisa*
Department of Education, University of Oslo, Norway
rogers.kaliisa@iped.uio.no

Kamila Misiejuk
Centre for the Science of Learning & Technology (SLATE), University of Bergen
kamila.misiejuk@uib.no

Sonsoles López-Pernas
School of Computing, University of Eastern Finland
sonsoles.lopez@uef.fi

Mohammad Khalil
Centre for the Science of Learning & Technology (SLATE), University of Bergen
mohammad.khalil@uib.no

Mohammed Saqr
School of Computing, University of Eastern Finland
mohammed.saqr@uef.fi

ABSTRACT

While learning analytics dashboards (LADs) are the most common form of LA intervention, there is limited evidence regarding their impact on students' learning outcomes. This systematic review synthesizes the findings of 38 research studies to investigate the impact of LADs on students' learning outcomes, encompassing achievement, participation, motivation, and attitudes. As we currently stand, there is no evidence to support the conclusion that LADs have lived up to the promise of improving academic achievement. Most studies reported negligible or small effects, with limited evidence from well-powered controlled experiments. Many studies merely compared users and non-users of LADs, confounding the dashboard effect with student engagement levels. Similarly, the impact of LADs on motivation and attitudes appeared modest, with only a few exceptions demonstrating significant effects. Small sample sizes in these studies highlight the need for larger-scale investigations to validate these findings. Notably, LADs showed a relatively substantial impact on student participation. Several studies reported medium to large effect sizes, suggesting that LADs can promote engagement and interaction in online learning environments. However, methodological shortcomings, such as reliance on traditional evaluation methods, self-selection bias, the assumption that access equates to usage, and a lack of standardized assessment tools, emerged as recurring issues. To advance the research line for LADs, researchers should use rigorous assessment methods and establish clear standards for evaluating learning constructs. Such efforts will advance our understanding of the potential of LADs to enhance learning outcomes and provide valuable insights for educators and researchers alike.

*Corresponding author.

CCS CONCEPTS

• Intelligent systems; • Automated systems; • Education;

KEYWORDS

Learning analytics dashboards (LADs), systematic review, impact, learning outcomes

1 INTRODUCTION

The field of Learning Analytics (LA) has emerged as a promising avenue for leveraging data-driven insights to enhance educational processes and outcomes. LA can provide students and teachers with valuable feedback and support, enabling them to make informed decisions and optimise their learning and teaching practices. However, despite the proliferation of studies within the LA domain, a critical gap remains in understanding the impact and effectiveness of LA interventions, specifically focusing on LA dashboards (LADs) [1].

LADs may be defined as "displays that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualisations" [2] (p. 37), that have the potential to empower students and teachers by offering valuable insights into their learning and teaching processes [3–5]. These interactive tools aim to visualise data and provide actionable information, enabling learners and educators to monitor progress, identify areas of improvement, and make data-informed decisions. Despite over a decade of advancements and innovations within the LA field, there is a dearth of compelling evidence demonstrating the effectiveness and impact of LA interventions [1, 6], with only a few individual studies yet reporting mixed results mostly based on small samples [7]. This lack of empirical evidence poses a significant challenge when attempting to justify investments in expensive LA infrastructure and the necessary human resource training.

https://dl.acm.org/doi/abs/10.1145/3636555.3636884

Understanding **Learning** Analytics Dashboards

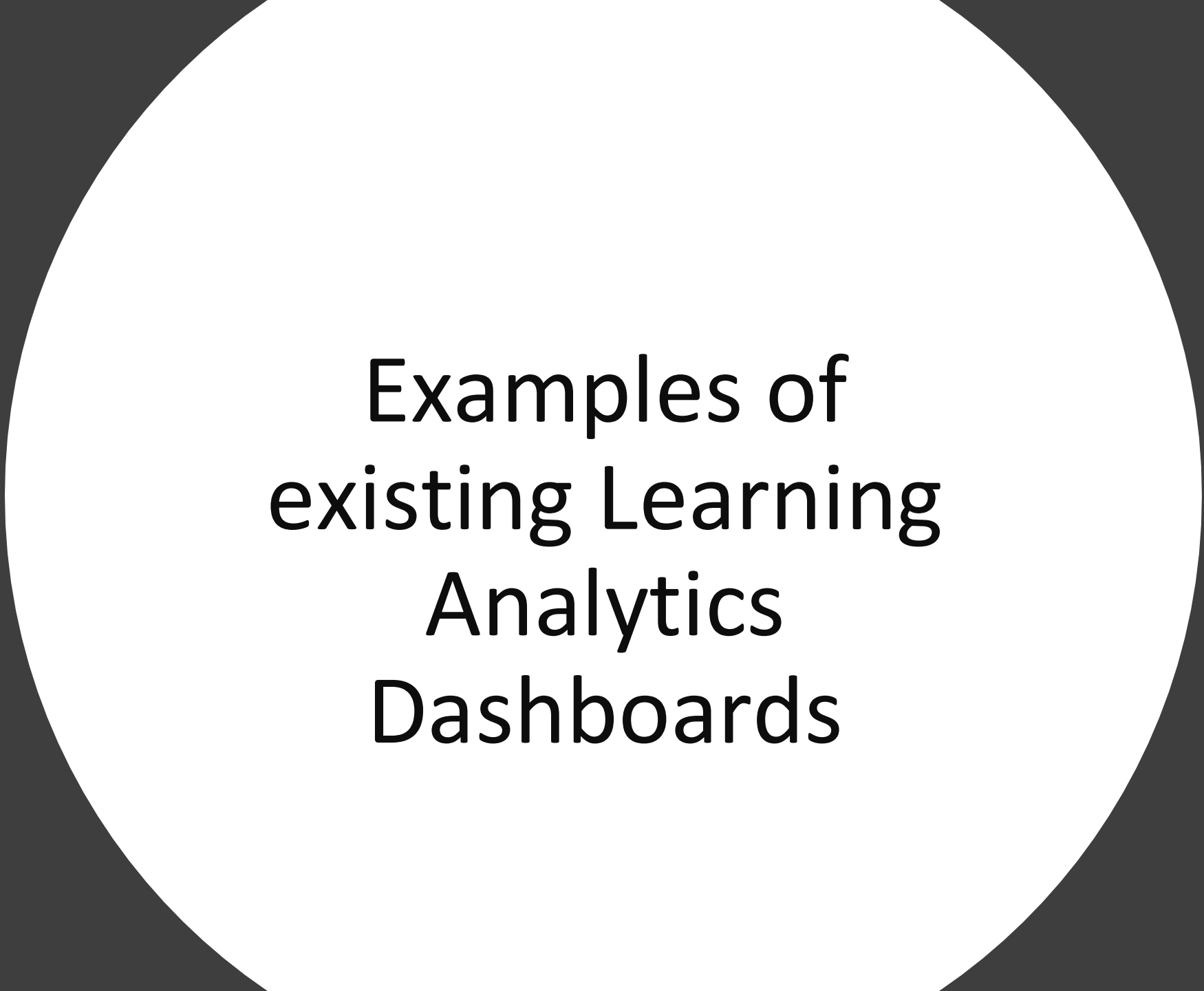"single displays that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple **visualizations**" (Schwendimann et al. 2016, p. 37)

Visualisation tools built with the purpose of empowering teachers and learners to make informed decisions about the learning process (Jivet et al., 2018)

Examples of existing Learning Analytics Dashboards

Signals

Summer 2009 - ANTH-34100-001 - XLST

[Add Intervention]   [Blackboard Stoplights]   [Intervention Emails]

Section Dashboard

Show/Hide Filter

Student	Int 1	Int 2	Int 3	Int 4	Int 5	Int 6	Int 7	Int 8	Int 9
Student A	🟡	🟡	🟡	🟢	🟢	🟢	🔴	🔴	🔴
Student B	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Student C	🟢	🟢	🟢	🟢	🟢	🟢	🔴	🔴	🔴
Student D	🟡	🟡	🟢	🟡	🟡	🟡	🟡	🟡	🟡
Student E	🟢	🟢	🟡	🟡	🟡	🟡	🔴	🔴	🔴
Student F	🟡	🟡	🟡	🟡	🔴	🟡	🔴	🔴	🔴
Student G	🟢	🟢	🟡	🟡	🟡	🟡	🟢	🟢	🟢
Student H	🟡	🟡	🟡	🟢	🟢	🟢	🔴	🔴	🔴
Student I	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢	🟢
Student J	🟢	🟢	🟢	🟢	🟢	🟢	🟡	🟡	🟢

PURDUE
UNIVERSITY

Course Signals at Purdue: Using Learning Analytics to Increase Student Success

Kimberly E. Arnold
Purdue University
519 Young Hall, 155 S. Grant Street
West Lafayette, IN 47907 USA
kimarnold@purdue.edu

Matthew D. Pistilli
Purdue University
517 Young Hall, 155 S. Grant Street
West Lafayette, IN 47907 USA
mdpistilli@purdue.edu

ABSTRACT

In this paper, an early intervention solution for collegiate faculty called Course Signals is discussed. Course Signals was developed to allow instructors the opportunity to employ the power of learner analytics to provide real-time feedback to a student. Course Signals relies not only on grades to predict students' performance, but also demographic characteristics, past academic history, and students' effort as measured by interaction with Blackboard Vista, Purdue's learning management system. The outcome is delivered to the students via a personalized email from the faculty member to each student, as well as a specific color on a stoplight – traffic signal – to indicate how each student is doing. The system itself is explained in detail, along with retention and performance outcomes realized since its implementation. In addition, faculty and student perceptions will be shared.

solutions should be focused on all students at an institution, not just a specific subpopulation. Finally, solutions implemented to enhance student success, and therefore persistence, needed to help integrate a student academically into the institution [6].

Helping a student become academically integrated to the institution is key, as Course Signals helps to promote integration in several ways. First, it allows faculty members to send personalized emails to students that contain information about their current performance in a given course. Second, faculty members can encourage students to visit various help resources on campus or office hours – activities that contribute to a student becoming more fully integrated into the institution. Third, it employs learner analytics to allow for the integration of real-time data on student performance and interaction with the LMS with demographic and past academic history information. This combination creates an intentionally created environment

https://zoomsense.io/#/scheduled

(Pozdniakov et al., 2022)

EAI

Interfac

Features

- Short-term or weekly predictions
- Uses static data (e.g. age, gender)
- Helps teachers to support students in their learning journey
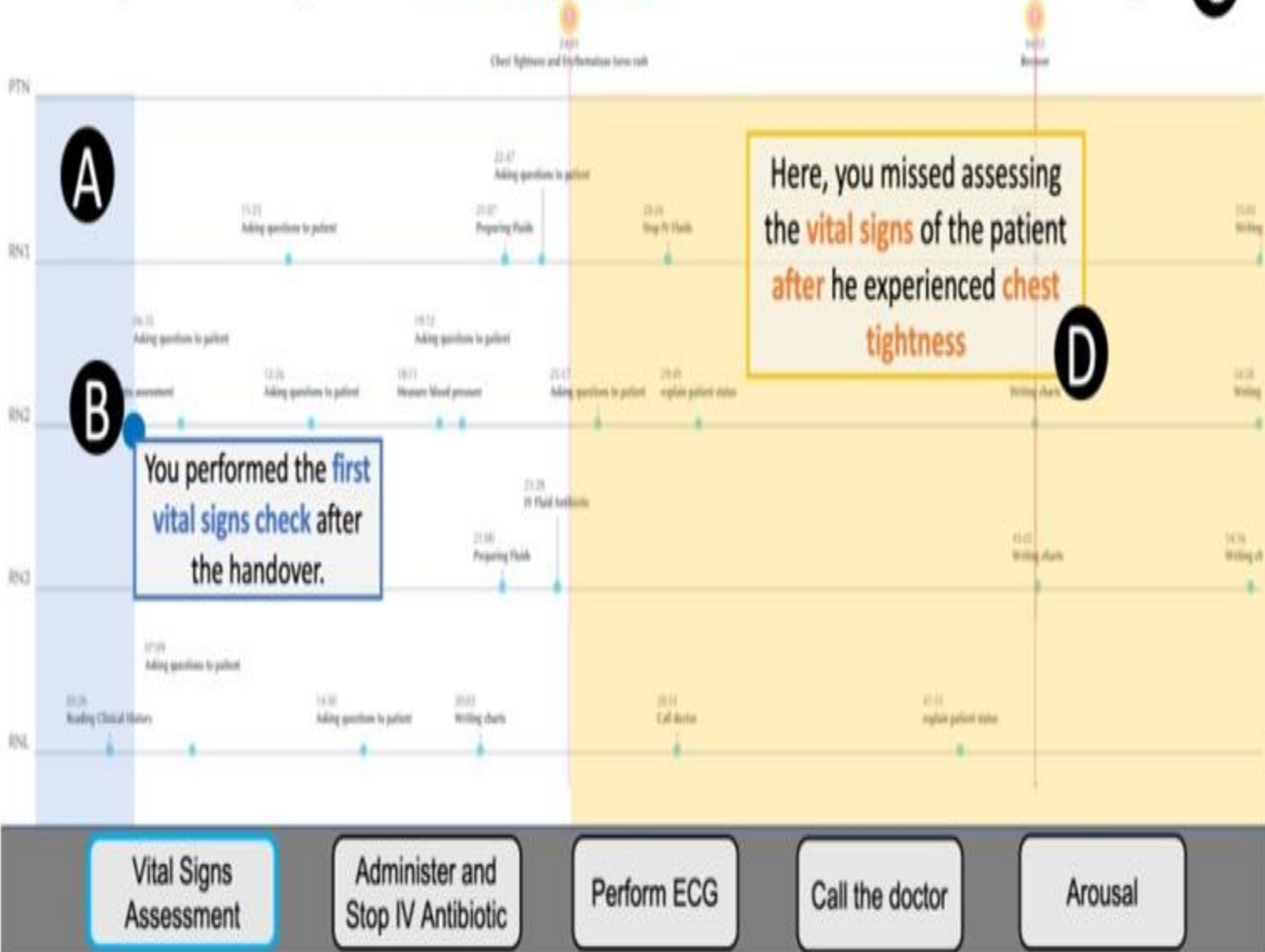
Beyond the Learning Analytics Dashboard: Alternative Ways to Communicate Student Data Insights Combining Visualisation, Narrative and Storytelling

Gloria Milena Fernández-Nieto*
University of Technology Sydney
Sydney, NSW, Australia
gloria.m.fernandeznieto@student.uts.edu.au

Simon Buckingham Shum
University of Technology Sydney
Sydney, NSW, Australia
Simon.BuckinghamShum@uts.edu.au

Kirsty Kitto
University of Technology Sydney
Sydney, NSW, Australia
Kirsty.Kitto@uts.edu.au

Roberto Martínez-Maldonado
Monash University
Melbourne, Australia
Roberto.MartinezMaldonado@monash.edu

(Fernandez Nieto et al., 2022)

Canvas
Discussion Analytics

This app helps the teacher to gain a quick glance of the Canvas discussion activities

Q CanvasLA Mock discussion: Technology & Learning   X   All

Dashboard   Discourse analytics   Participants   Network   Sentiments

Participant
Total participants
5

Discussion
Total interactions
29

Sentiment
Positive

Discourse analytics

school	20	smartboard	13	language	6
teaching	5	like	5	using	5
learn	5	class	5	norway	5

Participants

29 %
10/14
71 %

● Active participants
● Inactive participants

ORIGINAL RESEARCH
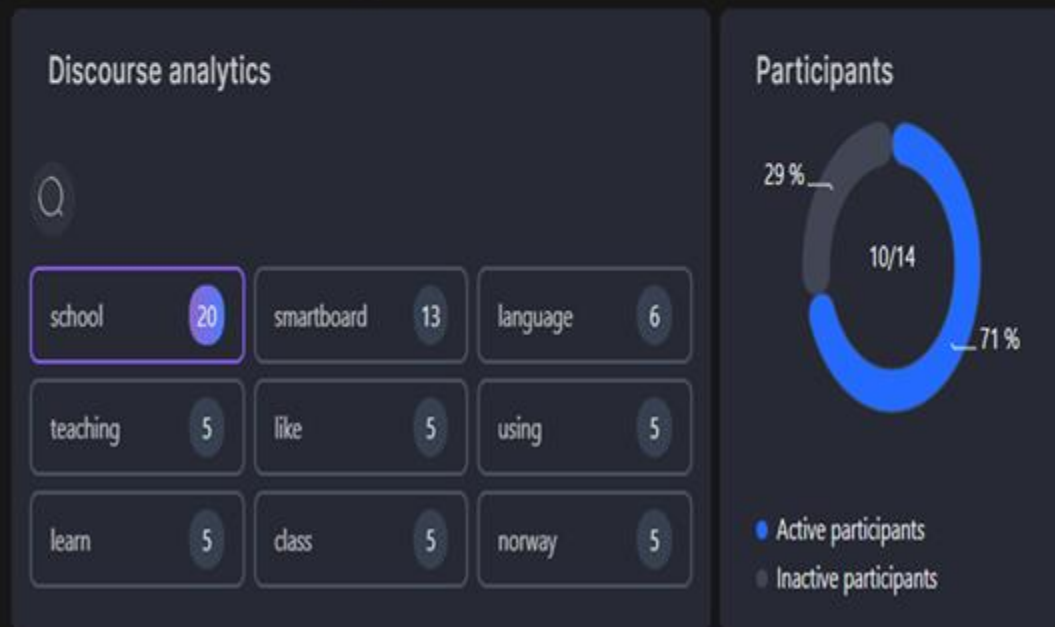
CADA: a teacher-facing learning analytics dashboard to foster teachers' awareness of students' participation and discourse patterns in online discussions

Rogers Kaliisa[1] · Jan Arild Dolonen[1]

Abstract

Despite the potential of learning analytics (LA) to support teachers' everyday practice, its adoption has not been fully embraced due to the limited involvement of teachers as co-designers of LA systems and interventions. This is the focus of the study described in this paper. Following a design-based research (DBR) approach and guided by concepts from the socio-cultural perspective and human-computer interaction (HCI), we design, test, and evaluate a teacher-facing LA dashboard, the Canvas Discussion Analytics Dashboard (CADA), in real educational settings. The goal of this dashboard is to support teachers' roles in online environments through insights into students' participation and discourse patterns. We evaluate CADA through 10 in-depth interviews with university teachers to examine their experiences using CADA in seven blended undergraduate and graduate courses over a one-year period. The findings suggest that engaging teachers throughout the analytics tool design process and giving them control/agency over LA tools can favour their adoption in practice. Additionally, the alignment of dashboard metrics with relevant theoretical constructs allows teachers to monitor the learning designs and make course design changes on the fly. The teachers in this study emphasise the need for LA dashboards to provide actionable insights by moving beyond *what things are* towards *how things should be*. This study has several contributions. First, we make an artefact contribution (e.g. CADA), an LA dashboard to support teachers with insights into students' online discussions. Second, by leveraging theory, and working with the teachers to develop and implement a dashboard in authentic teaching environments, we make an empirical, theoretical and methodological contribution to the field of learning analytics and technology enhanced learning. We synthesise these through practical design and implementation considerations for researchers, dashboard developers, and higher education institutions.

