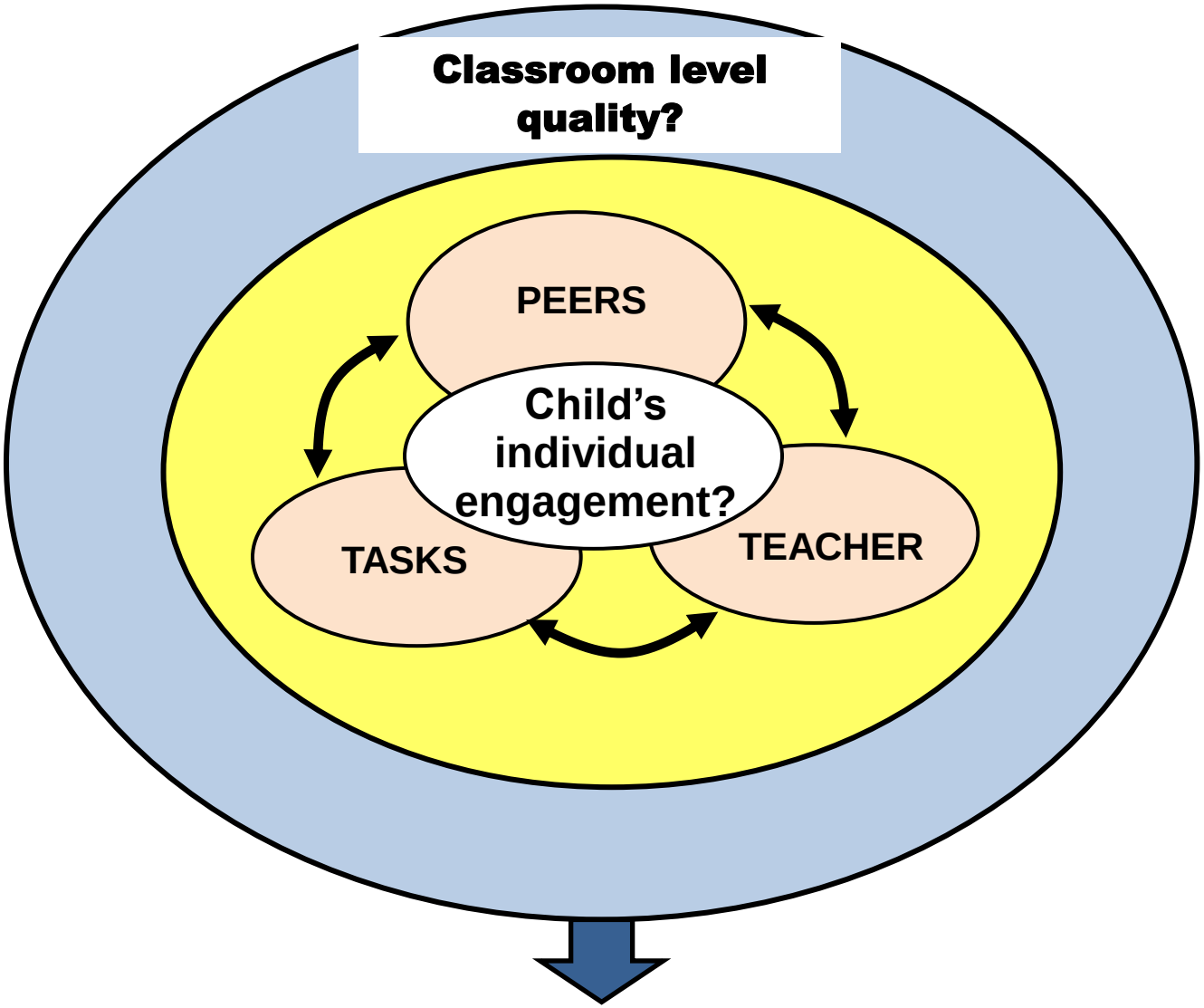
Measuring classroom quality with the CLASS in four different countries

Marja-Kristiina Lerkkanen - University of Jyväskylä

What is the aim of the study?



Some Observation Instruments of the Classroom Quality

ECERS-R - The Early Childhood Environmental Rating Scale - Revised (Harms, Clifford, & Cryer, 1999)

ITERS-R - The Infant Toddler Environment Rating Scale – Revised (Harms, Cryer, & Clifford, 2006)

Quality of the early childhood environment , e.g., use of space, materials and routines, activities;
Ranking: 1-7

CLASS - Classroom Assessment Scoring System; (Pianta, La Paro & Hamre, 2006, 2008)

Classroom quality (teacher-student interactions) affecting student outcomes.
2-4 Domains (e.g. *Emotional support, Organizational support, Instructional support)* and Dimensions.

Ratings (1-7):Time sampling with ratings low (1-2), middle (3-5), high (6-7).

