Assessing the Reliability of a Problem-Solving Rubric when using Multiple Raters

T. Ryan Duckett, Uchenna Asogwa, Matthew W. Liberatore,

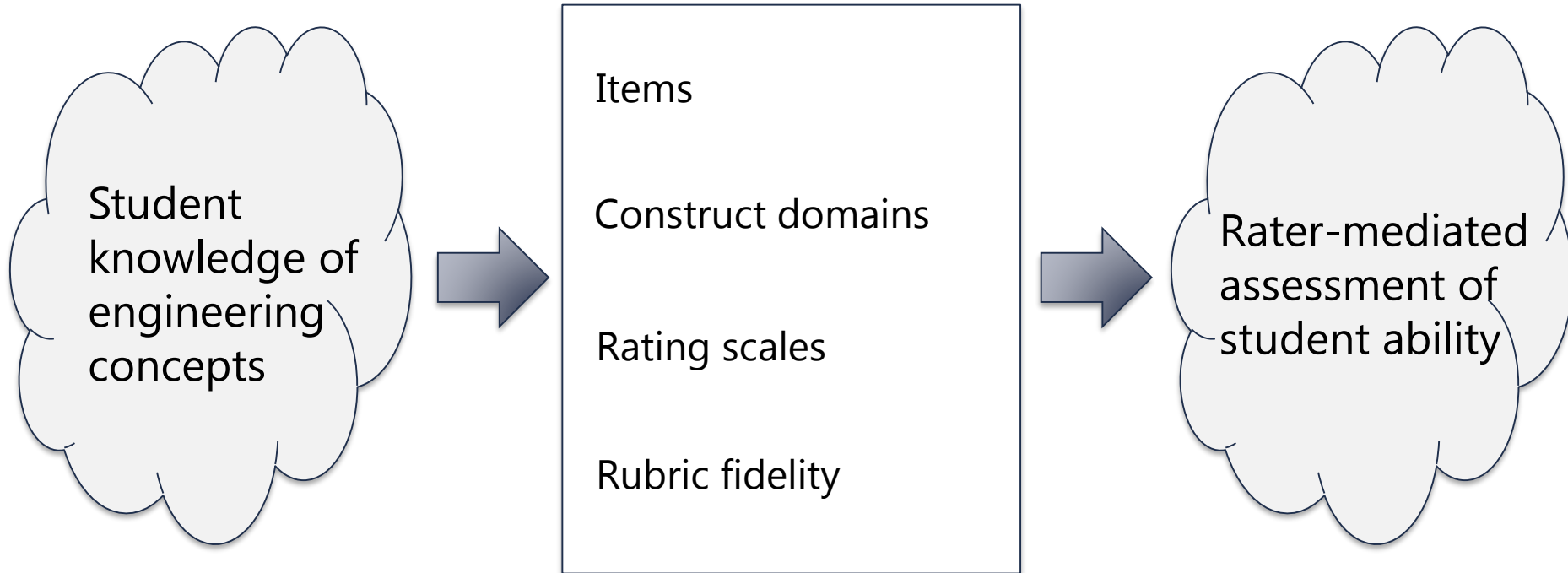Gale Mentzer, & Amanda Malefyt

timothy.duckett@utoledo.edu

http://www.utoledo.edu/engineering/chemical-engineering/liberatore/

June 2019

Conceptual framework

Student knowledge of engineering concepts → Items, Construct domains, Rating scales, Rubric fidelity → Rater-mediated assessment of student ability

Study design

	Data Collected	Full rating plan	Iterative, inter-rater reliability study
# Raters:	N/A	4	5
Participant N:	113 (39% female) 2 MW schools Undergrad MEB	70	20
Problem Type:	Traditional Innovative	Traditional Innovative	Traditional

Typical homework problem

Exercise 3.3.2: Methanol reactor.

About

The synthesis of methanol from carbon monoxide and hydrogen includes nitrogen as an inert carrier gas. The feed to the reactor is 425 mol/min with 102 mol/min CO, 0.143 mol fraction of N_2, and the balance H_2. In the reactor, a single-pass conversion of CO is 75.8%. The reactor effluent goes to a condenser for further separation.