Keywords Teacher-facing learning analytics dashboard · Asynchronous online

Why Learning Analytics Dashboards

*The Key Intervention for LA

✓ Teacher-facing LADs

  ○ Reflection & decision making

  ○ Learning design adaptation

  ○ Timely feedback

✓ Student-facing LADs

  ○ Engagement & motivation

  ○ Track progress

  ○ Sense making

  ○ Awareness

Overall objective **'Support & Improve Learning'** [Manly & Ochoa, 2023].

Existing Studies

Staying on target: A systematic literature review on learner-facing learning analytics dashboards

Natercia Valle[1] | Pavlo Antonenko[1,2] | Kara Dawson[1,2] | Anne Corinne Huggins-Manley[2]

License to Evaluate: Preparing Learning Analytics Dashboards for Educational Practice

Ioana Jivet
Open University of the Netherlands
Heerlen, the Netherlands
ioana.jivet@ou.nl

Maren Scheffel
Open University of the Netherlands
Heerlen, the Netherlands
maren.scheffel@ou.nl

A checklist to guide the planning, designing, implementation, and evaluation of learning analytics dashboards

Rogers Kaliisa[1*], Ioana Jivet[2] and Paul Prinsloo[3]

A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective

Wannisa Matcha, Nora'ayu Ahmad Uzir, Dragan Gašević, and Abelardo Pardo

A Review of Learning Analytics Dashboard Research in Higher Education: Implications for Justice, Equity, Diversity, and Inclusion

Kimberly Williamson
Cornell University
Ithaca, NY, USA
khw44@cornell.edu

Rene F. Kizilcec
Cornell University
Ithaca, NY, USA
kizilcec@cornell.edu

Participatory and Co-Design of Learning Analytics: An Initial Review of the Literature

Juan Pablo Sarmiento
sarmiento@nyu.edu
NYU LEARN
New York University
New York City, New York, USA

Alyssa Friend Wise
alyssa.wise@nyu.edu
NYU LEARN
New York University
New York City, New York, USA

Gaps in existing work

- Most dash-board **studies** focus on assessing the tool's **usability** (Jivet 2018).

- **Little attention to evaluating the effects of LADS on students' learning outcomes e.g., cognitive and emotional** [Manly & Ochoa, 2023].

- Few studies have conducted **a quantitative review** focusing on the **impact** of LADs on students' learning outcomes.

- Lack of quantitative evidence poses a challenge to **justify investments** in expensive LAD infrastructure & human resource training.

- Identifying **absence or presence of evidence** provides guidance for future LADs research.

Research Question

What is the impact of LADs on students' learning outcomes?

- Performance
- Participation
- Motivation
- Attitudes
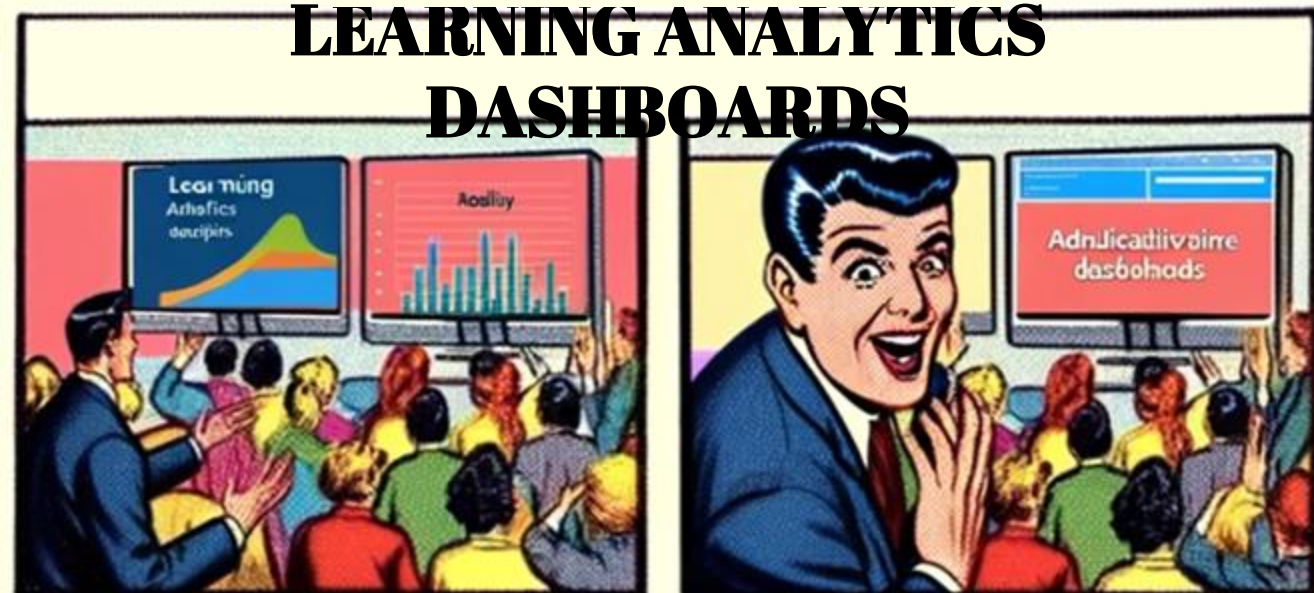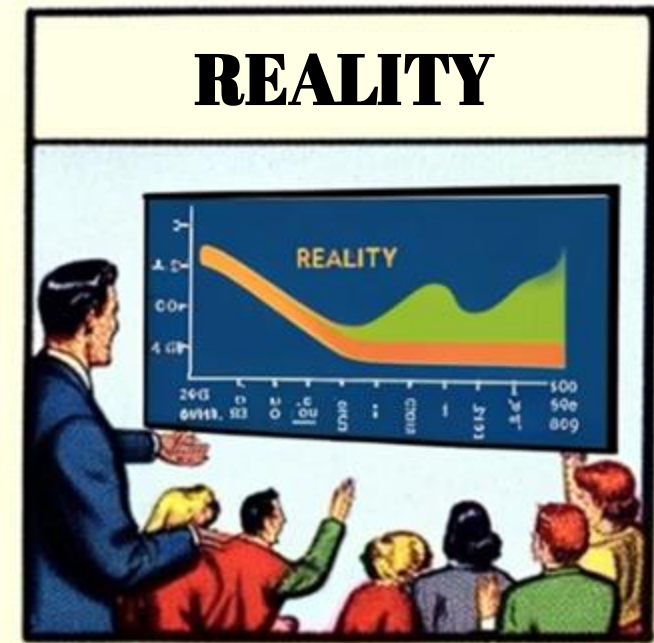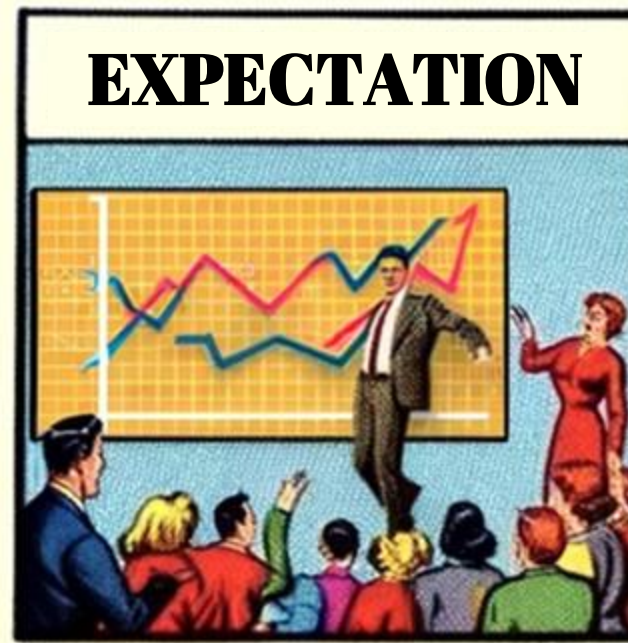
Methodology: A Systematic Quantitative Review

Records identified through database search (N = 812)
Scopus (n = 485) | ACM (n = 86) | Web of Science (n = 153) | ERIC (n = 88)

Duplicates removed (n = 284)

Records screened by abstract (n = 528)

Records excluded (n = 470)

Full-text assessed for eligibility (n = 58)

Records excluded (n = 20)
Relevant statistics not reported (n = 4)
Not testing the impact of a dashboard (n = 12)
Not relevant target group (n = 3)
Duplicate (n = 1)

Records included in review (n = 38)

Identification | Screening | Eligibility | Included

❖ Bottom-up identification of learning outcomes

❖ We found very few studies with **congruent research setups** and all the **statistical** information necessary to allow **a meta-analysis**

❖ Thus, we extracted the reported quantitative metrics (e.g., sample, effect size, mean) and reported them descriptively

❖ Studies with enough info., we classified the variables for **sub-group analysis** and converted the **effect size** to a common unit (Cohen's d) to facilitate comparison.

❖ We used **Cohen's d,** to categorise the effect size:

  ➢ a value over 0.8 (large)
  ➢ a value of 0.5 (medium)
  ➢ a value of 0.2 (small)
  ➢ a value below 0.2 (negligible)

Findings

Studies per year
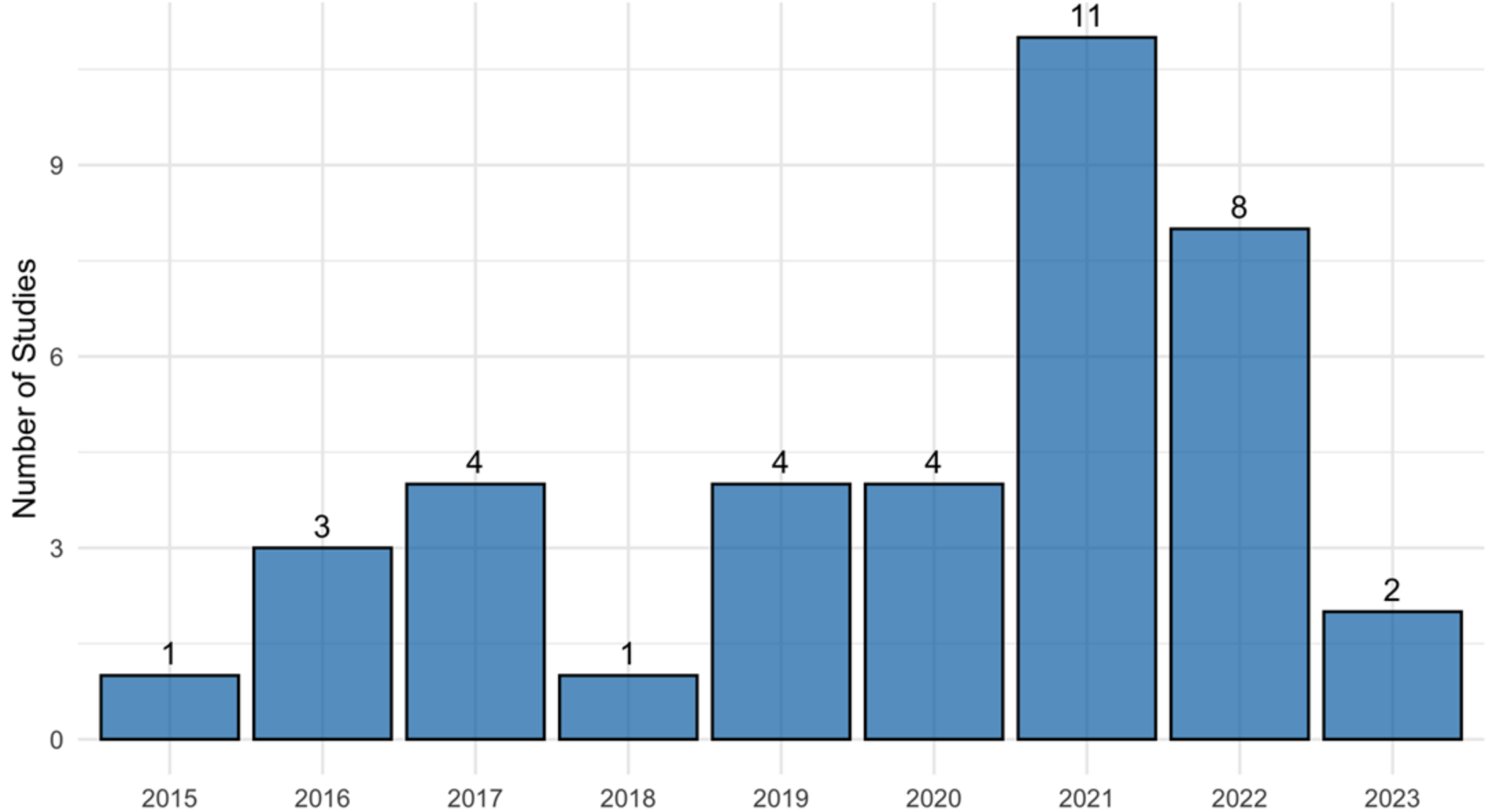
Discipline and stakeholders
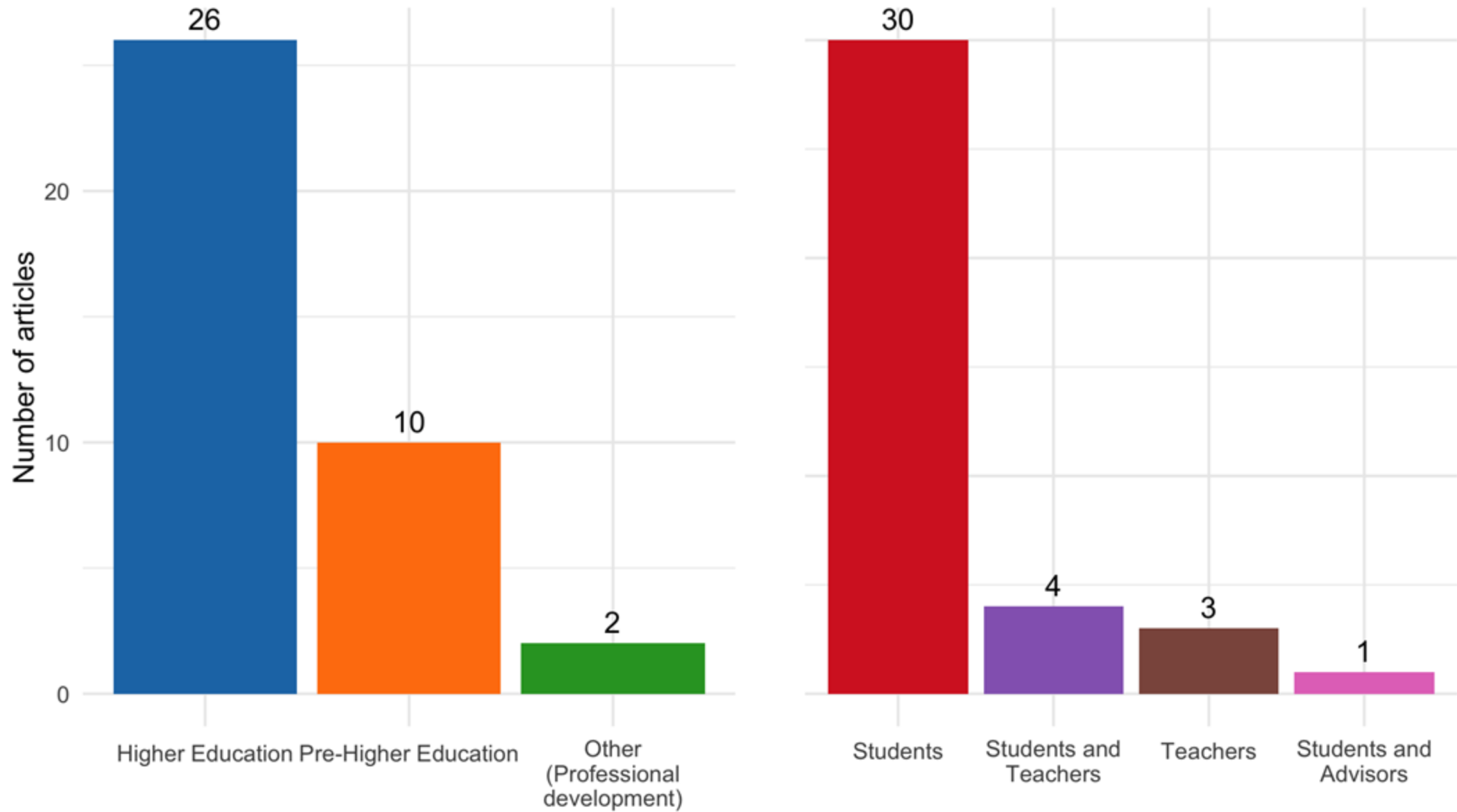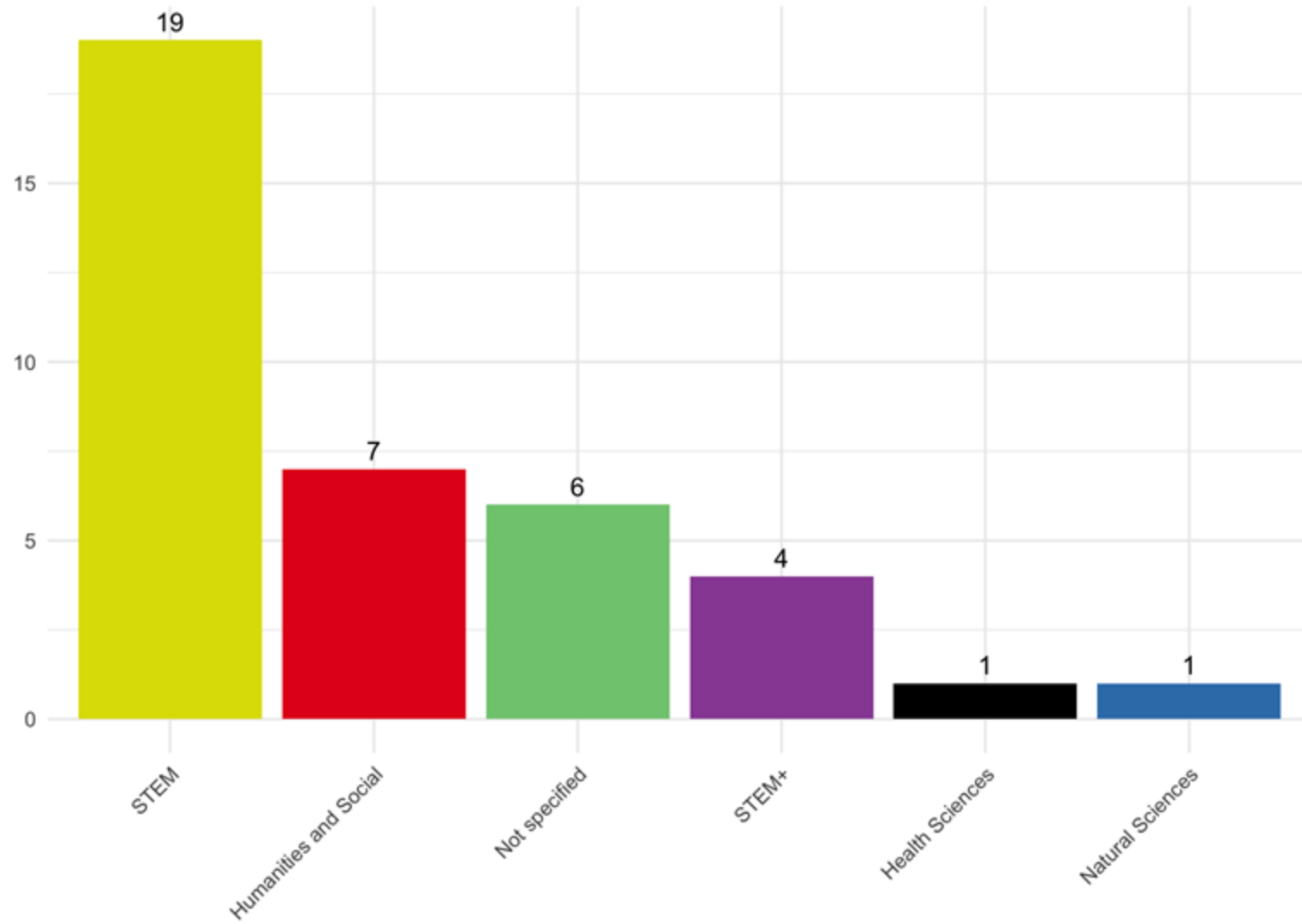
Study design
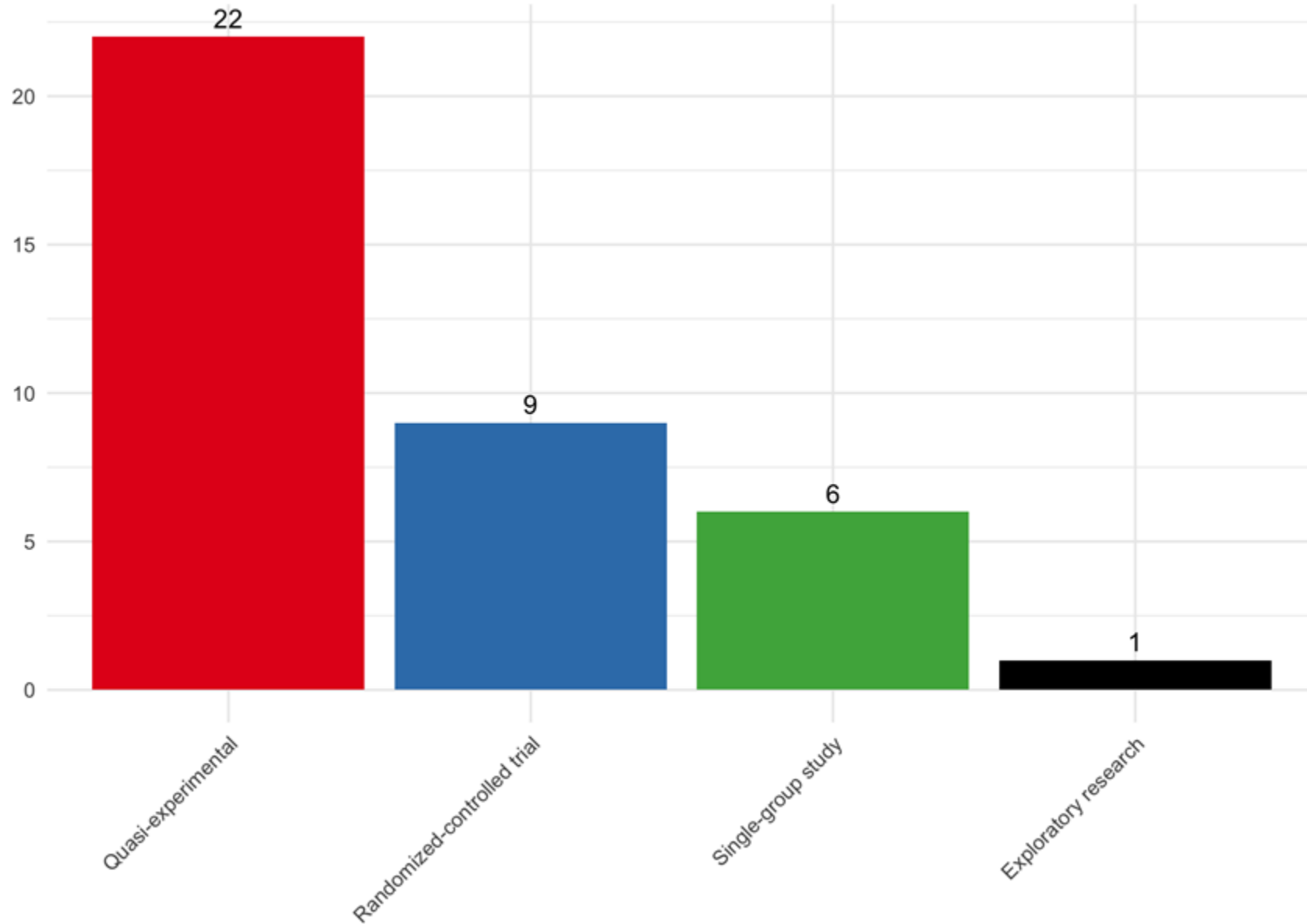
Achievement

