SCANTRON.

Enter the 4th Dimension: Outcome-Based Evaluation

Tobias Raupach, MME
University Medical Centre
Göttingen, Germany
University College London, UK

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

UCL

SMART STARTS HERE

The four dimensions of teaching quality

Global ratings are suboptimal indicators of teaching quality.

There are four dimensions of teaching quality, and each of these can be assessed separately.

Specific tools targeting structural and procedural aspects of teaching as well as individual teacher performance are available.

Is there an adequate outcome measure?

Schiekirka et al. BMC Med Educ 2015; 15: 30
Gibson et al. Acad Med 2008; 83:787-793
Roff, Med Teach. 2005; 27: 322-325; Marsh, Br J Psychol 1982; 52: 77-95

What is the (desired) outcome?

Multidimensional Outcome Considerations in Assessing the Efficacy of Medical Educational Programs

Phyllis Blumberg
Teaching and Learning Center
University of the Sciences in Philadelphia
Philadelphia, Pennsylvania, USA

- Educational outcomes (students: learning style)
- Clinical career outcomes (students: competencies)
- Environmental outcomes (university: culture)

Teach Learn Med 2003; 15: 210-214

Exam performance as a ‚surrogate' for learning outcome?

In order to be valid, exam results need to be objective and reliable.

Another prerequisite for the validity of exam results is complete coverage of the underlying construct.

In addition, exam format needs to be aligned to learning objectives and teaching formats.

Performance gain during a module/course can only be assessed if initial performance levels are taken into account.

Downing & Haladyna, Med Educ 2004; 38: 327-333
Kern et al.: Curriculum development for medical education – A six-step approach (1998)

Alternative ways to estimate learning outcome

Exploring the Learning Curve in Medical Education: Using Self-Assessment as a Measure of Learning

Britta M. Thompson and John C. Rogers
Acad Med. 2008;83(10 Suppl):S86–S88.

Validity of singular self-assessments?

Average correlation between self-assessments and objective ratings: r = 0.39 (range -0.05 to 0.82)

Influencing factors:

gender
student level
study design
type of learning objective
subject matter
design of the data collection tool
type of 'objective' data

Falchikov & Boud, Rev Educ Res 1989; 59: 395-430
Colthart et al. Med Teach 2008; 30: 124-145

From singular self-assessments to comparative self-assessments (CSA)

The validity of singular student self-assessments is limited owing to a number of confounding factors.

However, within a given individual, the ability to self-assess is relatively stable over time.

Thus, repeated/comparative student self-assessments might be used to estimate learning outcome.

Measuring learning outcome on the level of specific learning objectives requires these objectives to be clearly operationalised.

Ward et al. Adv Health Sci Educ Theory Pract 2002; 7: 63-80
Fitzgerald et al. Med Educ 2003; 37: 645-649

Alternative ways to estimate learning outcome

2011; 33: e446–e453 MEDICAL TEACHER

WEB PAPER

Towards outcome-based programme evaluation: Using student comparative self-assessments to determine teaching effectiveness

T. RAUPACH, C. MÜNSCHER, T. BEIßBARTH, G. BURCKHARDT & T. PUKROP
University Hospital Göttingen, Germany

Estimating Learning Outcomes From Pre- and Posttest Student Self-Assessments: A Longitudinal Study

Sarah Schiekirka, Deborah Reinhardt, Tim Beißbarth, PhD, Sven Anders, MD, MME, Tobias Pukrop, MD, and Tobias Raupach, MD, MME