(a) Draw and label a process flow diagram and number the streams.

Solution ∧

Step 1. The process flow diagram involves a single process unit - a reactor with one inlet stream and one exit stream.

Stream 1 ⟶ | Equilibrium Reactor | ⟶ Stream 2

Stream 1
CO
H_2
N_2

Stream 2
CO
H_2
N_2
CH_3OH
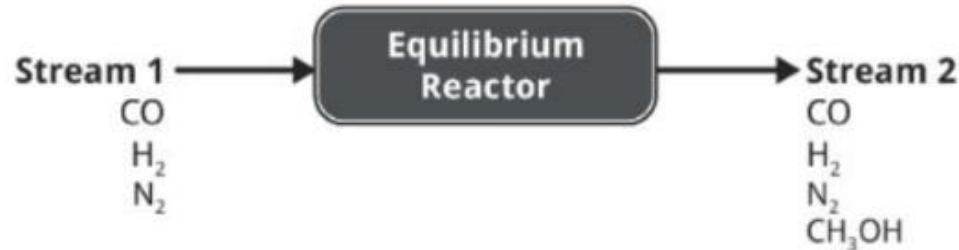
(b) Calculate the component molar flow rates for all of the components exiting the reactor (mol/min).

Established rating tool: PROCESS

Problem-solving domain
Problem definition
Represent the problem
Organize information
Calculations
Solution completion
Solution accuracy

Grigg, S. J. & Benson, L. European Journal of Engineering Education, 2014. *39(6)*: 617-635.

Detailed PROCESS rubric

Problem-solving domain	Tasks performed	Level of completion				Source of Error
		Missing	**Inadequate**	**Adequate**	**Accurate**	
		0 points	**1 point**	**2 point**	**3 point**	
Problem definition	Identify unknown	Did not explicitly identify problem	Completed few problem/system tasks with many errors	Completed most problem/system desks with few errors	Clearly and correctly identified and defined problem/system	

Grigg, S. J. & Benson, L. European Journal of Engineering Education, 2014. *39(6)*: 617-635.

Accuracy in assessment

$$\log \left(\frac{Pnijjk}{Pnijk{-}1} \right) \ = \ \ B_n \ \ - \ D_i \ \ - \ C_j \ - \ F_k$$

B_n is the ability of student *n*.

D_i is the difficulty of item *i*.

C_j is the severity of judge *j*.

F_k is the extra difficulty overcome in being observed
    at the level of category *k,* relative to category *k-1*.