Achievement: Effect size descriptives (count of votes)
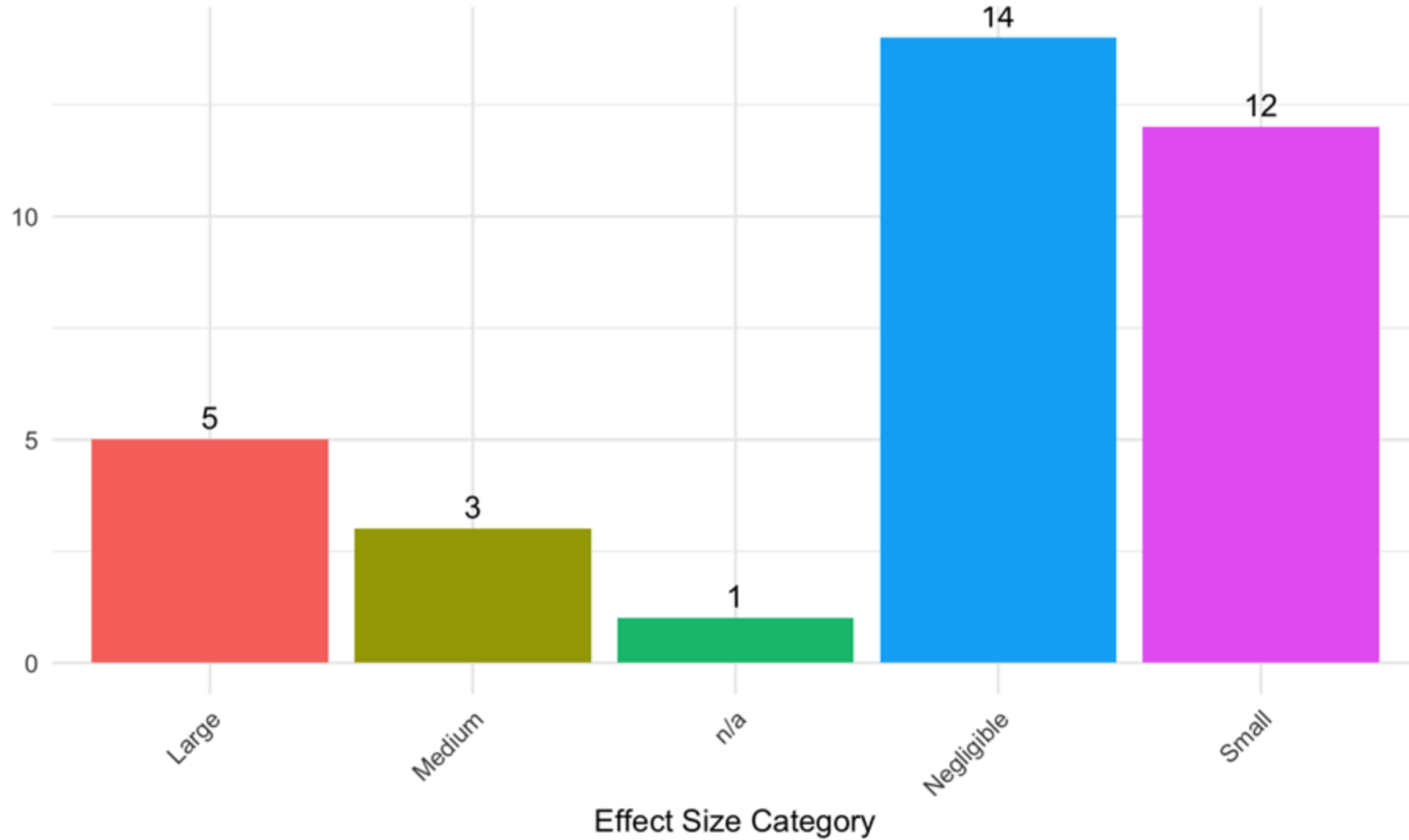
Achievement: Effect size type

Achievement: Average Effect Size by Measure
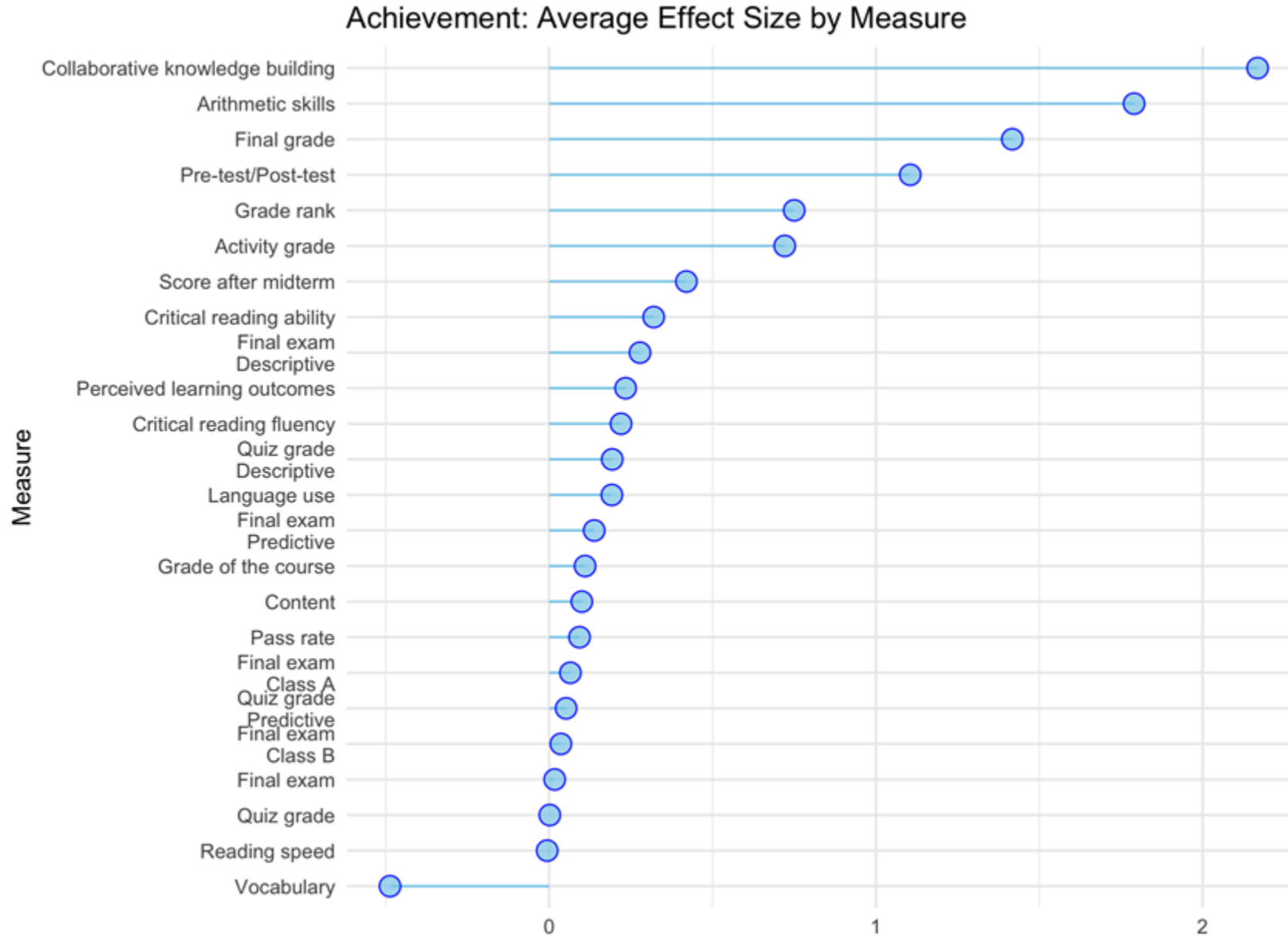
Achievement: Effect size versus sample size
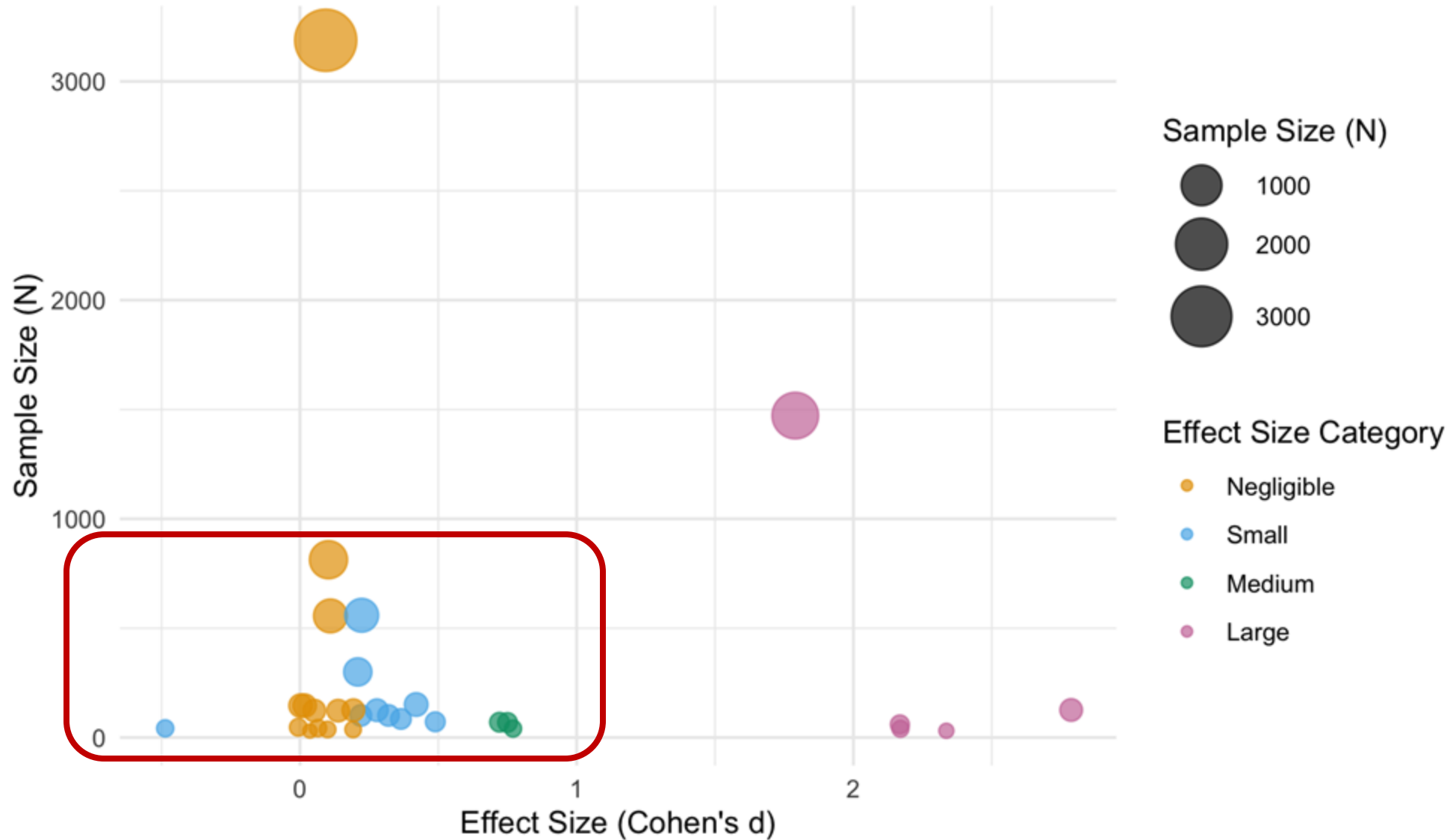
Achievement: Effect size by study design

