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”
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.

www.coarc.com
Reliability

Consistency

- Repeatability
- Reproducibility
- Dependability
- Trustworthiness
- Producing similar results
Quantifying reliability

Correlation coefficient

- Establishes consistency
- Reliable instruments have a high positive correlation

www.coarc.com
Quantifying reliability

Percent agreement

- Establishes consensus
- Reliable instruments have a high percent agreement

www.coarc.com
Percent agreement

Test score = \frac{\text{correct}}{\text{total}}

Agreement = \frac{\text{observed}}{\text{possible}}
Percent agreement

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

### Percent agreement

<table>
<thead>
<tr>
<th>Category</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>80% – 100%</td>
</tr>
<tr>
<td>Good</td>
<td>60% – 79%</td>
</tr>
<tr>
<td>Moderate</td>
<td>40% – 59%</td>
</tr>
<tr>
<td>Slight</td>
<td>20% – 39%</td>
</tr>
</tbody>
</table>

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

Consensus

Degree of agreement among raters

Agree which performance is better and which is worse

Agree on ratings of performance
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.

www.coarc.com
Improving inter-rater agreement

- Structured and standardized forms or checklists
- Explicit rubric descriptors, aka scoring guidelines
  - Clearly distinguish between different levels of performance
outstanding
satisfactory
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

www.coarc.com
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

<table>
<thead>
<tr>
<th>Scale</th>
<th>Description</th>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Strongly Agree</td>
<td>1</td>
<td>Strongly Agree</td>
</tr>
<tr>
<td>9</td>
<td>Agree</td>
<td>2</td>
<td>Agree</td>
</tr>
<tr>
<td>8</td>
<td>Neutral</td>
<td>3</td>
<td>Slightly Agree</td>
</tr>
<tr>
<td>7</td>
<td>Disagree</td>
<td>4</td>
<td>Undecided</td>
</tr>
<tr>
<td>6</td>
<td>Strongly Disagree</td>
<td>5</td>
<td>Slightly Disagree</td>
</tr>
<tr>
<td>5</td>
<td>Agree</td>
<td>6</td>
<td>Disagree</td>
</tr>
<tr>
<td>4</td>
<td>Neutral</td>
<td>7</td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>3</td>
<td>Disagree</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Strongly Disagree</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Items stated in negative statements:

<table>
<thead>
<tr>
<th>Scale</th>
<th>Description</th>
<th>Scale</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Strongly Agree</td>
<td>1</td>
<td>Strongly Agree</td>
</tr>
<tr>
<td>9</td>
<td>Agree</td>
<td>2</td>
<td>Agree</td>
</tr>
<tr>
<td>8</td>
<td>Neutral</td>
<td>3</td>
<td>Slightly Agree</td>
</tr>
<tr>
<td>7</td>
<td>Disagree</td>
<td>4</td>
<td>Undecided</td>
</tr>
<tr>
<td>6</td>
<td>Strongly Disagree</td>
<td>5</td>
<td>Slightly Disagree</td>
</tr>
<tr>
<td>5</td>
<td>Agree</td>
<td>6</td>
<td>Disagree</td>
</tr>
<tr>
<td>4</td>
<td>Neutral</td>
<td>7</td>
<td>Strongly Disagree</td>
</tr>
<tr>
<td>3</td>
<td>Disagree</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Strongly Disagree</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
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.
SMUCKER'S Uncrustables
PEANUT BUTTER & STRAWBERRY JAM SANDWICH
4 Sandwiches
Net WT. 8 OZ (223g)
Improving inter-rater agreement

Get feedback on

• Rubric descriptors
• Operational definitions
• Number of items
• Number of categories
Limitations to improving inter-rater agreement

- Motivation
- Interest
- Observational skills
- Desire to adhere

Cannot control
Resources

- [http://srmo.sagepub.com/](http://srmo.sagepub.com/)

---

**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.
Questions and Answers
Sarah Varekojis, PhD, RRT, FAARC
Varekojis.16@osu.edu