ECERS-E - The Early Childhood Environment Rating Scale – Extension (Sylva, Siraj-Blatchford, & Taggart, 2014)

Was developed to supplement the ECERS-R in the area of curriculum.
The Four Curricular Subscales Extension: *Literacy; Mathematics; Science and Environment ; Diversity*
Ranking: 1-7

ECCOM - Early Childhood Classroom Observation (Stipek & Byler, 2005)

Teacher practices affecting student outcomes and motivation in preschool and Kindergarten

Ratings (1-5): Child centered, Teacher directed, Child dominated practices (in climate, management, instruction)

Classroom Assessment Scoring System - CLASS

(by Pianta, Hamre et al.)


University of Virginia, Charlottesville, USA

- **Objective:** Measurement of classroom level variables that produce benefits for children's development
- **Development:** Based on a review of teacher education, educational environments, and observational research literatures
- **Versions:** **Infant; Toddler; Pre-K; K-3; Middle school; Secondary**
- **Reliability:** A standardized metric and training protocol
- **Validity:** Focuses observation on adult/teacher-child interactions that have been shown to be associated with student gains
- **Professional development:** *MyTeachingPartner* provides common language and a lens through which to view classroom processes

Four studies

	Portugal	The Netherlands	Germany	Finland
Age group	Infants (under 3-year)	Day care and preschool (2-3-year-olds)	Preschools (3-6-year-olds)	Kindergarten (6-year-olds)
Number of classrooms	90	271	63	49
CLASS version	Infant (4 dimensions)	Toddler (8 dimensions)	Pre-K (10 dimension)	Pre-K (10 dimension)
Cycles	4 x 20-min	4 x 20-min	4 x 18-min	10 x 20-min + 10 min rating blocks
Other instruments	ITERS-R obs	Teacher and classroom characteristics	STRS teacher ratings	ECCOM obs, teacher ratings of practices

Reliability

- Are observers reliable users of the CLASS?
  → Training procedure in Virginia; an average interrater reliability of 87%, within one point of master codes
- How much consistency is there across users of the CLASS?
  → double-coding procedure (15-20%)
  → Some dimensions are more challenging to code than some others
- How stable are scores on the CLASS?

→ across at least 4 cycles?
→ across 2 days in week?
→ across the school year fall and spring???
→ across children and/or days???

Construct validity

- Does the CLASS measure constructs of importance in classrooms?

→The CLASS has been developed based on literature review on classroom practices shown to relate to students' social and academic development, and extensive piloting.

→ However, the contruct needs to be tested again when adapting the intrument to new (cultural) context →*the results has shown that construct is not necessarily very clear.*

→Differences in factor structure between CLASS versions (e.g. negative climate is in different domain in CLASS Pre-K vs. Secondary) and between some countries (US vs. Finland vs. France)

Criterion validity

- How does the CLASS relate to other measures of classroom quality and associated constructs?

  → To do concurrent observation with another observational instument(s) → Correlations between CLASS vs. ITERS-R/ECERS-R interaction effects/ECCOM teacher practices

  → To use other information (e.g. teacher's ratings, colleague ratings, student ratings, (parents ratings) of teacher practices or style) and see correlational relationships → However, there might be some limitations with this information too.

Predictive validity

- How does the CLASS relate to child development?

→Global classroom quality?

→Each domain? → number of studies

→Each dimenssion to predict child outcomes?

→Which aspects of the variance are related to different child outcomes?

→What is the most important level of information to transform for practices?

Predicting child development ("Pre-K", 4-year-olds)

	Classroom Processes			Structural features		
	CLASS: Emotional support	CLASS: Instructional support	ECERS-R Total	Level of teacher education	Teacher-child ration max 10:1	NIEER index
Language comprehension		X				
Productive language skills		X	X			
Phonological skills (rhimes)		X				
Letter naming		X				
Math		X				
Social skills	X					
Behavioural problems	X					

(Mashburn, et al. 2008, *Child Development*)

Some limitations of the CLASS…

CLASS scores from 4 cycles are averaged across data.

→Do we miss some information of reality?

→Do we miss the variation during the day or between the days?

→Does morning differ from the afternoon or Monday from Friday?

→Do we expect normal distributions from the domains?

What do we actually observe?

How do we select which activities to observe?

→Do we observe only indoor activities?

→What about outdoor activities or other context (e.g., field trips)?

Some more limitations….

How do we select teachers/classrooms?

To whom we focus our observation? Teacher(s)-child(ren)? Two very different teachers in classroom?

→ Are teachers randomly selected vs. participating voluntarily vs. teachers send their own videotapes?

→Do we get different results from live observations (20 min cycle – 5-10 min blocks rating) vs. video recorded ratings when every minute will be coded?

→ Are 20 minutes cycles too long? Do we miss something? (compare to ISI 20 second coding procedure by Connor & Morrison)

→Are 20 minutes cycles too short? (compare to ECCOM)