Correlation: Effect Size vs. Sample Size

$R^2 = 0.0017$, $p = 0.83$
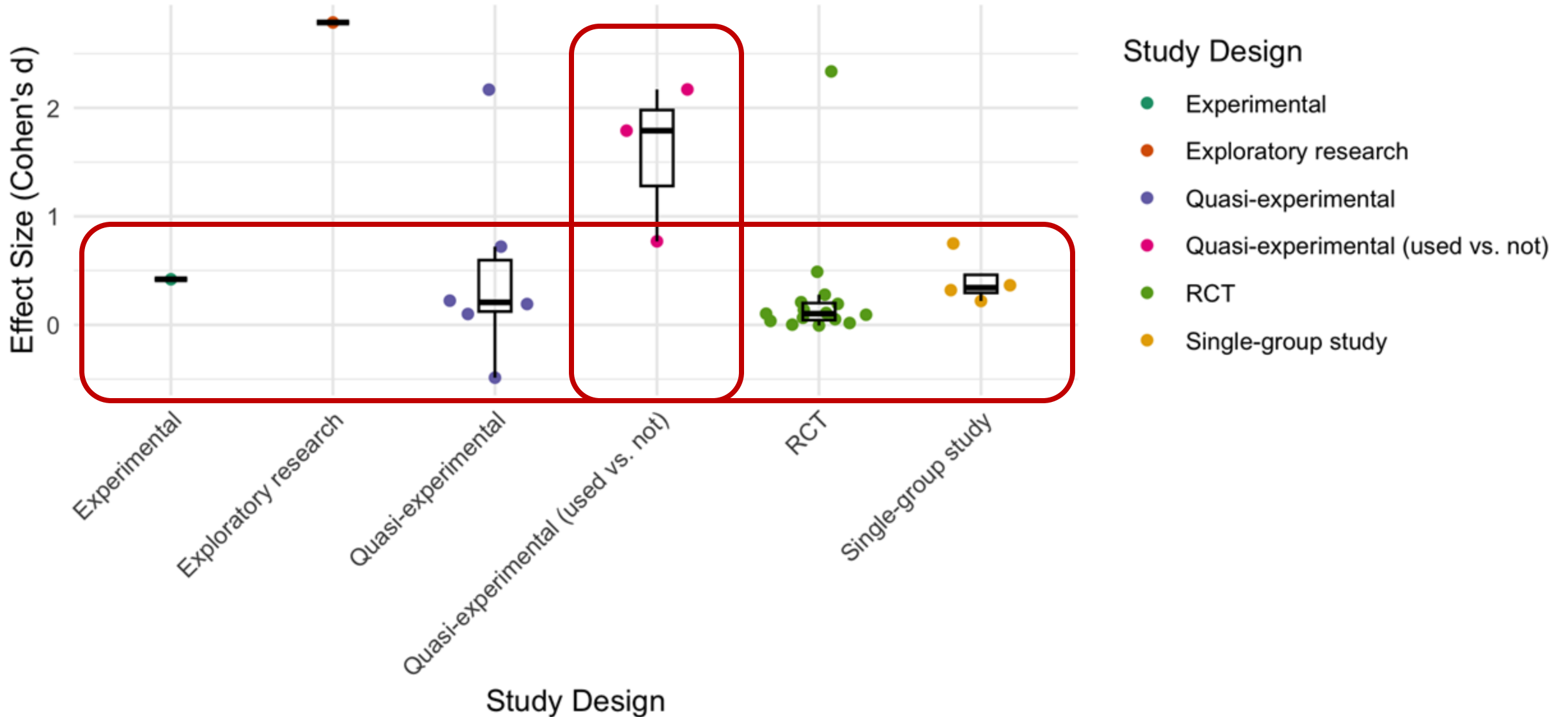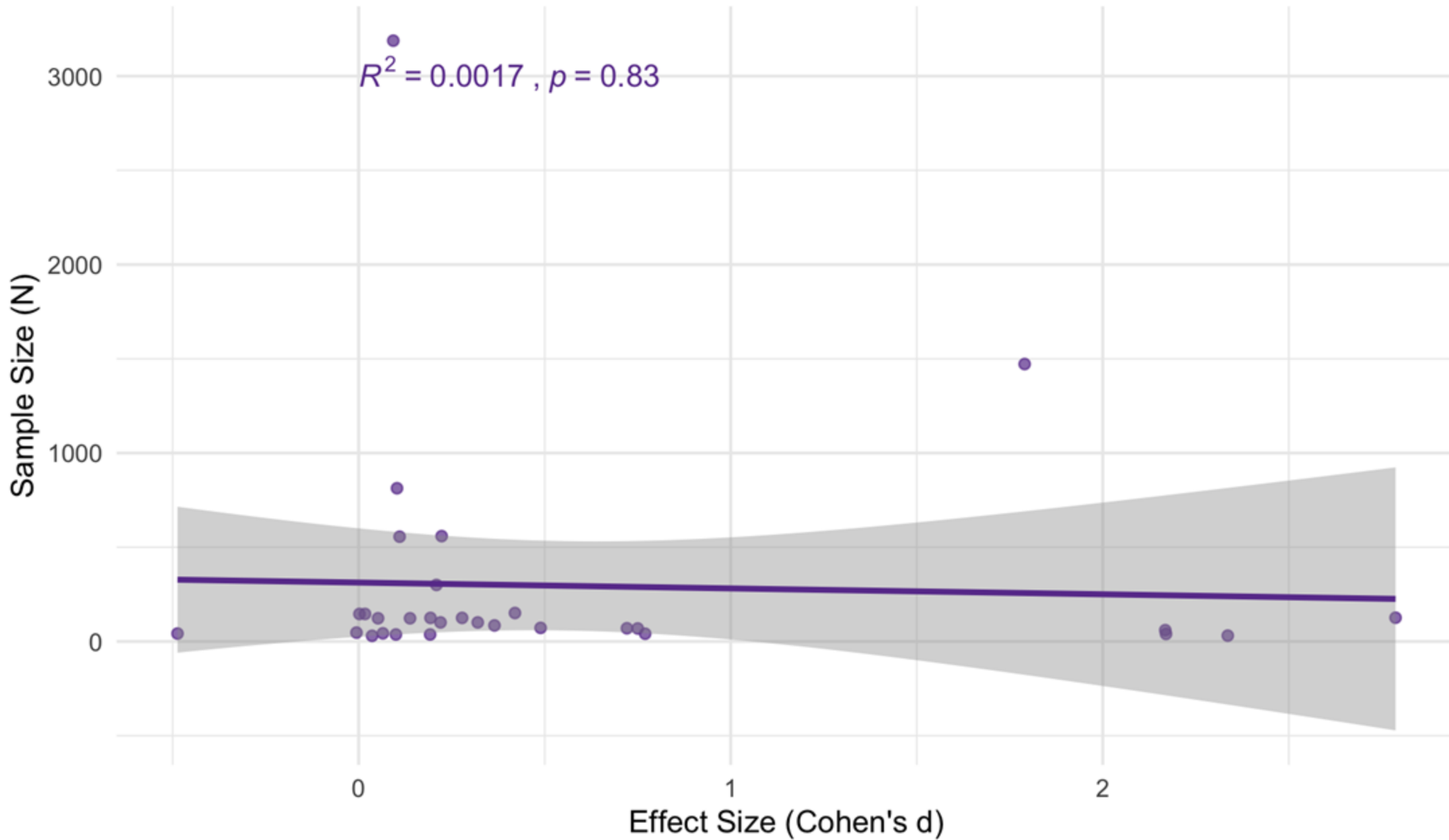
Participation

Participation: Effect size descriptives (count of vote)
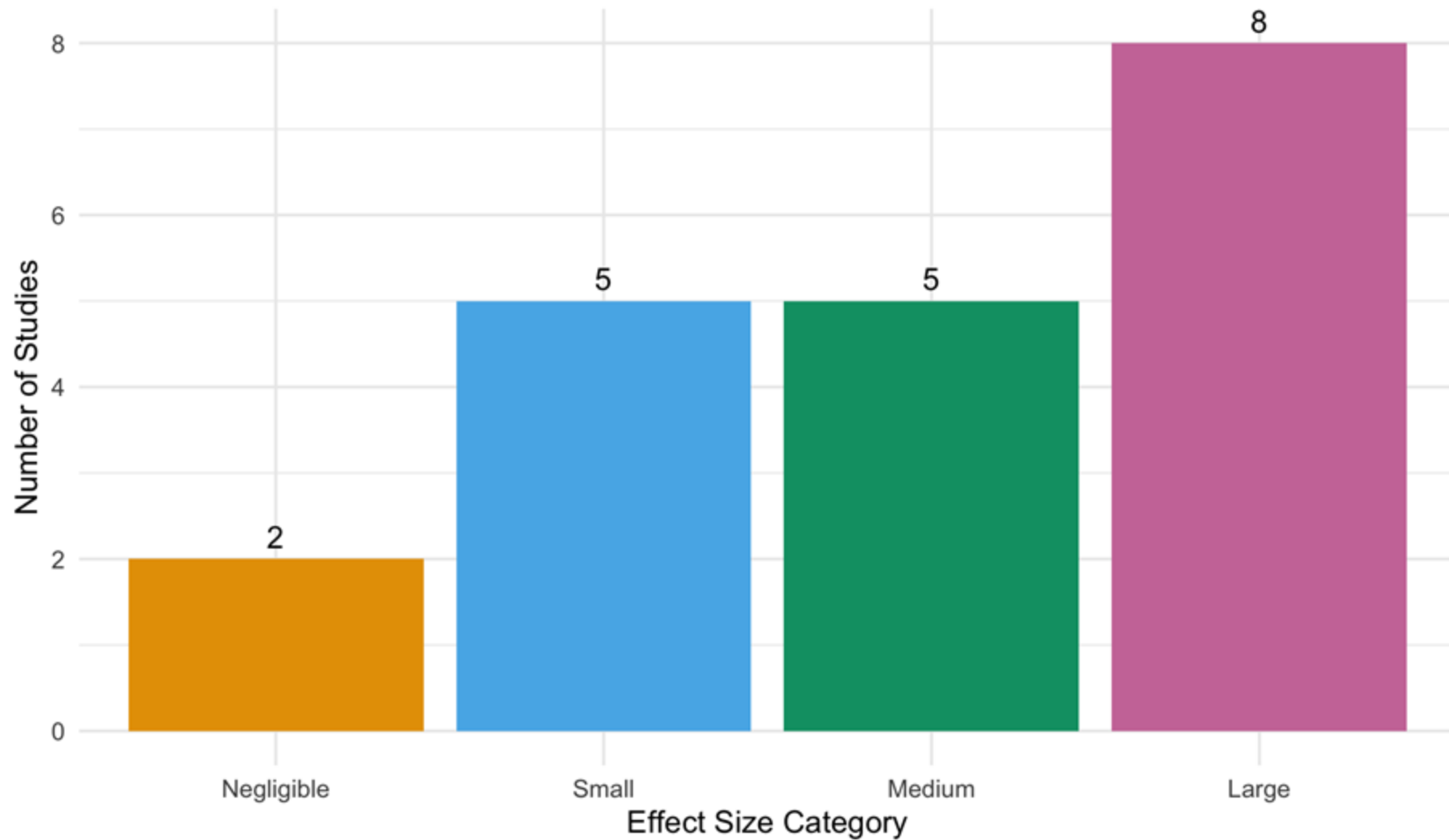
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Participation: Effect size type
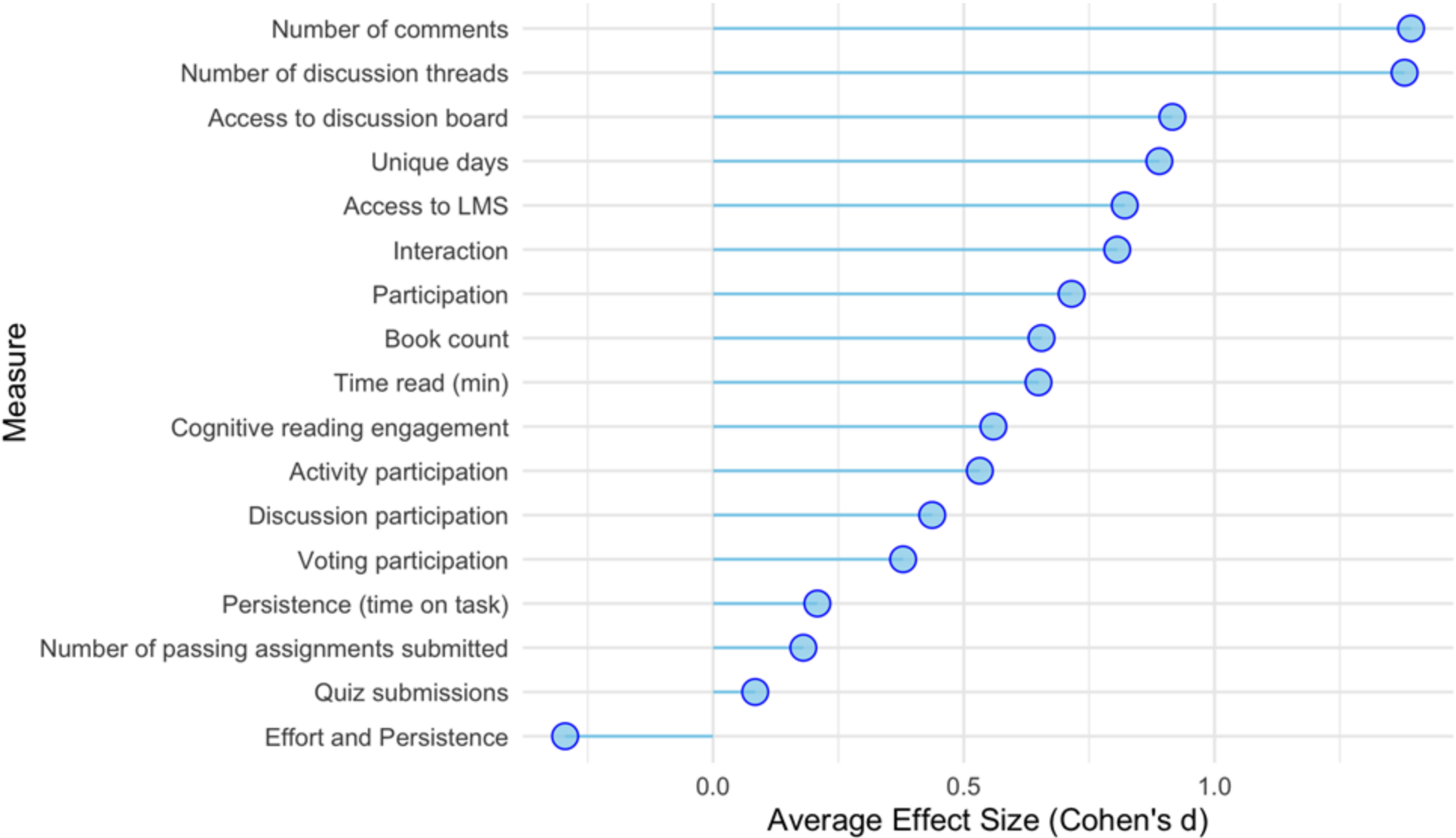
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Participation: Effect size versus sample size

Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Correlation: Effect Size vs. Sample Size

$R^2 = 0.35$, $p = 0.0075$

Sample Size (N)

