



Working Together to Ensure Consistent Student Evaluation

Sarah M. Varekojis, PhD, RRT, FAARC

Objectives

- Discuss the purpose of ensuring inter-rater agreement for student evaluations.
- Discuss the role of preceptor training in aiding inter-rater agreement.
- Describe methods for improving inter-rater agreement.



So be honest – how do you feel about evaluations the students complete about you?

Student evaluation of instruction

While I know we all appreciate constructive feedback we can use to improve our instruction and our courses, in the end I think everyone can feel...

Judged


Uneasy

Uncertain


Apprehensive

Found
deficient or
inadequate

Anxiety about
how they will
rate you



“Important to quantify the reliability of scoring when performance assessment scores are used for consequential, high-stakes decisions”



Sage Reference: Encyclopedia of Research Design
(<https://us.sagepub.com/en-us/nam>)

Standard 3.07

- The program must develop and implement processes that **REDUCE INCONSISTENCY** among individuals who perform clinical evaluations.

Evidence of Compliance

Standard 3.07

- Evidence of Compliance:
 - Documentation of an **INTER-RATER RELIABILITY PLAN** that includes a description of **EVALUATOR TRAINING** and records of training participation by clinical evaluators;
 - Documentation of review and analysis of clinical evaluations completed by individuals performing clinical evaluations;
 - Documentation of implementation of an **ACTION PLAN TO REDUCE INCONSISTENCY** when variability is identified.

Reliability

Consistency

Repeatability

Reproducibility

Dependability

Trustworthiness

Producing similar results

Quantifying reliability

#1

Correlation coefficient

Establishes consistency

Reliable instruments have a high positive correlation

Quantifying reliability



Percent agreement

Establishes consensus

Reliable instruments have a high percent agreement

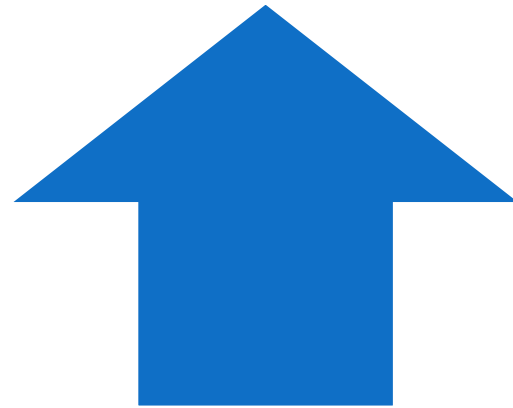
Percent agreement



Test score =
correct / total



Agreement =
observed / possible



Percent agreement

“An adequate level of agreement is generally considered to be 70%.”

Sage Reference: Encyclopedia of Research Design
(<https://us.sagepub.com/en-us/nam>)