Rasch fundamentals of measurement
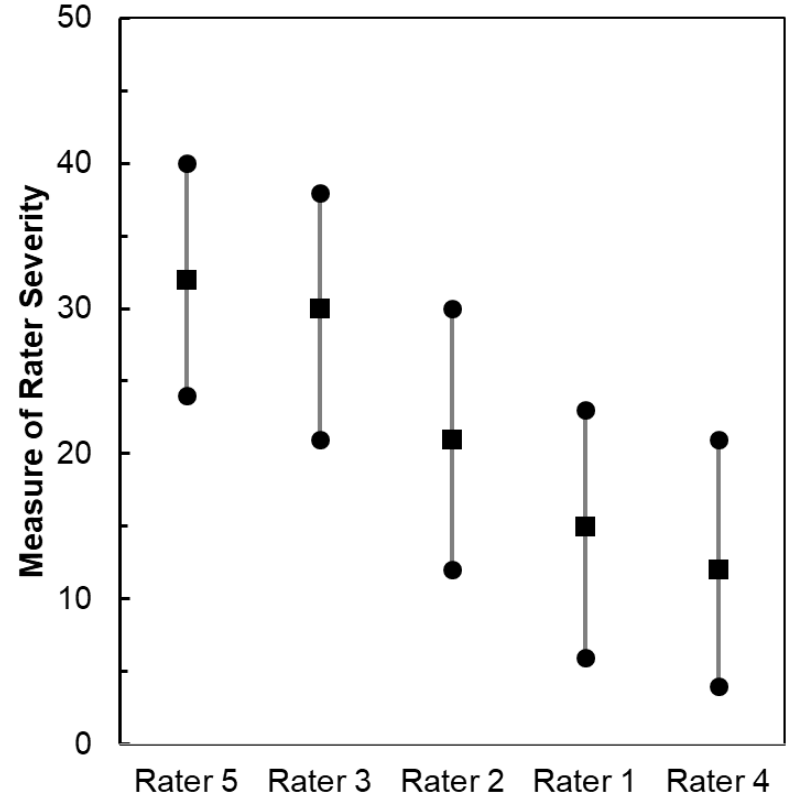
Principle of Invariance
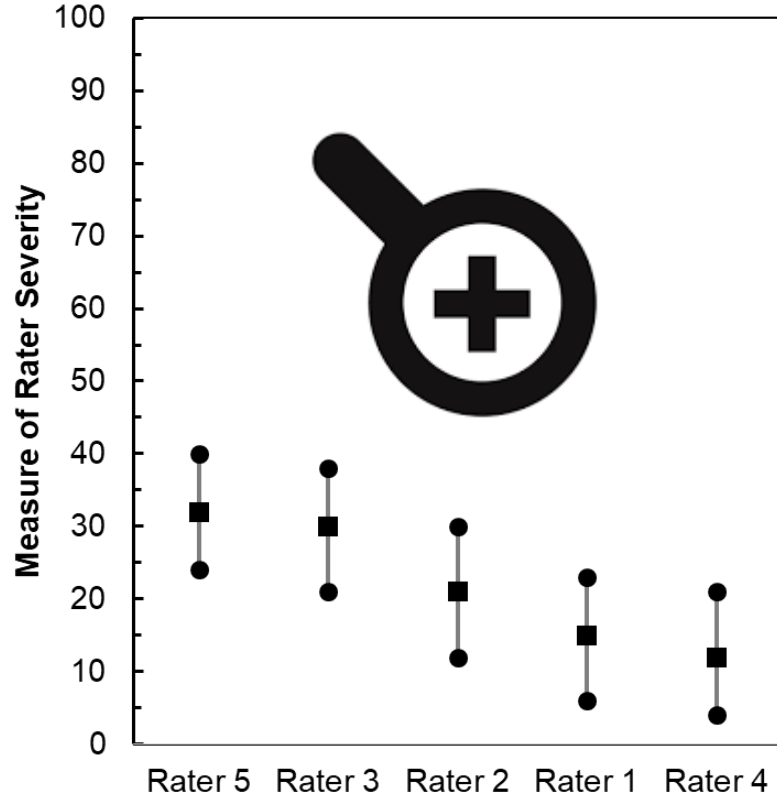
Monotonicity

Unidimensionality

Local independence

Rasch creates common measure

Overall Measure	- Rater Severity -	+ Student Ability +	- PROCESS Item Difficulty -	Rating Category
100		A		(3)
90		B C		--
80				
70		D E	Solution Accuracy	2
60				
50		F G H	Identify	--
		I	Organize	
40	Rater 3 Rater 5	J K	Allocate	
30		L	Represent	1
20	Rater 2	M N O	Solution Completion	
10	Rater 1 Rater 4	P Q		--
		R S		
0		T		(0)

9

Measuring rater bias

Raters discuss similar scores

Scores for Student E:

