Improving feedback from teaching assistants

Towards timely, fair and high quality feedback at scale

WMSEE, 9 April 2024

Torgeir Dingsøyr
Department of Computer Science

Introductory course in software engineering

- 7,5 ECTS - one of four courses in spring semester

- About 500 students, organised in 72 development teams. Students from 8 study programs.

- 30 teaching assistants (6 "senior", 24 "junior"), two professors, two adjunct lecturers

- Project work with over 500 meetings between teams and supervisor

- 144 demonstrations of products

- Feedback on over 1000 "deliverables", portfolio assessment

Feedback in the course: Student perceptions

"Uneven feedback. The teaching assistants use evaluation criteria in different ways. I asked for clarification as much which was written in the feedback was wrong, but was told that we could only complain on the whole grade and not get any more feedback"

"It seems like the ones giving feedback look for particular aspects not mentioned in the assignment, for example in the retrospectives. Here those who had structured the deliverable into "works well" and "could be improved" got a good score, while those who had structured the report after topics got a lower score as this was interpreted as less structured"

"The feedback often seemed arbitrary; you were criticized for things not stated in the assignment such as lack of figure text. Some of the feedback is weighted very strange and seems unsupported"
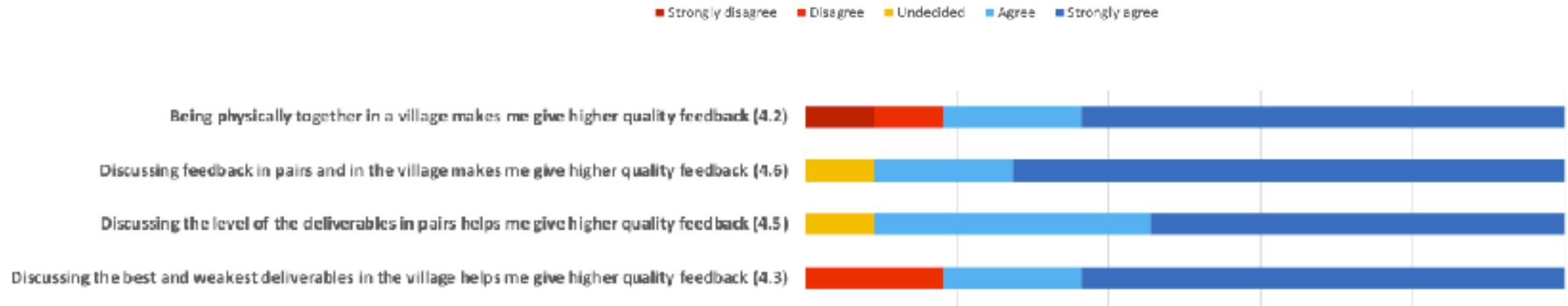
Change in work process

- Old process:

    - TAs evaluate and provide feedback individually

- New process:

    - TAs work in pairs

    - Discussion first in pairs, then with in group of six TAs

What teaching assistants say

Legend: Strongly disagree · Disagree · Undecided · Agree · Strongly agree

- Being physically together in a village makes me give higher quality feedback (4.2)
- Discussing feedback in pairs and in the village makes me give higher quality feedback (4.6)
- Discussing the level of the deliverables in pairs helps me give higher quality feedback (4.5)
- Discussing the best and weakest deliverables in the village helps me give higher quality feedback (4.3)

Change from 2022 to 2023
On average: lower scores from students, higher from TAs

	Students						Teaching assistants					
	Useful feedback			Clear evaluation criteria			Clear evaluation criteria			Good support in evaluation		
	2023	2022	Diff.	2023	2022	Diff.	2023	2022	Diff.	2023	2022	Diff.
Deliverable 1: Group contract (pass / not pass)	4,3	4,2	0,1	3,1	4,2	-1,1	4,9	4,3	0,6	5,0	4,2	0,8
Deliverable 2: Theory test (5%)	4,3	3,3	1,0	3,5	4,1	-0,6						
Deliverable 3: Prestudy report (pass / not pass)	3,1	3,9	-0,8	3,7	3,7	0,0	4,5	3,9	0,6	4,8	4,1	0,7
Deliverable 4: Sprint review1 (5%)	2,7	3,4	-0,7	3,2	3,1	0,1	4,1	3,4	0,7	4,7	4,1	0,6
Deliverable 5: Retrospective1 (5%)	2,6	3,6	-1,0	3,1	3,0	0,1	4,0	3,4	0,6	4,8	4,2	0,6
Deliverable 6: Sprint review2 (5%)	3,4	3,8	-0,4	3,1	3,3	-0,2	4,2	3,5	0,7	4,6	4,4	0,2
Deliverable 7: Retrospective2 (5%)	3,4	3,7	-0,3	3,1	3,3	-0,2	3,3	3,6	-0,3	3,9	4,1	-0,2
Deliverable 8: Individual reflection report (35%)	3,3	3,3	0,0	2,8	3,0	-0,2	3,7	3,6	0,1	4,6	4,5	0,1
Average	3,4	3,7	-0,3	3,2	3,4	-0,2	4,1	3,7	0,4	4,6	4,2	0,4

First step: Confident TAs.
For students: Reduction probably due to more critical feedback

Course structure: 2023
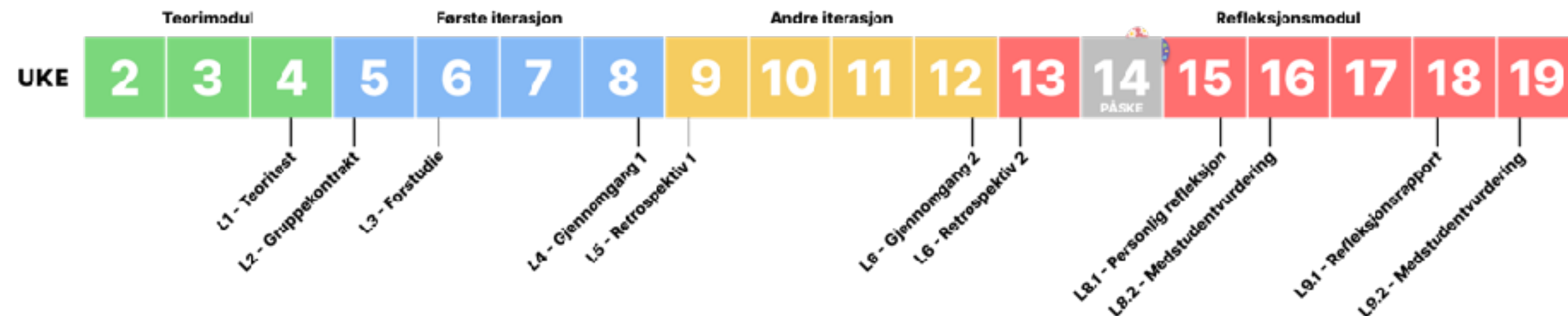
D1 - Theory multiple choice test (pass / not pass)
D2 - Team contract (pass / not pass)
D3 - Prestudy report (5%)
D4 - Sprint review1 (5%)
D5 - Sprint retrospective1 (5%)

D6 - Sprint review2 (5%)
D7 - Sprint retrospective2 (5%)
D8 - Individual reflection report (35%) + 3 peer reviews
D9 - Team reflection report (40%) + **2 peer reviews**