Acad Med 2013; 88: 369-375

Data collection

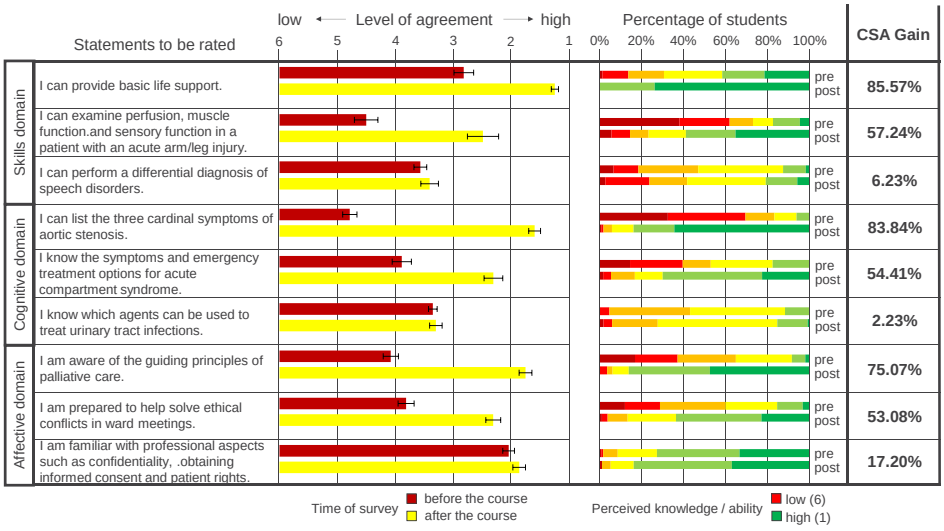
Using an online survey tool, students were invited to self-assess their individual performance levels at the beginning and at the end of each teaching module. Statements used for self-assessments addressed the following three domains:

- Factual knowledge
- Practical skills
- Affective learning objectives including professionalism

$$\text{CSA Gain}\,[\%] = (\mu_{pre} - \mu_{post}) \,/\, (\mu_{pre} - 1)$$

Raupach et al. Med Teach 2011; 33: e446-e453

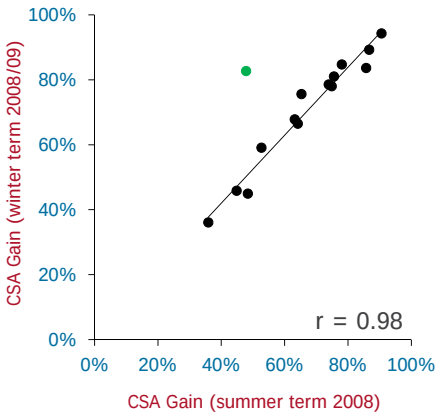
Data presentation

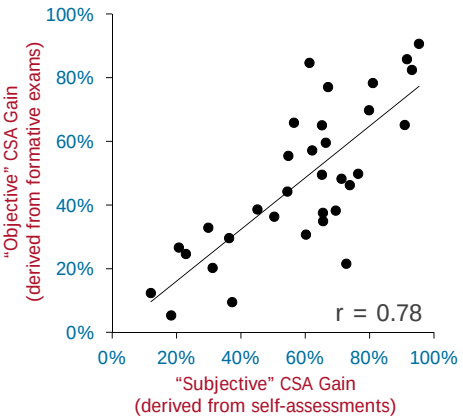


Raupach et al., Med Teach 2011

Reliability & Validity

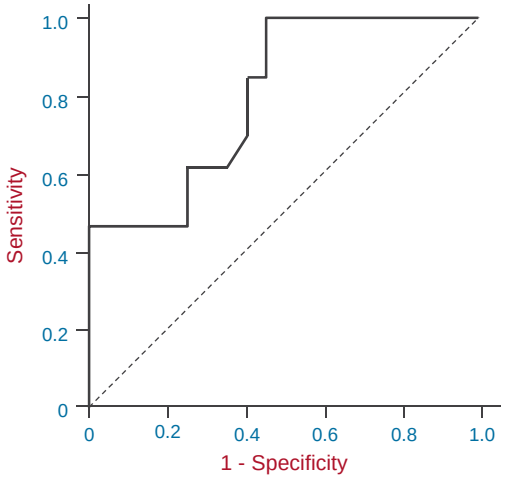
Each dot represents one specific learning objective.