Percent agreement

Excellent

- 80% – 100%

Good

- 60% – 79 %

Moderate

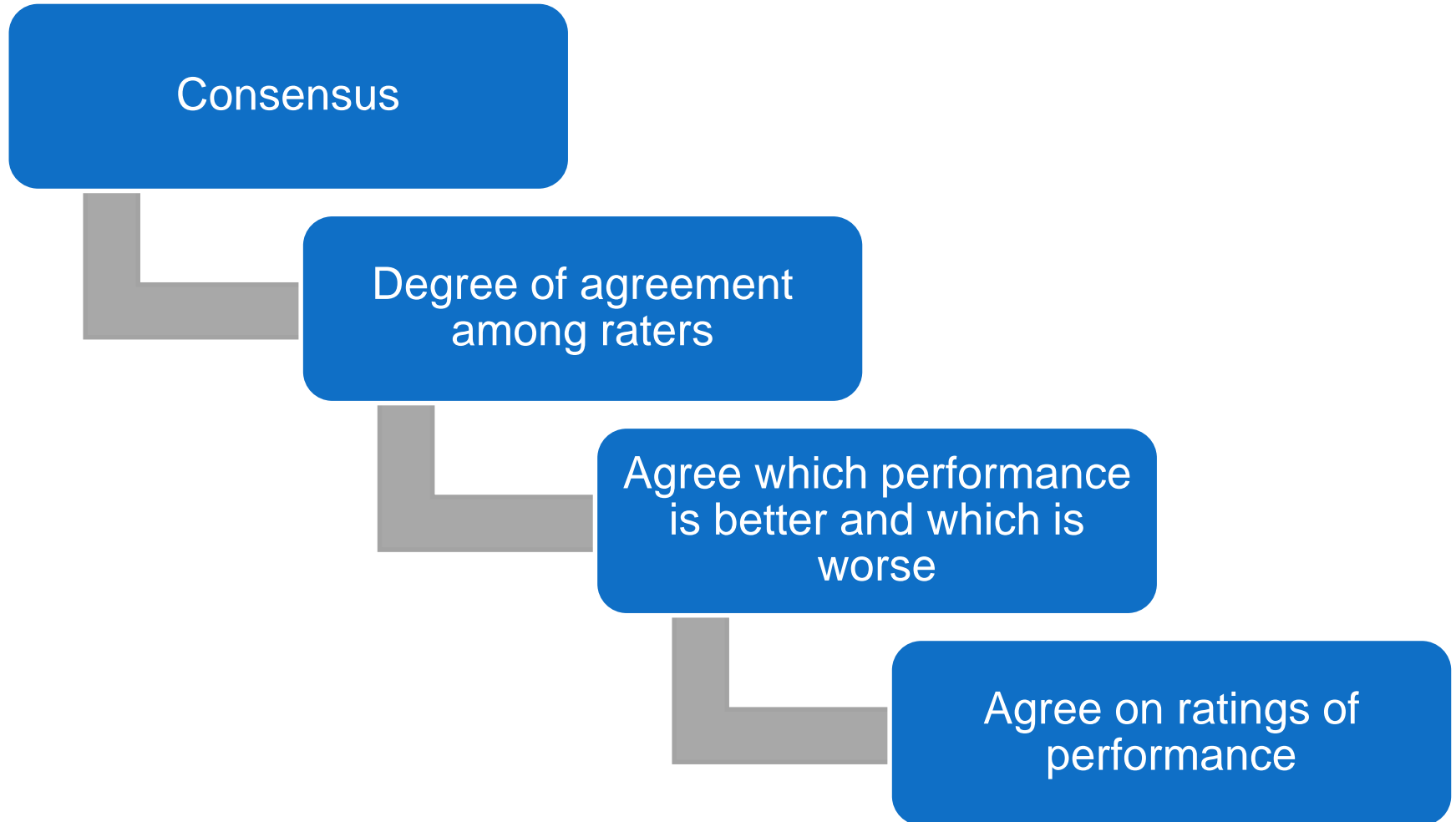
- 40% – 59%

Slight

- 20% – 39%

Marques JF, McCall C. The application of inter-rater reliability as a solidification instrument in a phenomenological study. *The Qualitative Report* 2005;10(3);439-62.

Inter-rater reliability



Inter-rater reliability

Similar results no matter where and when the assessment is carried out



Similar results independent of who does the scoring

Improving inter-rater agreement

Preceptor training

Establish roles and responsibilities.

Discuss student and preceptor expectations.

Improve skills in teaching, coaching, mentoring, leadership, communication, and **evaluation**.

Creating an optimal learning environment, conflict resolution, communication skills and teamwork.

Process for **assessment**, planning and implementation of clinical learning experiences.

Expectations for **formative and summative evaluation** processes.

Preceptor training

Understanding rating scale

- Definitions and meanings for each value of the scale.

Reduce scoring or rater bias

- Score using your criteria and expectations, not theirs.

Improving inter-rater agreement

Structured and standardized forms or checklists

Explicit rubric descriptors, aka scoring guidelines

- Clearly distinguish between different levels of performance



THE OHIO STATE UNIVERSITY

utstanding
atisfactory
nsatisfactory



Improving inter-rater agreement

Operationally define terms

- Clear, concise, detailed definition of a measure
- How you will measure in a specific instance

Operational definition

Weight:

- Number displayed on digital screen
- How much a needle deflects
- Number of coins needed to balance



Improving inter-rater agreement

- Avoid:
 - Vague terms
 - Often vs. “every day” or “once a day”
 - Grammatical ambiguity
 - “I saw her duck.”



Improving inter-rater agreement

- Avoid:
 - Excessive complexity
 - “One difference between television news reporting and the coverage provided by newspapers is the time factor between the actual happening of an event and the time it takes to be reported. The problem is that instantaneous coverage is physically impossible for newspapers.”
 - “Television news reporting differs from that of newspapers in that television can provide instantaneous coverage of events as they happen, and newspapers cannot.”

