Norwegian University of Science and Technology

Educating Reflective Systems Developers at Scale

Towards “productive feedback” in a semi-capstone large-scale software engineering course

WMSEE, 31 May 2023

Torgeir Dingsøyr
Department of Computer Science

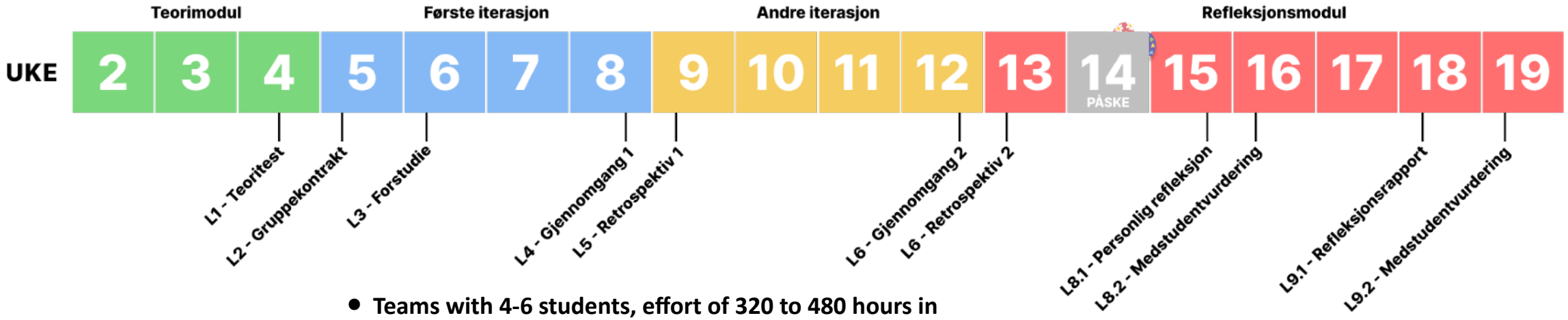
Introductory course in software engineering

- Students should be able to plan and manage small software development projects
- Learning objectives on requirements engineering, software quality, software architecture and development process. Focus on methods relevant for a small co-located development team: Agile development methods (Scrum, Kanban, XP).
- 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
- 7,5 ECTS - one of four courses in spring semester

Course structure

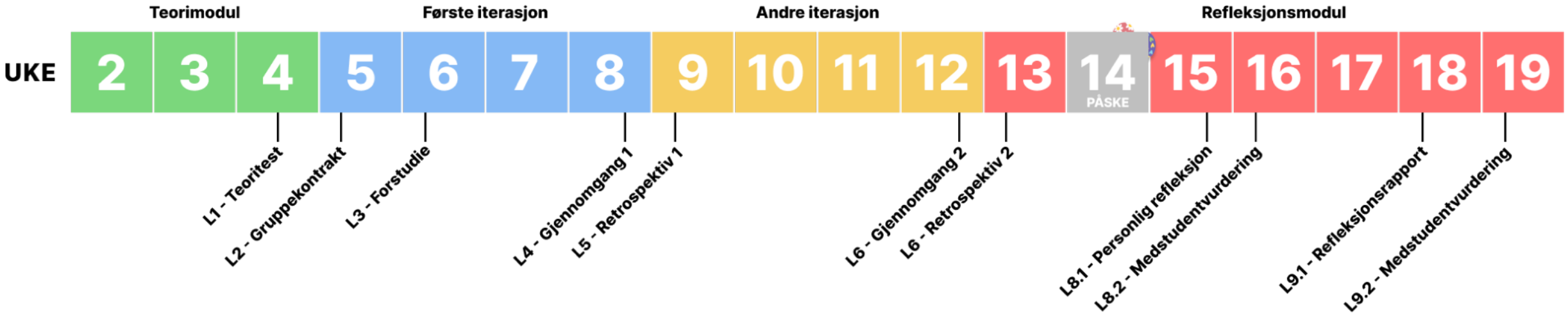
- On software engineering
- Agile principles
- Requirements engineering
- Scrum
- Software quality
- Technology stack

- Introduction to literature search
- Video on “reflective writing”
- On use of AI
- Individual essay on 7 topics
- 3 peer reviews on individual essay
- 2 peer reviews on group report



- Teams with 4-6 students, effort of 320 to 480 hours in product development
- 4 main products, 72 variants
- All Scrum practices mandatory + Pair programming
- Weekly supervision meeting: Supervisor & PO
- New user stories in second iteration

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**

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"

Evalueringsrubrikk for team: <X>

Evalueringskriterier	Ikke bestått	Tilfreds-stillende	Nokså god	God	Meget god	Fremragende
GENERELT						
Bruker teamet relevante kilder for å begrunne valg?						
PROBLEMBESKRIVELSE						
Klarer teamet å beskrive problemet til produkteieren tydelig?						
VALG AV TEKNOLOGISTAKK						
Klarer teamet å reflektere over og begrunne valg av teknologi?						
PRODUKTKVALITET						
Forklarer teamet valg av aktiviteter for å sikre produktkvalitet?						
Presenterer teamet fornuftige testaktiviteter for første og tredje smidige testkvadrant?						
LANSERINGSPLAN						
Presenterer gruppa en realistisk lanseringsplan?						
Har teamet definert et tydelig mål, tidsestimater og brukerhistorier for første sprint? Utviklingsoppgaver for tre første brukerhistoriene?						