Raupach et al. Med Teach 2011; 33: e446-e453; Schiekirka et al. Acad Med 2013; 88: 369-375

ROC analysis



- Cut-off: 54.7%
- Specificity: 57%
- Sensitivity: 100%
- PPV: 59%
- NPV: 100%

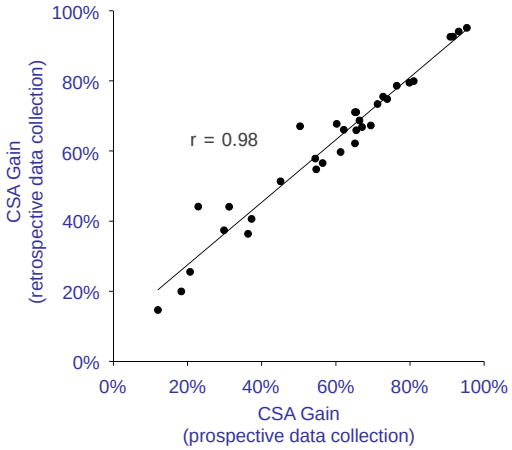
Schiekirka et al. Acad Med 2013; 88: 369-375

Practicability

RESEARCH ARTICLE Open Access

Assessment of two different types of bias affecting the results of outcome-based evaluation in undergraduate medical education

Sarah Schiekirka[1,2], Sven Anders[3] and Tobias Raupach[1,4*]



Schiekirka et al. BMC Med Educ 2014; 14: 149

A 'word' of caution

Annals of Anatomy 208 (2016) 222–227

Contents lists available at ScienceDirect

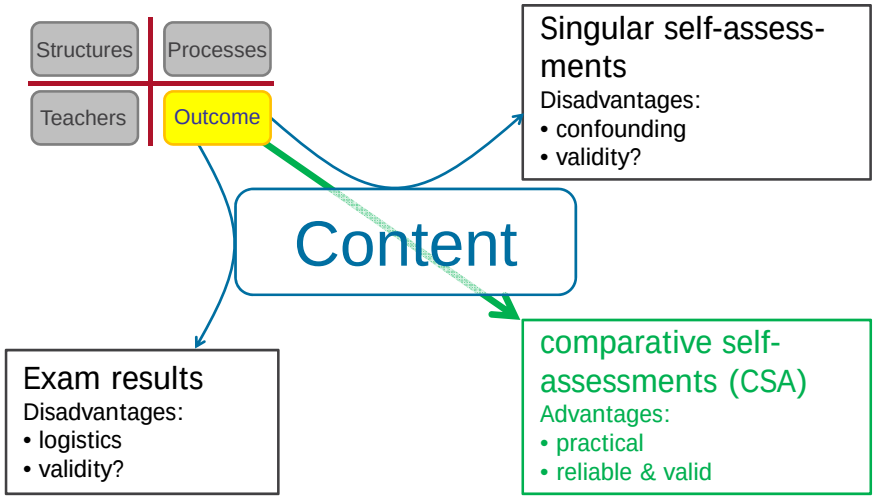
Annals of Anatomy

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

Influence of the wording of evaluation items on outcome-based evaluation results for large-group teaching in anatomy, biochemistry and legal medicine

Sven Anders[a,*], Katharina Pyka[a], Tjark Mueller[a], Nicole von Streinbuechel[b], Tobias Raupach[c,d]

Summary



Summary

Estimating student learning outcome from comparative student self-assessments…

…produces values ranging from -100% to +100%,

…takes initial performance levels into account,

…provides results for different domains of teaching (knowledge, skills, affective learning objectives),

…is robust against a number of potential confounders.

…helps to differentiate between learning objectives with favourable and suboptimal learning outcome, thus facilitating an increase in teaching quality.

Prof. Dr. Tobias Raupach, MME
Consultant Cardiologist
Teaching Co-ordinator
Clinic for Cardiology & Pneumology
University Medical Centre Göttingen
Phone: ++49-551-3963393
Email: raupach@med.uni-goettingen.de