Improving inter-rater agreement

- Avoid:
 - Double-barreled questions



Operational definition

What is Outstanding?

No prompting, no errors, answered questions, mastery obvious.



What is Satisfactory?

Occasional prompting (___ or less prompts), no critical errors, minor intervention required



What is Unsatisfactory?

Frequent prompting (___ or more prompts) or intervention required, harmful errors, unable to perform

Improving inter-rater agreement

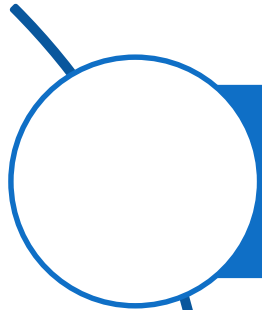
More discrete items within dimensions –
distinct and separate items



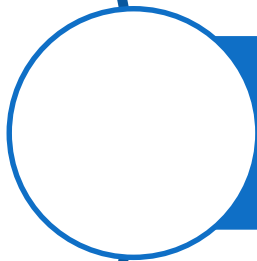
List of 22
professional
behaviors

Overall
professional
behaviors
rating

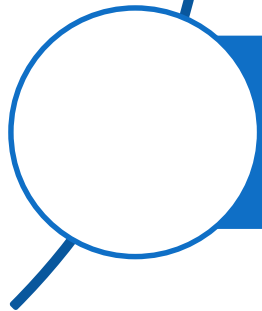
Improving inter-rater agreement



Sufficient range or variation in scores across dimensions



Raters struggle to assign scores when there are fewer score categories, wanting to create their own



Suggest 5-7 scores

Improving inter-rater agreement



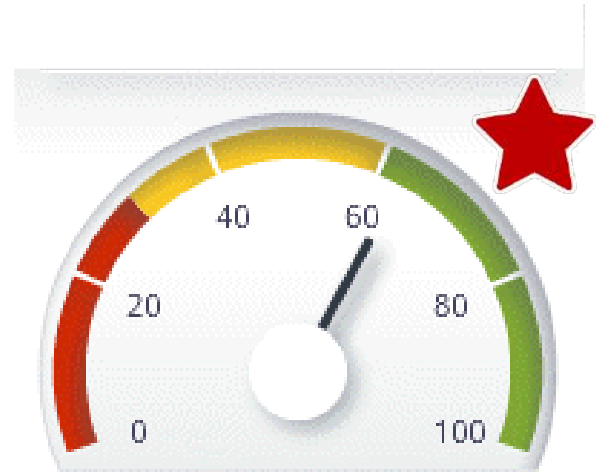
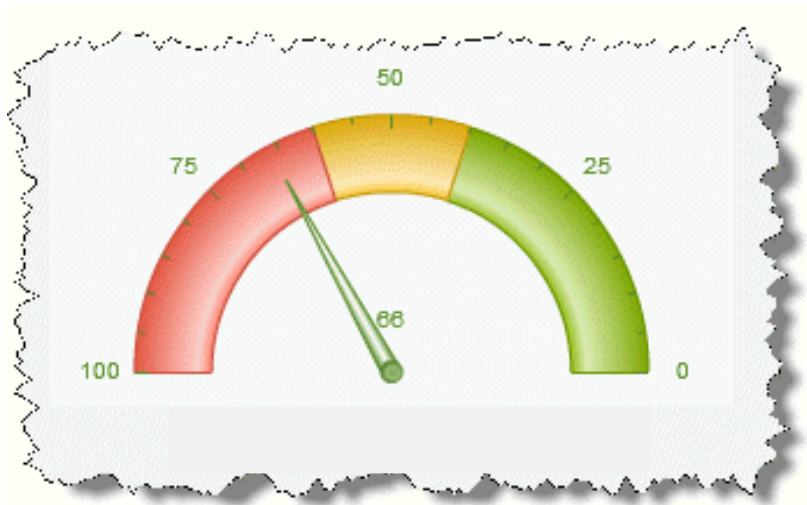
10 9 8 7 6 5 4 3 2 1
 Strongly Agree Agree Neutral Disagree Strongly Disagree

Items stated in negative statements:

1 2 3 4 5 6 7 8 9 10
 Strongly Agree Agree Neutral Disagree Strongly Disagree

Five-Point Scale (on Importance)	Seven-Point Scale (on Agreement)
1 = Very Important	1 = Strongly Agree
2 = Important	2 = Agree
3 = Moderately Important	3 = Slightly Agree
4 = Of Little Importance	4 = Undecided
5 = Unimportant	5 = Slightly Disagree
	6 = Disagree
	7 = Strongly Disagree

Improving inter-rater agreement



Improving inter-rater agreement

- Focus on the critical steps
 - Making a PB&J sandwich



Directions:

Step 1: Getting the Bread

- Take out two slices of bread from the bag.
- Place the two slices, lying down, on the plate.