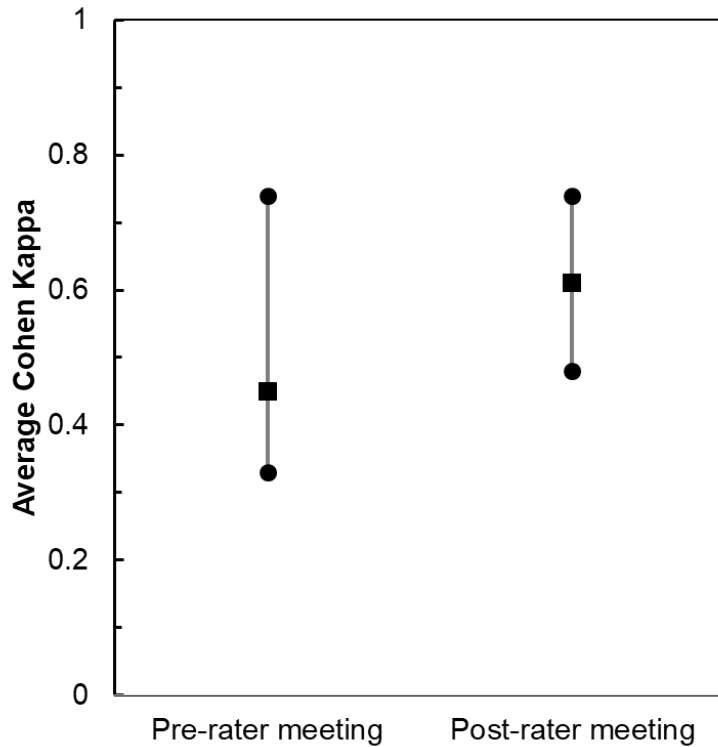| Rater	Problem definition	Represent problem	Organize knowledge	Calculate	Solution completion	Solution accuracy
Rater 1	3	3	3	3	3	1
Rater 2	3	2	3	3	3	1
Rater 3	3	3	3	3	3	3
Rater 4	3	3	3	3	3	1
Rater 5	3	3	3	2	2	1

Raters identify differences

Scores for Student M:

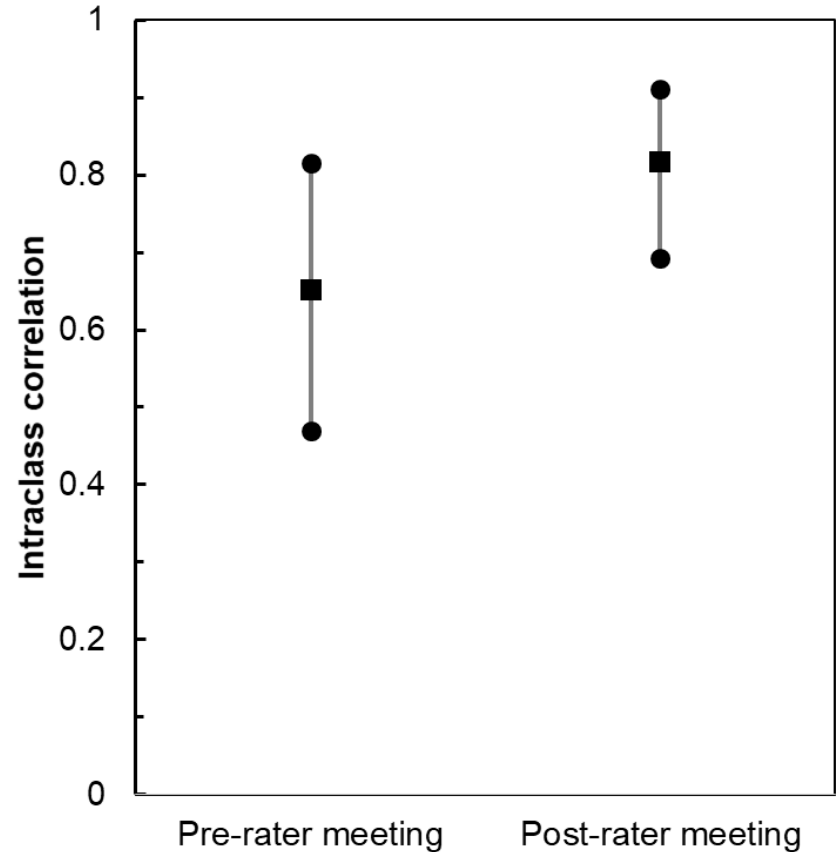| Rater	Problem definition	Represent problem	Organize knowledge	Calculate	Solution completion	Solution accuracy
Rater 1	3	3	3	2	3	1
Rater 2	3	2	1	1	2	1
Rater 3	3	3	3	3	2	3
Rater 4	2	3	3	2	1	1
Rater 5	3	3	1	3	2	1

Improving rater agreement

Level of agreement: Weak → Moderate

Conclusions

Iterative reliability evaluation → Accuracy of assessment

Identify source of measurement errors

Greater adherence to measurement principles

Thank you and...

Katherine Roach, Caleb Sims, Lindsey Stevens, countless TAs

DUE-1712186

University of Toledo IRB protocols 202214