Effect Size (Cohen's d)
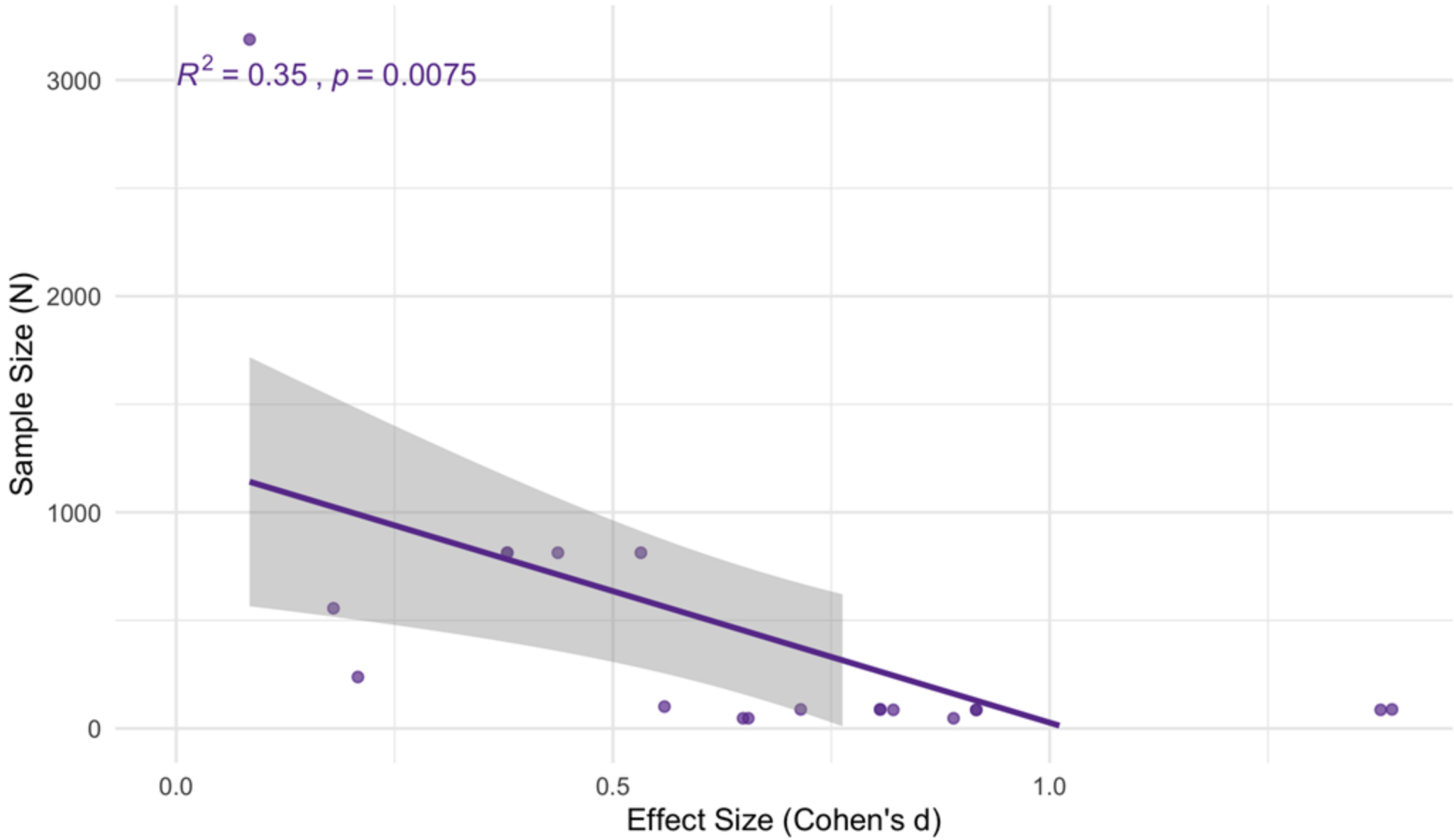
Motivation

Motivation: Effect size descriptives (count of vote)
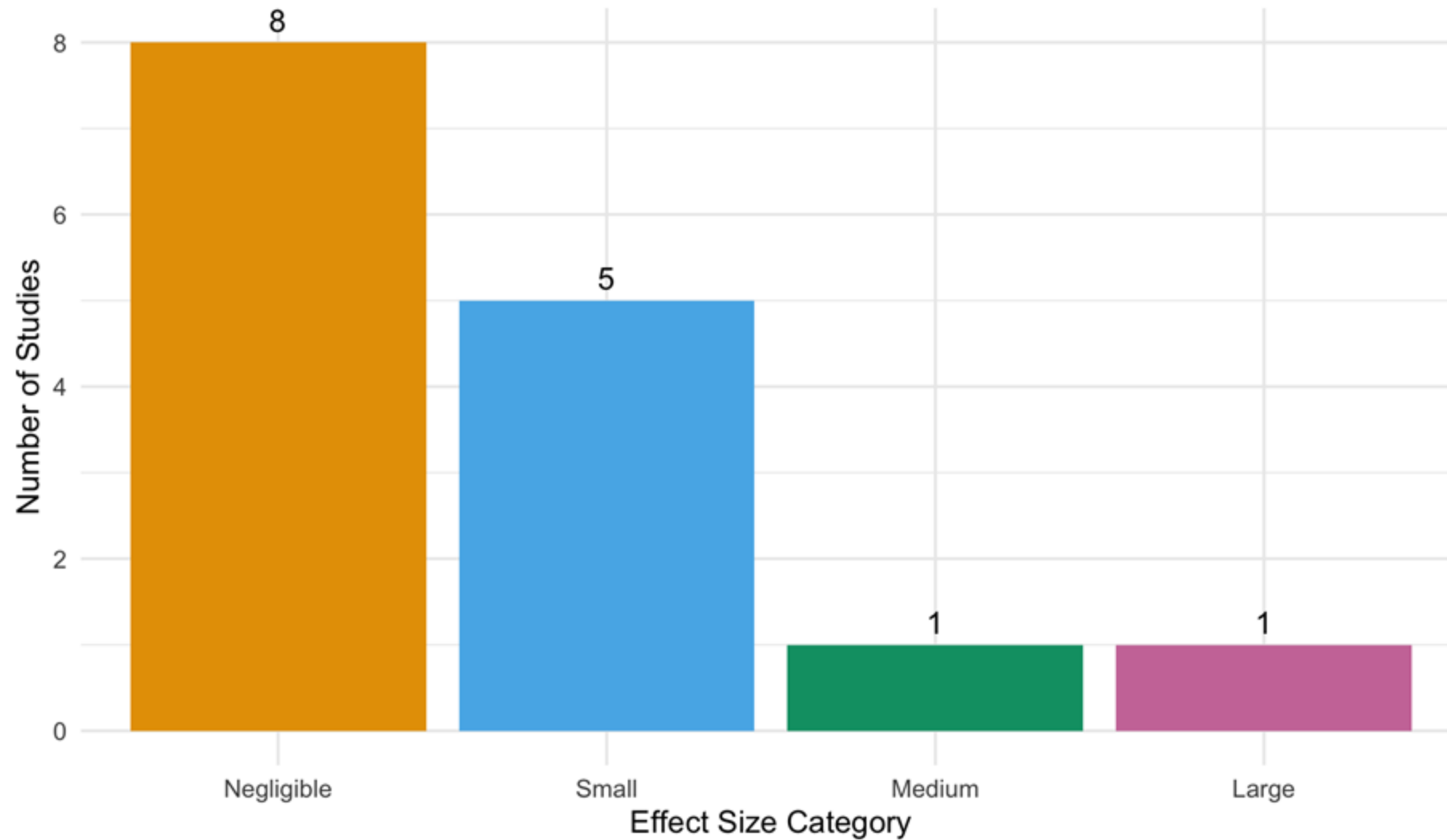
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Participation: Effect size type
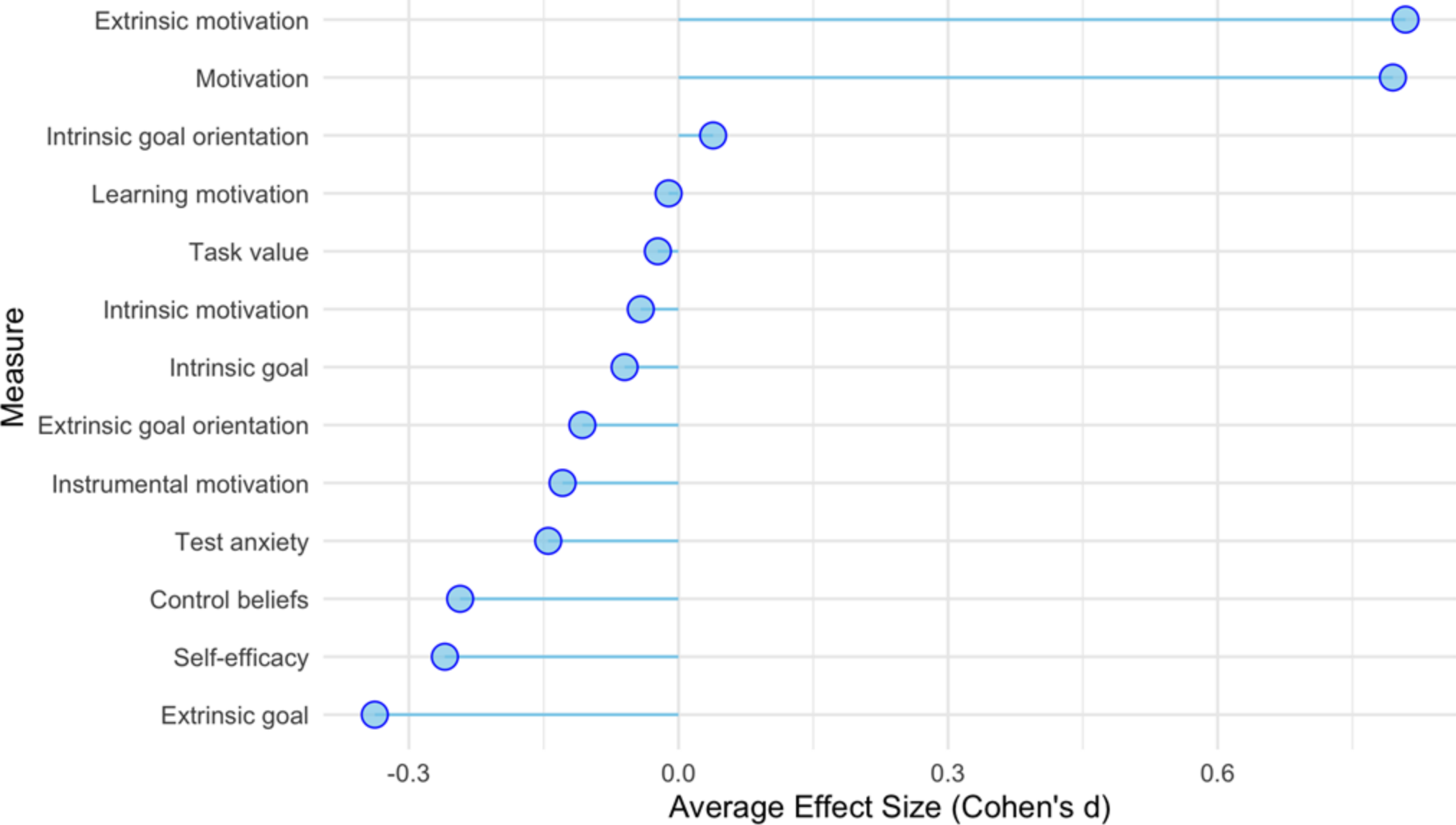
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Motivation: Effect size versus sample size
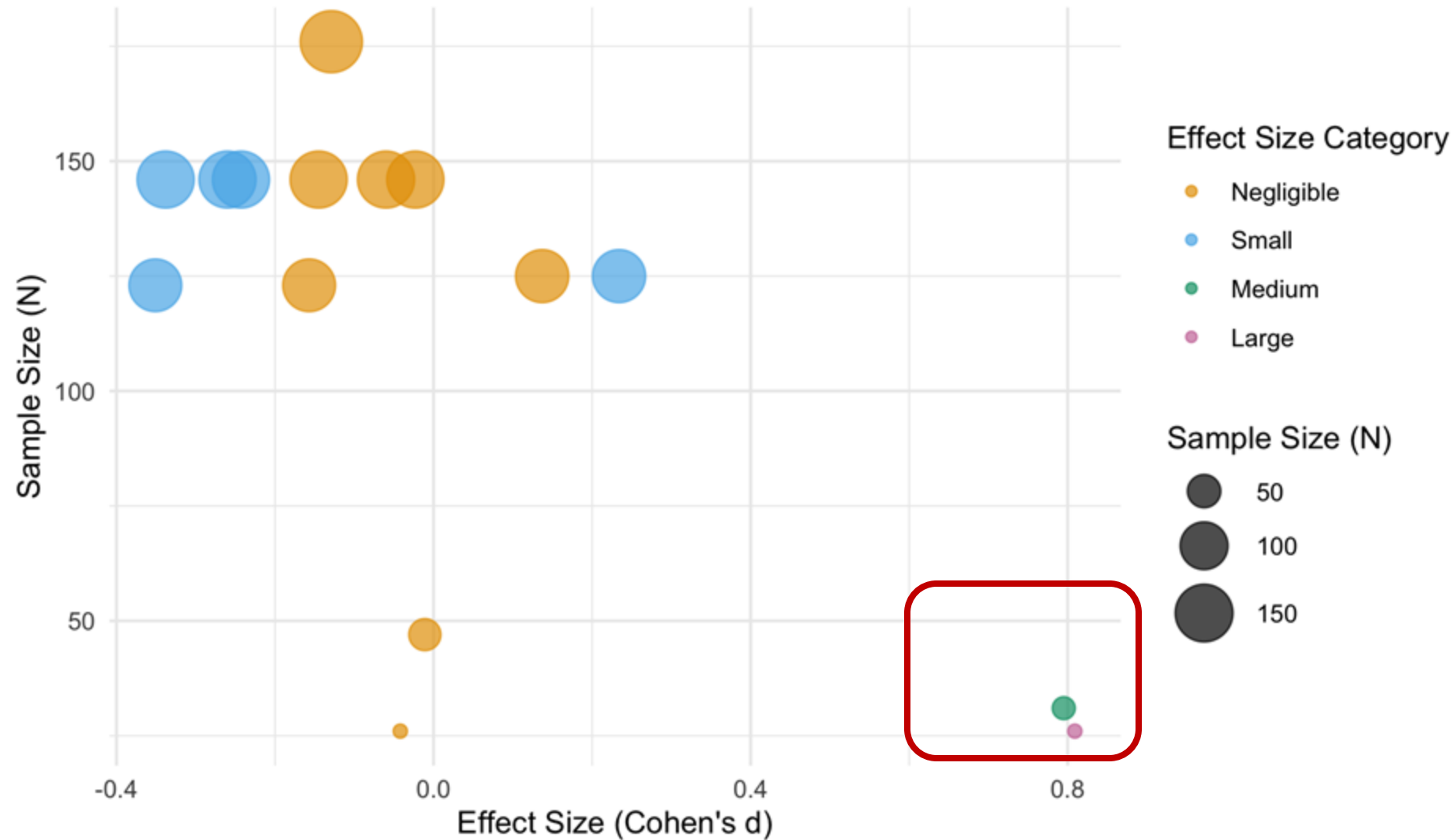
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Attitude