Step 2: Adding the Peanut Butter

- Get the jar of Peanut Butter and the knife.
- Open the jar of Peanut Butter.
- Use the knife to scoop out a small amount of Peanut Butter from the jar.
- Transfer the Peanut Butter that is now on the knife to one slice of bread and spread the scoop of Peanut Butter evenly on this slice with the knife.
 - If more Peanut Butter is desired repeat taking the Peanut Butter from the jar and spreading it on the same slice of bread.
- Put the knife down off to the side.
- Put down the slice of bread with the Peanut Butter on it back on the plate, next to the other slice of bread, with the Peanut Butter side up.

Step 3: Cleaning the Knife

- Clean the knife using water from the sink and wipe it clean and dry with the napkin or paper towel.
 - CAUTION!! The knife is sharp so be extra careful when cleaning it to avoid injury.

Step 4: Adding the Jelly

- Get the jar of Jelly and the knife.
- Open the jar of Jelly.
- Use the knife to scoop out a small amount of Jelly from the jar.
- Transfer the Jelly that is now on the knife to the slice of bread that has nothing on it and spread the scoop of Jelly evenly on this slice with the knife.
 - If more Jelly is desired repeat taking the Jelly from the jar and spreading it on the same slice of bread.
- Put the knife down off to the side.
- Put down the slice of bread with the Jelly on it back on the plate, next to the other slice of bread, with the Jelly side up.

Step 5: Putting the Sandwich Together

- Pick up both slices of bread by the sides that are not covered in Peanut Butter or Jelly.
- Place the two slices of bread together with the Peanut Butter and Jelly facing each other and touching.
 - Fit the two slices so that the shapes match and all the sides are aligned.
- Place the sandwich back down on the plate.

Step 6: Enjoy the Sandwich

- Eat the sandwich and enjoy.
 - Cut sandwich in half with the knife if desired.





Improving inter-rater agreement

Field test

Get feedback on

- Rubric descriptors
- Operational definitions
- Number of items
- Number of categories

Limitations to improving inter-rater agreement

Cannot
control

- Motivation
- Interest
- Observational skills
- Desire to adhere

Resources

- <http://srmo.sagepub.com/>

The screenshot shows the SAGE Research Methods website. At the top, a green banner reads "SAGE researchmethods" with a logo, followed by "Learn more >>" and "SAGE Research Methods is migrating to a new and improved platform". Below this is a navigation bar with links: "About", "What's New?", "Tutorials", "Help", "Migration FAQ", "User Resources", "Member Sign in", "Not a Member?", "Sign up for save and search tools.", and "Ohio State University". The main content area features a search bar with "Advanced Search" and a "Browse" section with buttons for "Content", "Authors & Editors", "Methodologies", "Methods Lists", and "Videos". The main content is divided into four columns: "Learning Research Methods", "Doing a Research Project", "Supervising a Research Project", and "Teaching Research Methods". Each column has a green header with a right-pointing arrow and a light blue body with descriptive text.

SAGE researchmethods Learn more >>
SAGE Research Methods is migrating to a new and improved platform

About What's New? Tutorials Help Migration FAQ User Resources Member Sign in Not a Member? Sign up for save and search tools. Ohio State University

SAGE researchmethods

Search: Advanced Search

Browse Content Authors & Editors Methodologies Methods Lists Videos

Learning research methods? The comprehensive introductory content in SAGE Research Methods covers all the basics.

Doing a research project? Our content will help you design and conduct your own research.

Supervising a project? SAGE Research Methods can support you as you support others!

Teaching research methods? SAGE Research Methods has all the tips and tools you need to inspire your students.

Learning Research Methods

Doing a Research Project

Supervising a Research Project

Teaching Research Methods

Questions and Answers

Sarah Varekojis

Sarah Varekojis, PhD, RRT, FAARC
Varekojis.16@osu.edu