Attitude: Effect size descriptives (count of vote)
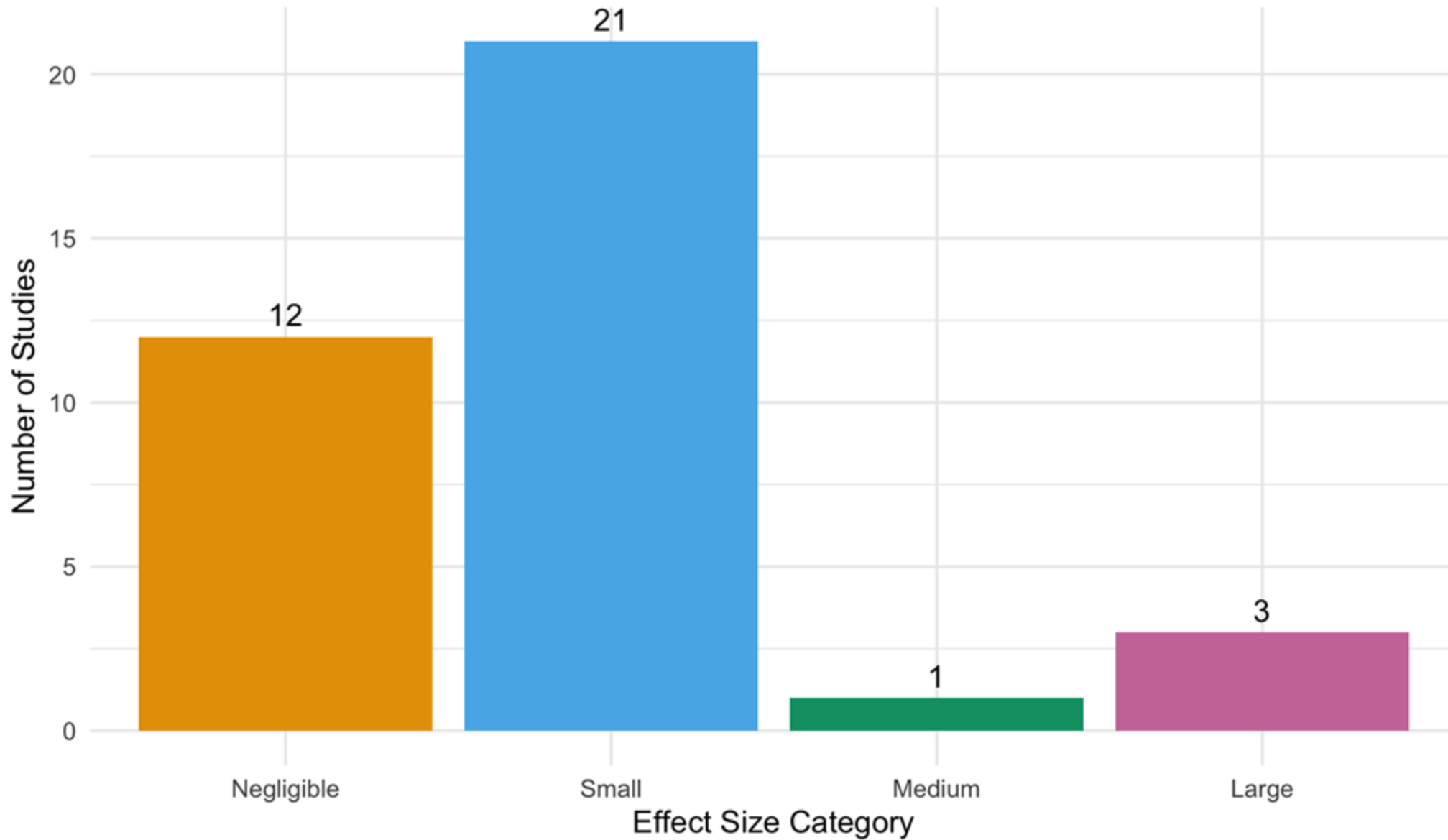
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Attitude: Effect size type
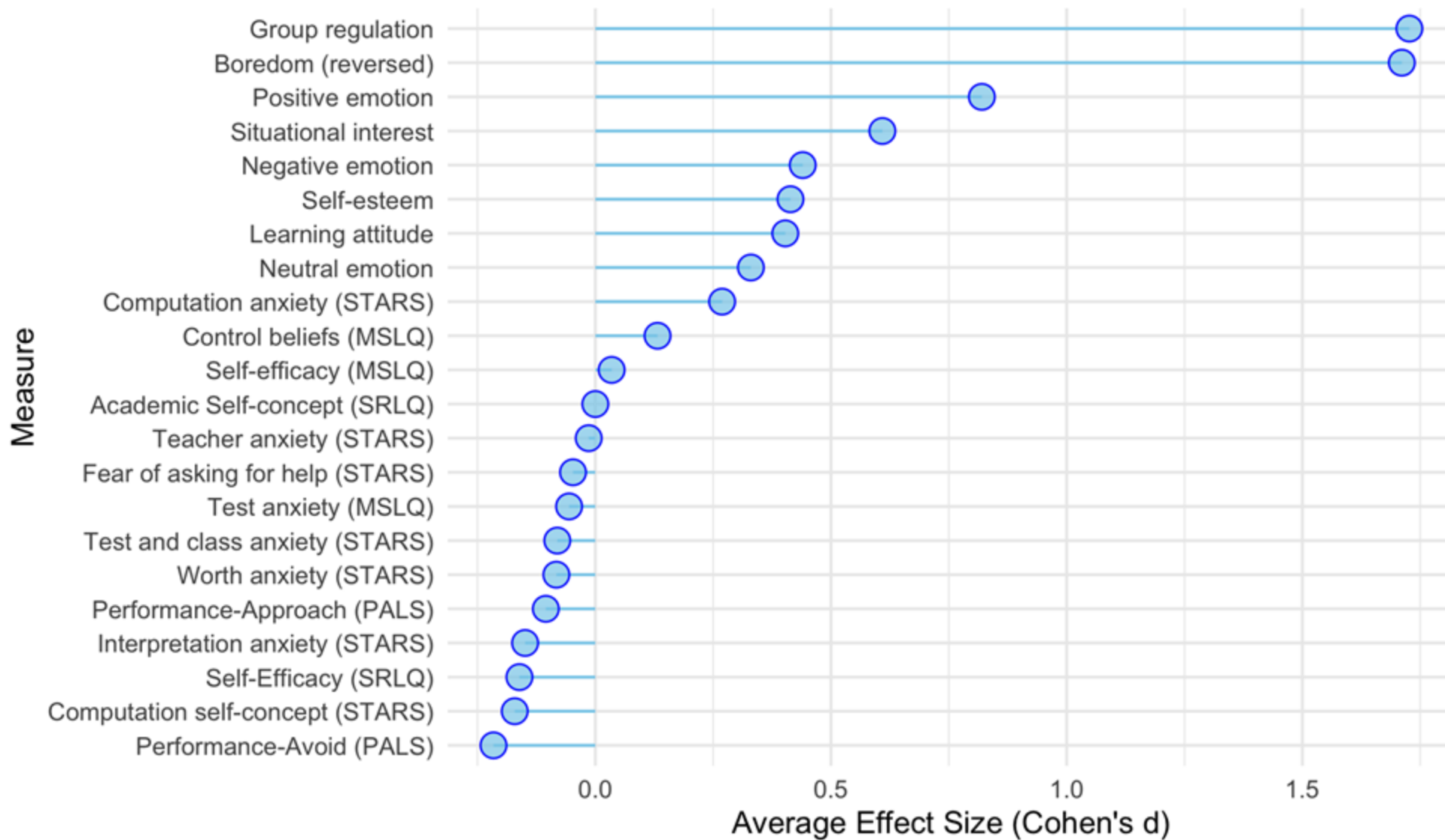
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Participation: Effect size versus sample size
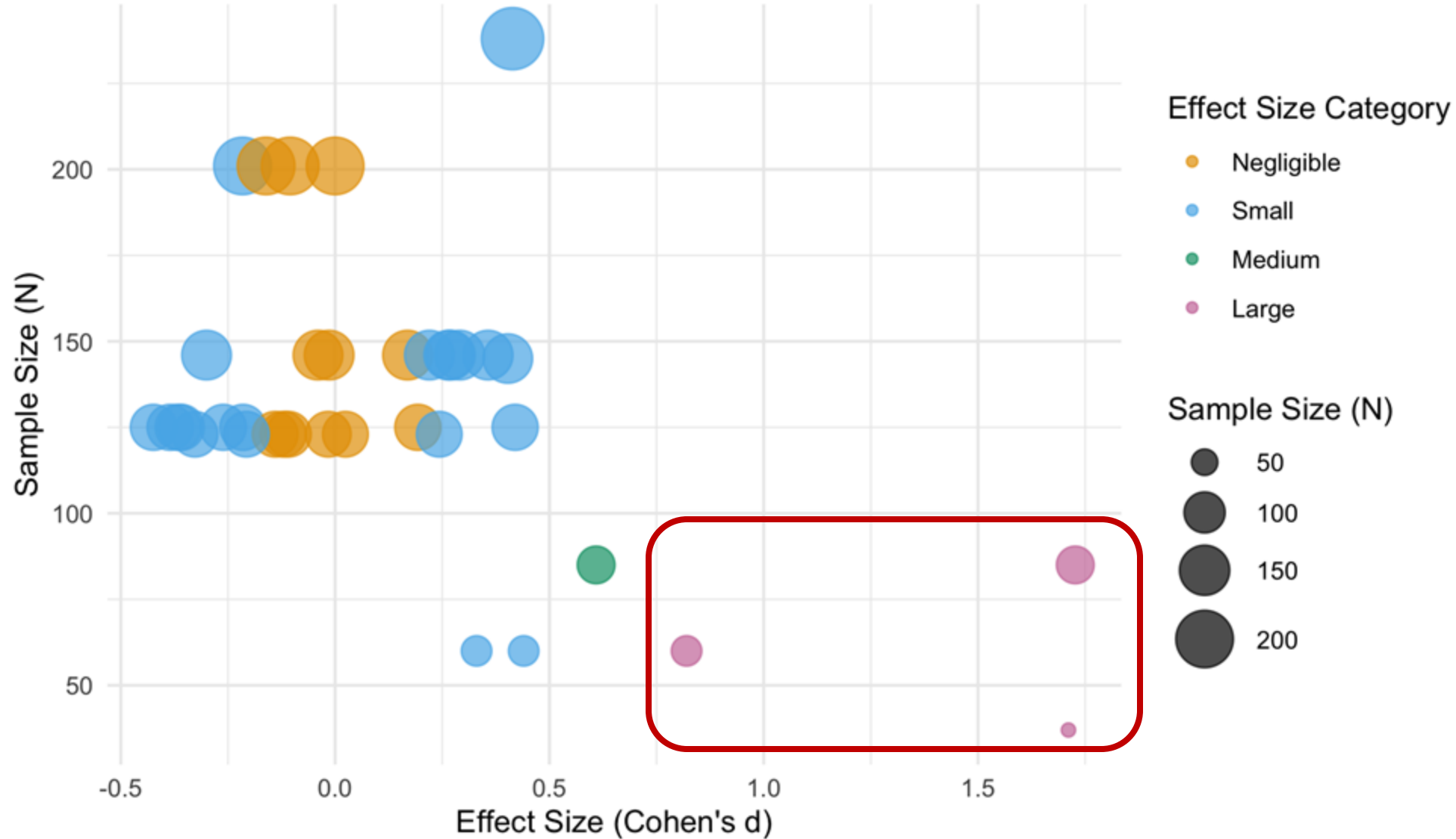
Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Attitudes: Correlation - Effect Size vs. Sample Size
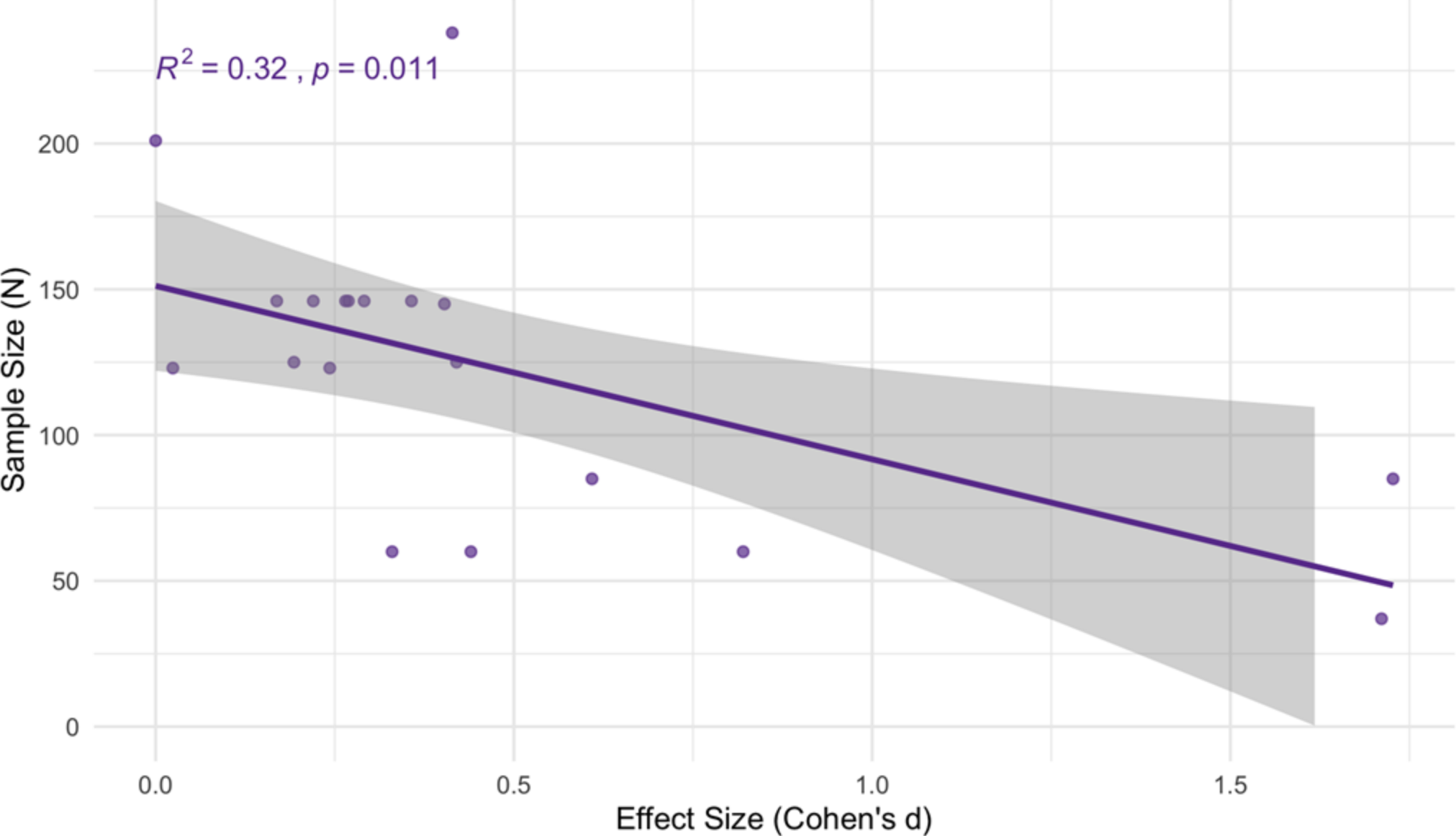
$R^2 = 0.32$, $p = 0.011$

Please note that a study may be represented more than once for demonstration purposes according to the number of outcome they evaluated.

Discussion

Performance

- As we currently stand, evidence is lacking that LADs has helped improve performance.

- Of course, **absence of evidence is not evidence of absence**. Yet, our study included no single article with well-powered controlled sample that assessed LAD with properly randomised sides.

- Randomization and control aside, most studies **reported small or negligible effect size,** with **confounders** that make it impossible to draw credible conclusions

Methodological issues

- LADs have either been combined with another type of intervention (and thus had an obvious confounding) or were assessed using a non-controlled design

- Another pattern compared dashboard users to non-users. A comparison that essentially measures the difference between two activity levels (not a comparison between a control and an experimental group).

- In controlled studies, many compared those who has the opportunity to use it (access) versus those who did not, regardless of whether everyone in the access group actively used it or not.

On Engagement, Attitude & Participation

- Slightly better results were reported about engagement, yet, with confounders that are hard to reconcile attributed to LADs.

- Motivation and attitude improvements were generally modest, and at times mixed. Yet again suffered the same drawbacks of lack of rigorous evaluation of two groups with confounders and randomizations, etc.

One step Back

One step Back

We are **<u>very excited</u>** about the opportunities for awareness, reflection, sensemaking, and impact that such dashboards provide and, above all, about the potential to improve learning, that is, to get better at getting better.

Impact remains especially **hard to demonstrate in** evaluation studies (Verbet 2013)

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. American Behavioral Scientist, 57(10), 1500-1509.

Previous research

- A considerable amount of exploratory work and small proof-of-concept studies, which very often do not reach the stage of being used (and **evaluated**) in authentic settings (Schwendimann 2016).

- Most dash-board **evaluations** focus on assessing the tool's usability and the impact on the behavioural competence. The effects on the cognitive and emotional levels received very little attention overall (Jivet 2018).

Schwendimann, Beat A., et al. "Perceiving learning at a glance: A systematic literature review of learning dashboard research." *IEEE transactions on learning technologies* 10.1 (2016): 30-41.

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In *Data Driven Approaches in Digital Education: 12th European Conference on Technology Enhanced Learning, EC-TEL 2017, Tallinn, Estonia, September 12–15, 2017, Proceedings 12* (pp. 82-96). Springer International Publishing.

Previous research

- Existing LADs are rarely grounded in learning theory, cannot be suggested to support metacognition, do not offer any information about effective learning tactics and strategies, and have significant limitations in how their **evaluation** is conducted and reported (W. Matcha 2019).

- Mostly are prototype and few are in early pilots with paucity of evidence on their **effectiveness** to affect learner outcomes (Susnjak 2022).

Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: a tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, *19*(1), 12.
Matcha, W., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. IEEE transactions on learning technologies, 13(2), 226-245.

Maybe there is no impact at all

Are there meta-analysis

● To the best of our knowledge, and the search we conducted, **not a single meta-analysis in any field exists**, let-alone proved that dashboards in their own right can, or has or will improve performance.

Review

Dashboards for improving patient care: Review of the literature

Dawn Dowding [a] [b] 👤 ✉, Rebecca Randell [c], Peter Gardner [d], Geraldine Fitzpatrick [e], Patricia Dykes [f], Jesus Favela [g], Susan Hamer [h], Zac Whitewood-Moores [i], Nicholas Hardiker [j], Elizabeth Borycki [k], Leanne Currie [l]

Eleven studies were included on CINAHL, Medline, Embase, Cochrane Library, PsychInfo, Sciencedirect and ACM Digital Library. A citation search and a hand search of relevant papers were also conducted.

Dowding, D., Randell, R., Gardner, P., Fitzpatrick, G., Dykes, P., Favela, J., ... & Currie, L. (2015). Dashboards for improving patient care: review of the literature. *International journal of medical informatics*, *84*(2), 87-100.

The results: Neither conclusive nor reliable

The authors described marked heterogeneity in the design of dashboards and users targeted and settings and concluded.

Although overall the majority of studies in this review indicated that the introduction of dashboards had a positive effect on outcomes and care processes (such as documentation of care processes, improved communication and access to information), there are a number of limitations with the study designs utilized to evaluate dashboards. With the exception of one study in the review which was rated as high quality, **the majority of studies had some element of potential bias**, with **5** studies being of low quality, meaning that **any significant results should be treated with caution**.

Dowding, D., Randell, R., Gardner, P., Fitzpatrick, G., Dykes, P., Favela, J., ... & Currie, L. (**2015**). Dashboards for improving patient care: review of the literature. *International journal of medical informatics*, *84*(2), 87-100.

JAMIA

JOURNAL ARTICLE   EDITOR'S CHOICE   FEATURED

Effectiveness of clinical dashboards as audit and feedback or clinical decision support tools on medication use and test ordering: a systematic review of randomized controlled trials

Charis Xuan Xie ✉, Qiuzhe Chen, Cesar A Hincapié, Léonie Hofstetter, Chris G Maher, Gustavo C Machado

Eleven randomized trials were included from 7 databases. Eight trials evaluated clinical dashboards as standalone interventions and provided conflicting evidence on changes in antibiotic prescribing and no effects on statin prescribing compared to usual care.

Xie, C. X. et al. (**2022**). Effectiveness of clinical dashboards as audit and feedback or clinical decision support tools on medication use and test ordering: a systematic review of randomized controlled trials. *Journal of the American Medical Informatics Association*, *29*(10), 1773-1785.

Clinical decision support tools

There is limited evidence that dashboards integrated into electronic medical record systems and used as feedback or decision support tools may be associated with improvements in medication use and test ordering.

Xie, C. X. et al. (2022). Effectiveness of clinical dashboards as audit and feedback or clinical decision support tools on medication use and test ordering: a systematic review of randomized controlled trials. *Journal of the American Medical Informatics Association, 29*(10), 1773-1785.

Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis

Katie-Jane Brickwood[1]; Greig Watson[1]; Jane O'Brien[1]; Andrew D Williams[1]

Cochrane Controlled Register of Trials, MEDLINE, PubMed, Scopus, Web of Science, Cumulative Index of Nursing and Allied Health Literature, SPORTDiscus, and Health Technology Assessments.

Controlled trials of adults comparing the use of a consumer-based wearable activity tracker with other non-activity tracker–based interventions were included.

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR mHealth and uHealth*, 7(4), e11819.

Results

Utilizing a consumer-based wearable activity tracker as either the primary component of an intervention or as part of a broader physical activity intervention **has the potential to increase physical activity participation**.

As the effects of physical activity interventions are often short term, the inclusion of a consumer-based wearable activity tracker may provide an effective tool to assist health professionals to provide ongoing monitoring and support.

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR mHealth and uHealth*, 7(4), e11819.

Personal dashboards: Trackers Increase Physical Participation

1. Daily Steps – all studies

Study or Subgroup	Intervention Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Ashe (2015) [29]	7,606	3,917	12	4,593	663	7	0.8%	0.91 [-0.08, 1.89]
Ashton (2017) [30] (1)	1,588.2	2,608.1	26	575.4	2,591.4	24	2.4%	0.38 [-0.18, 0.94]
Brakenridge (2016) [32] (2)	440.8	2,278.4	66	-157.8	2,015.2	87	7.2%	0.28 [-0.04, 0.60]
Cadmus-Bertram (2015) [21]	6,695	2,708	25	6,188	2,423	26	2.5%	0.19 [-0.36, 0.74]
Finkelstein (2016) [34] (3)	570	2,562.2	197	-480	2,516.4	201	18.1%	0.41 [0.21, 0.61]
Finkelstein (2016) [34] (4)	-130	2,601.3	203	-480	2,516.4	201	18.7%	0.14 [-0.06, 0.33]
Finkelstein (2016) [34] (5)	-300	2,575.2	199	-480	2,516.4	201	18.5%	0.07 [-0.13, 0.27]
Lyons (2017) [40]	6,193.7	3,183.5	20	4,586	2,476.1	20	1.9%	0.55 [-0.08, 1.19]
Martin (2015) [41] (6)	408	2,701	32	-616	2,385	16	2.1%	0.39 [-0.22, 0.99]
Melton (2016) [42]	10,674	2,703	21	10,870	2,426	36	2.6%	-0.08 [-0.61, 0.46]
Poirier (2016) [44]	5,411	2,277	107	4,751	1,834	110	10.3%	0.32 [0.05, 0.59]
Skrepnik (2017) [22]	5,537.9	3,101.2	107	4,825.4	2,425.1	101	9.9%	0.25 [-0.02, 0.53]
Thorndike (2014) [50] (7)	7,886	3,622	50	7,600	3,492	49	4.9%	0.08 [-0.31, 0.47]
Total (95% CI)			1065			1079	100.0%	0.23 [0.15, 0.32]

Heterogeneity: Tau² = 0.00; Chi² = 12.38, df = 12 (P = 0.42); I² = 3%
Test for overall effect: Z = 5.21 (P < 0.00001)

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR mHealth and uHealth*, 7(4), e11819.
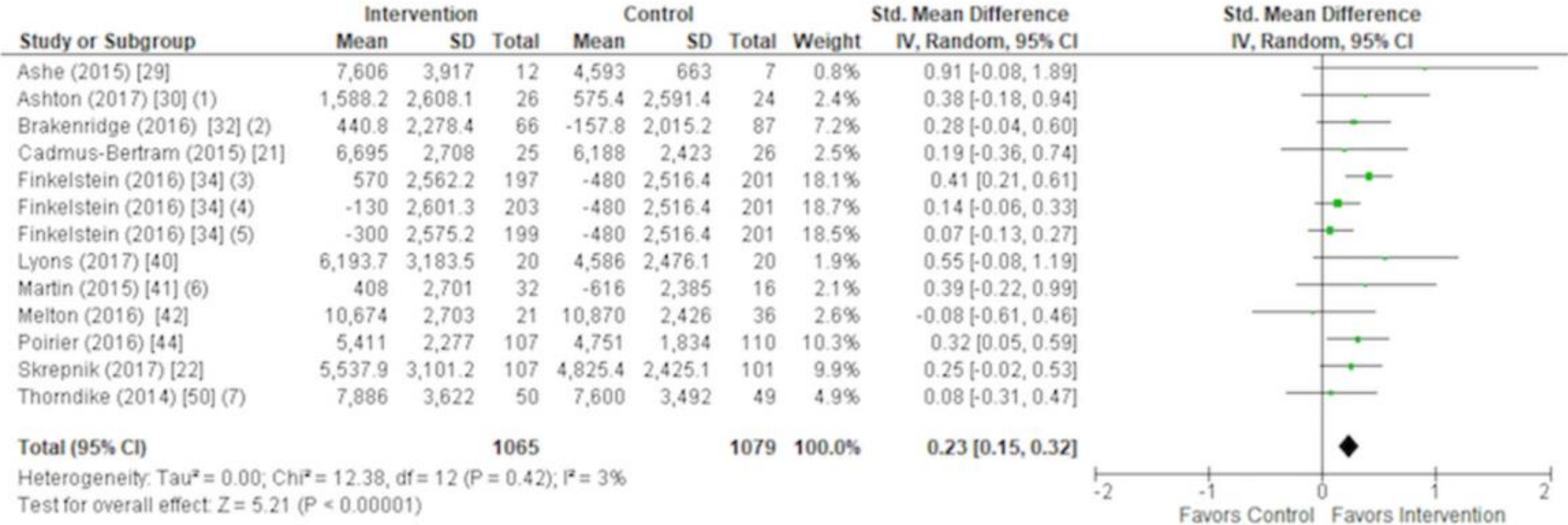
Maybe it is time to ask **for whom it works** rather that did it work?

Aggregating averages don't reflect impact

Most research is typically conducted by calculating the average scores across a sample of students to establish the "**state of affairs**". The average reflects the central tendency where data tend to cluster.
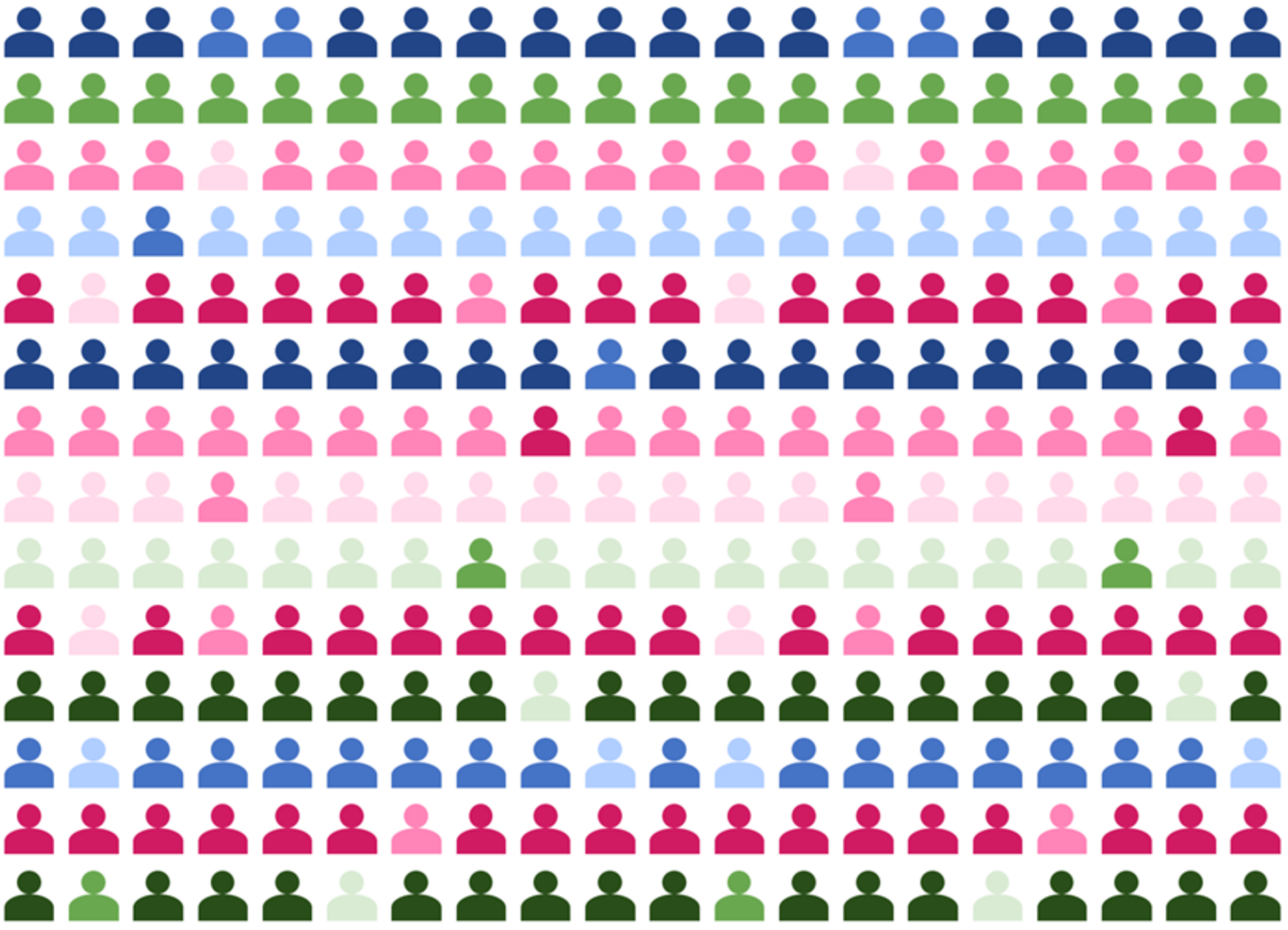
Check for updates

Behavioural science is unlikely to change the world without a heterogeneity revolution

Christopher J. Bryan [1] ✉, Elizabeth Tipton [2] ✉ and David S. Yeager [1] ✉

Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Beh[...] tho[...]
behaviour, 5(8), 980-989.
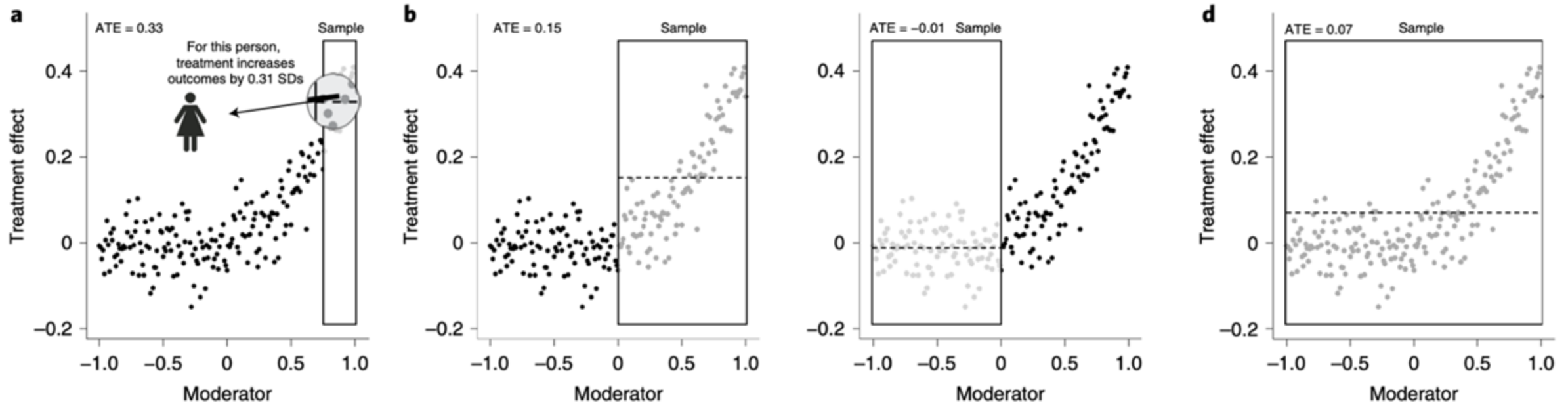
Aggregating averages mixes results

$V_{\text{Wilcoxon}} = 2422.00$, $p = 0.85$, $\hat{r}_{\text{biserial}}^{\text{rank}} = -0.02$, $\text{CI}_{95\%}$ [-0.24, 0.20], $n_{\text{pairs}} = 100$

$\hat{\mu}_{\text{mean}} = 4.80$

$\hat{\mu}_{\text{mean}} = 4.83$

Post intervention

Baseline
(n = 100)

PostIntervention
(n = 100)
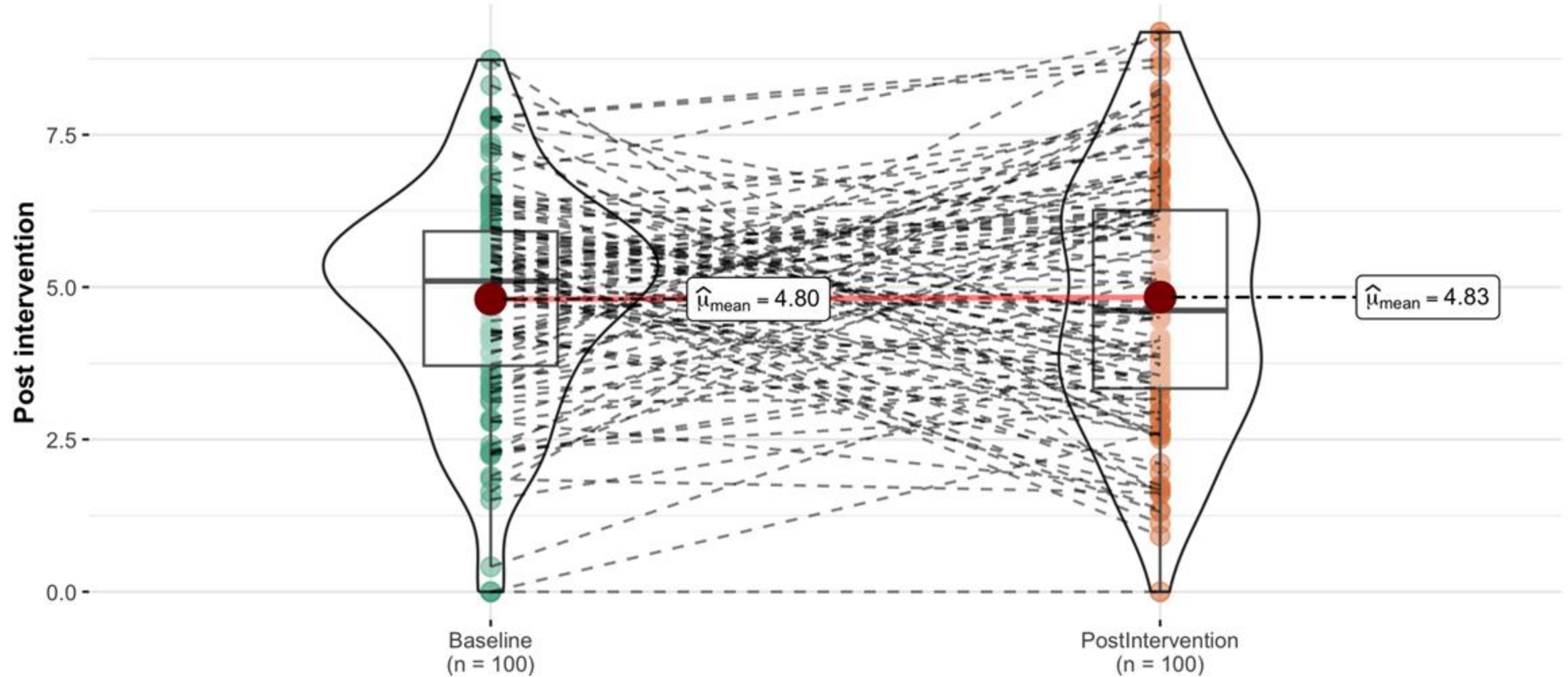
We need tools that capture the heterogeneous effects

Improvers

$V_{\text{Wilcoxon}} = 0.00$, $p = 2.48e\text{-}09$, $\hat{r}_{\text{biserial}}^{\text{rank}} = -1.00$, $CI_{95\%}$ [-1.00, -1.00], $n_{\text{pairs}} = 47$

$\hat{\mu}_{\text{mean}} = 5.75$

$\hat{\mu}_{\text{mean}} = 3.73$
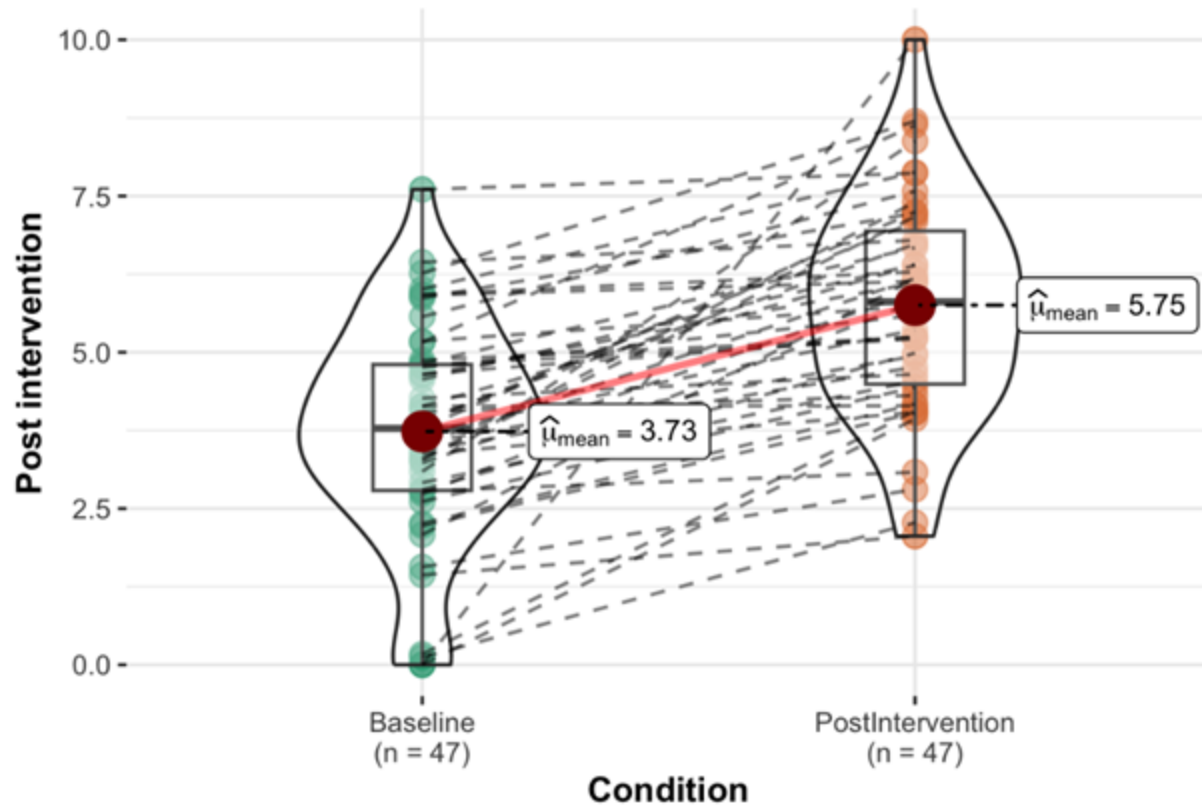
Worseners

$V_{\text{Wilcoxon}} = 1431.00$, $p = 2.46e\text{-}10$, $\hat{r}_{\text{biserial}}^{\text{rank}} = 1.00$, $CI_{95\%}$ [1.00, 1.00], $n_{\text{pairs}} = 5$
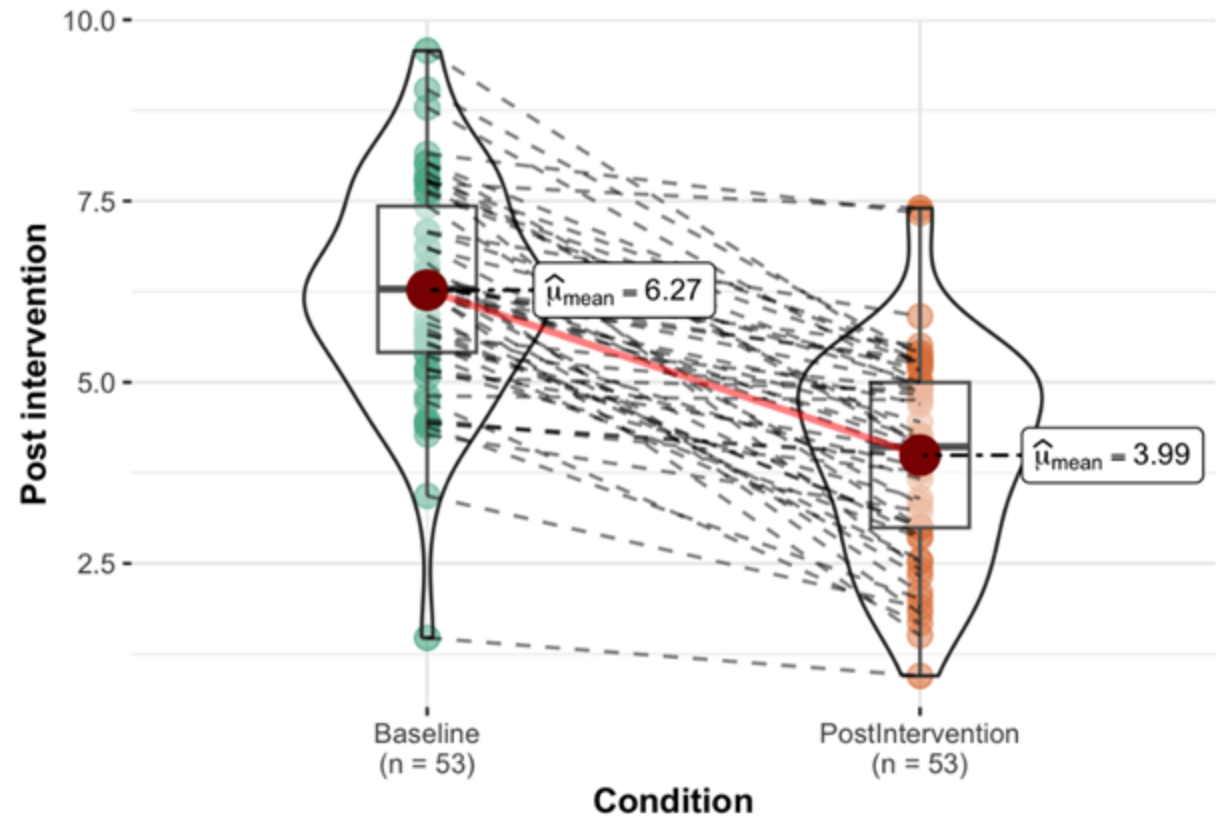
$\hat{\mu}_{\text{mean}} = 6.27$

$\hat{\mu}_{\text{mean}} = 3.99$

nature

Brain–phenotype models fail for individuals who defy sample stereotypes

Abigail S. Greene[1,2 ✉], Xilin Shen[3], Stephanie Noble[3], Corey Horien[1,2], C. Alice Hahn[3], Jagriti Arora[3], Fuyuze Tokoglu[3], Marisa N. Spann[4], Carmen I. Carrión[5], Daniel S. Barron[6,7,8,9], Gerard Sanacora[7], Vinod H. Srihari[7], Scott W. Woods[7], Dustin Scheinost[1,3,10,11,12] & R. Todd Constable[1,3,10,13 ✉]

Greene, A. S., Shen, X., Noble, S., Horien, C., Hahn, C. A., Arora, J., … & Constable, R. T. (2022). Brain–phenotype models fail for individuals who defy sample stereotypes. Nature, 1-10.
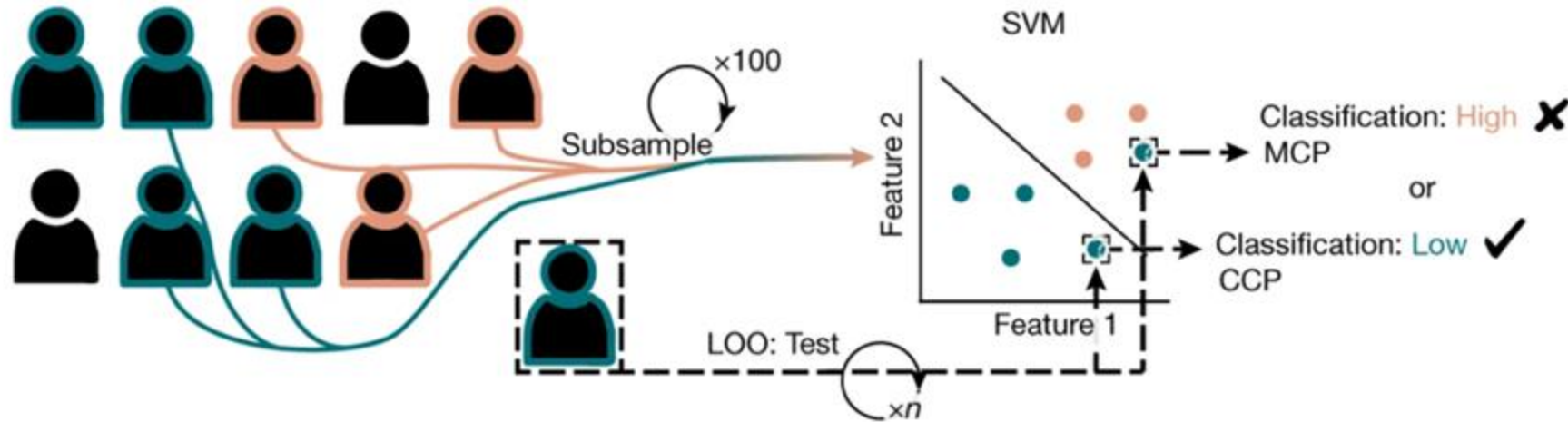
nature

Brain–phenotype models fail for individuals who defy sample stereotypes

Models fail when applied to people who defy these stereotypes.

Models systematically fail because they aren't predicting unitary cognitive constructs. They're predicting measures of these constructs intertwined with sociodemographic and clinical covariates—stereotypes.
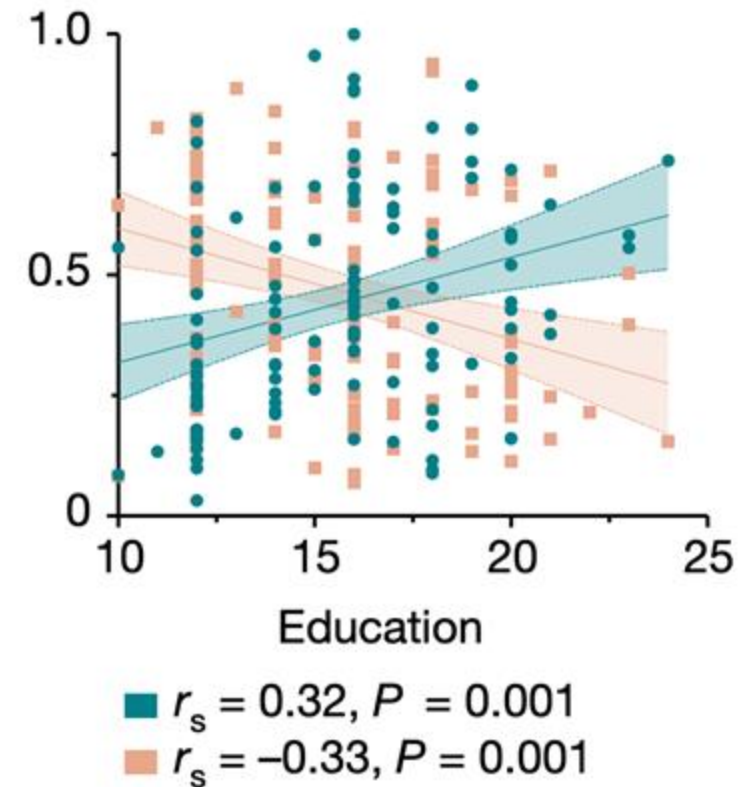
$r_s = 0.32, P = 0.001$
$r_s = -0.33, P = 0.001$

Greene, A. S., Shen, X., Noble, S., Horien, C., Hahn, C. A., Arora, J., ... & Constable, R. T. (2022). Brain–phenotype models fail for individuals who defy sample stereotypes. Nature, 1-10.

Idiographic artificial intelligence to explain students' self-regulation: Toward precision education

Mohammed Saqr [a], Rongxin Cheng [b], Sonsoles López-Pernas [a], Emorie D Beck [b]

We developed N=1 machine learning models for
each and every person using EMA data as well other
data to predict 3 outcome

- Their Effort in doing their studies
- Motivation
- Metacognition

Saqr, M., Cheng, R., López-Pernas, S., & Beck, E. D. (2024). Idiographic artificial intelligence to explain students' self-regulation: Toward precision education. *Learning and Individual Differences*, *114*, 102499.

Predicting effort in studying

	M (SD)	N	Range
Random Foest	0.45 (0.26)	17	(0.07, 0.95)
Elastic Net	0.49 (0.25)	16	(0.02, 0.9)
BISCWIT	0.49 (0.23)	17	(0.13, 0.9)

Saqr, M., Cheng, R., López-Pernas, S., & Beck, E. D. (2024). Idiographic artificial intelligence to explain students' self-regulation: Toward precision education. *Learning and Individual Differences*, *114*, 102499.
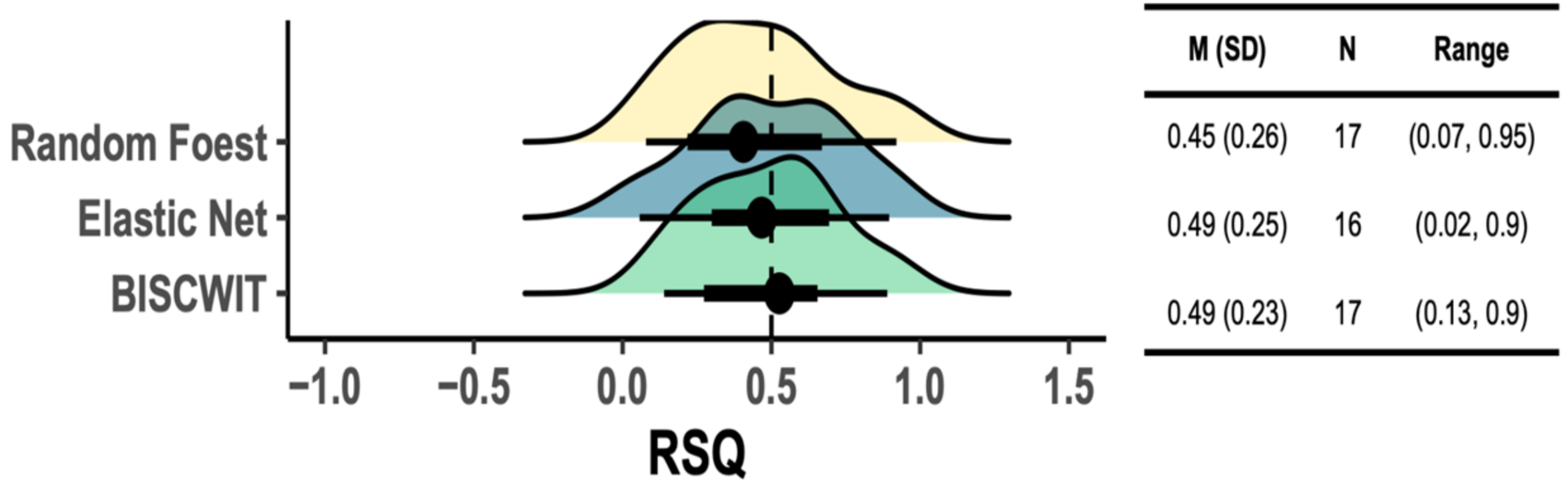
Top 5 predictors for everyone in the sample
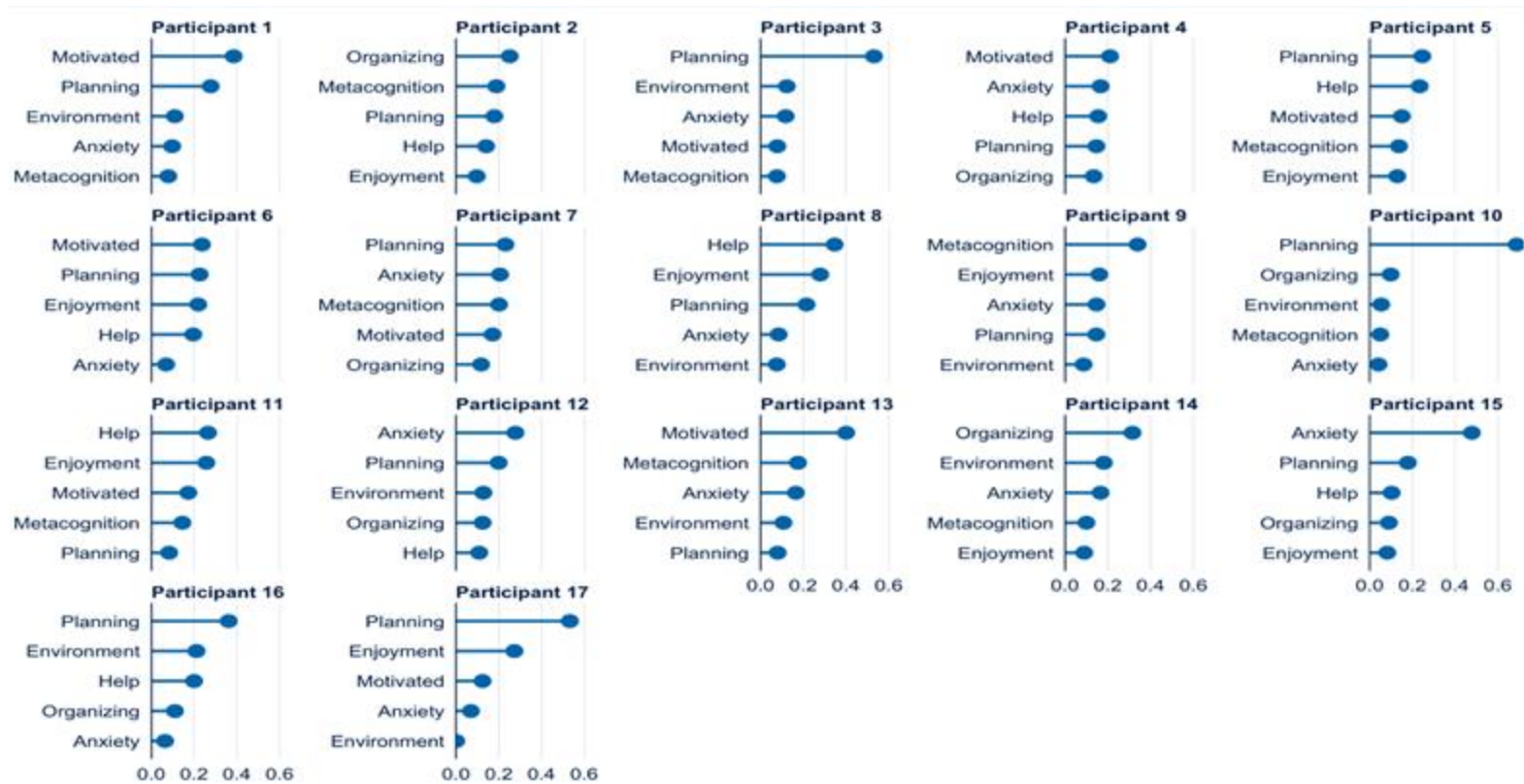
Surprisingly, not a single student shared the same order of the top predictors for any outcome with another student, Nor any one shared the order of the average model.

Saqr, M., Cheng, R., López-Pernas, S., & Beck, E. D. (2024). Idiographic artificial intelligence to explain students' self-regulation: Toward precision education. *Learning and Individual Differences*, *114*, 102499.
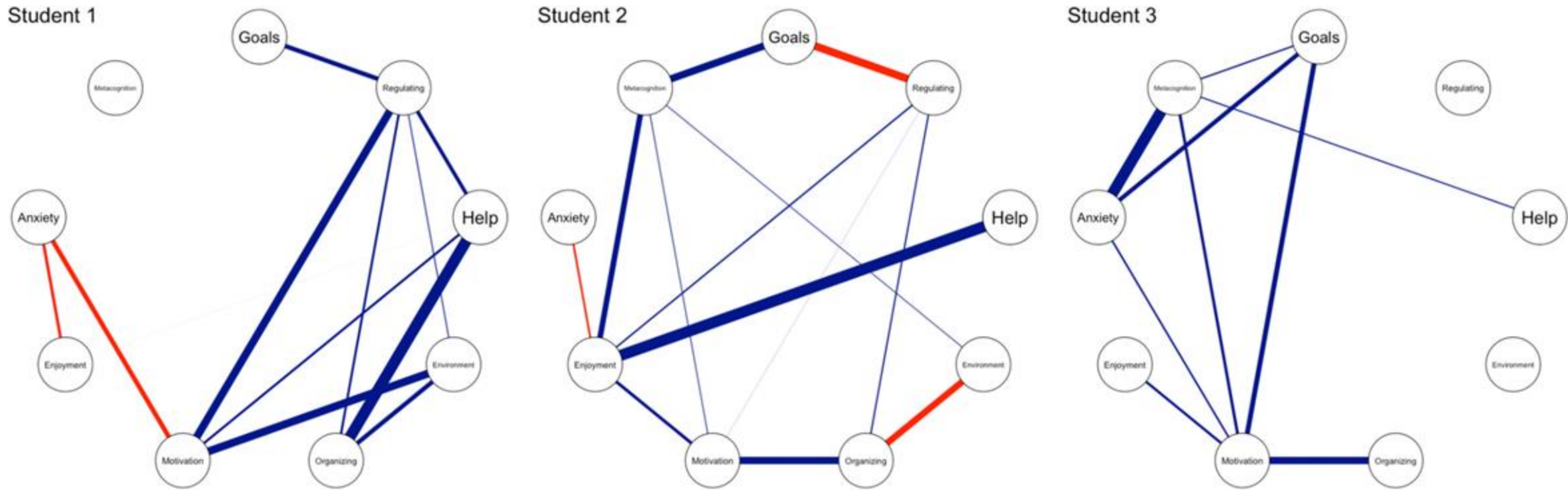
Behavioural science is unlikely to change the world without a heterogeneity revolution

Christopher J. Bryan ⓘ[1]✉, Elizabeth Tipton ⓘ[2]✉ and David S. Yeager ⓘ[1]✉

- We need to acknowledge that most effects are heterogeneous.
- So, the variation in effect estimates across studies that defines the replication crisis is to be expected as long as heterogeneous effects are studied without a systematic approach to sampling and moderation.

Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. Nature human behaviour, 5(8), 980-989.

cemmap

centre for microdata methods and practice

Generic machine learning inference on heterogenous treatment effects in randomized experiments

Victor Chernozhukov
Mert Demirer
Esther Duflo
Ivan Fernandez-Val

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018). *Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India* (No. w24678). National Bureau of Economic Research.

Science

Current Issue   First release papers   Archive   About ⌄   Submit manuscript

HOME > SCIENCE > VOL. 380, NO. 6644 > WHERE AND WITH WHOM DOES A BRIEF SOCIAL-BELONGING INTERVENTION PROMOTE PROGRESS IN COLLEGE?

🔓 | **RESEARCH ARTICLE** | SOCIAL BELONGING

Where and with whom does a brief social-belonging intervention promote progress in college?

GREGORY M. WALTON iD , MARY C. MURPHY iD , CHRISTINE LOGEL iD , DAVID S. YEAGER iD , J. PARKER GOYER iD , SHANNON T. BRADY iD , KATHERINE T. U. EMERSON iD , DAVID PAUNESKU iD , OMID FOTUHI iD , [...], AND NATASHA KROL   +27 authors   Authors Info & Affiliations

Walton, G. M., Murphy, M. C., Logel, C., Yeager, D. S., Goyer, J. P., Brady, S. T., ... & Krol, N. (2023). Where and with whom does a brief social-belonging intervention promote progress in college?. *Science*, *380*(6644), 499-505.

Interventions works for certain groups of people

A randomized controlled trial to systematically explain and understand these heterogeneous effects in a brief online intervention across 22 universities and colleges (see the Perspective by Bowman). The intervention was designed to remedy students' concerns about belonging through a reading-and-writing activity that emphasized how worries about fitting in, struggling in class, and feeling homesick during the college transition are common and improve over time. They found that the intervention **improved retention and persistence in school, particularly among historically underrepresented students, when the school context offered students opportunities to belong**.

Walton, G. M., Murphy, M. C., Logel, C., Yeager, D. S., Goyer, J. P., Brady, S. T., ... & Krol, N. (2023). Where and with whom does a brief social-belonging intervention promote progress in college?. *Science*, *380*(6644), 499-505.

Rogers Kaliisa, Kamila Misiejuk,
Sonsoles López-Pernas, Mohammad Khalil, **Mohammed Saqr**

Thanks