KVALITATIV TILBAKEMELDING OG TOTALEVALUERING
Tre til fire setninger med kvalitativ tilbakemelding med fokus på hva som er gjort bra i leveransen, og hva som kunne vært gjort bedre.

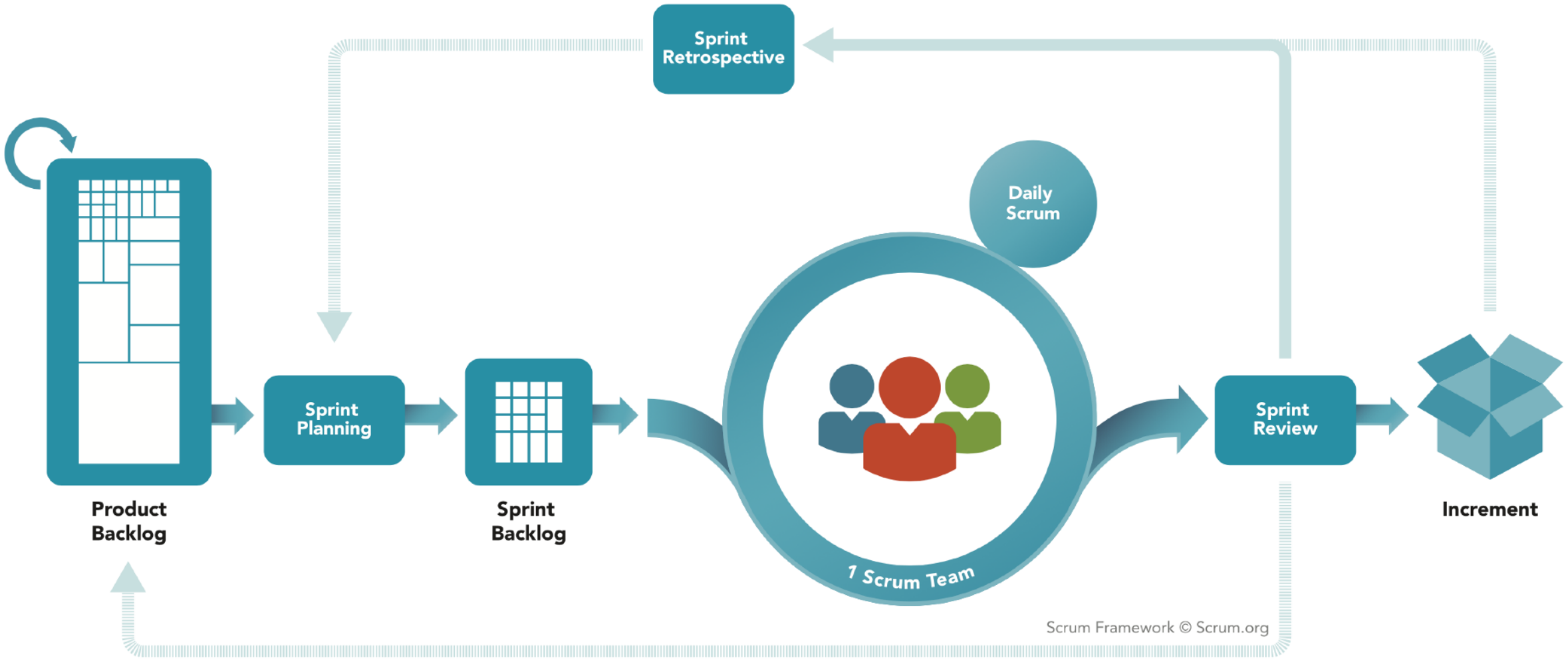
	Ikke bestått	Tilfredsstillende	Nokså god	God	Meget god	Fremragende
TOTALEVALUERING						

Feedback in literature

- Students were critical to feedback received, which was perceived as having an “overly negative tone” or was described as “vague” or “ambiguous” (Price et al. 2010)
- Teachers recognised the importance of feedback but had few ideas of the effect of the feedback apart from seeming to think that a large volume of feedback lead to increased learning (Price et al. 2010)
- Use of rubrics: “their complexity of language may confuse students and some students do not even read rubrics because they find them too complicated” (Bacchus et al. 2020)
- While 73% of students always read rubrics, only 43% understood what was required from reading them (study cited in Bacchus et al. 2020)
- *Productive feedback:* Students make meaning of information about the quality of their performance and they develop skills allowing them to capitalize on similar learning opportunities in the future (Esterhazy et al. 2019)

Price, M., Handley, K., Millar, J., and O'donovan, B., "Feedback: all that effort, but what is the effect?," *Assessment & Evaluation in Higher Education,* vol. 35, pp. 277-289, 2010
Bacchus, R., Colvin, E., Knight, E. B., and Ritter, L., "When rubrics aren’t enough: Exploring exemplars and student rubric co-construction," *Journal of Curriculum and Pedagogy,* vol. 17, pp. 48-61, 2020..
Esterhazy, R., Nerland, M., and Damşa, C., "Designing for productive feedback: an analysis of two undergraduate courses in biology and engineering," *Teaching in Higher Education*, pp. 1-17, 2019.

Feedback opportunities in iterations: Scrum



Source: www.scrum.org

Coping with scale: 2020

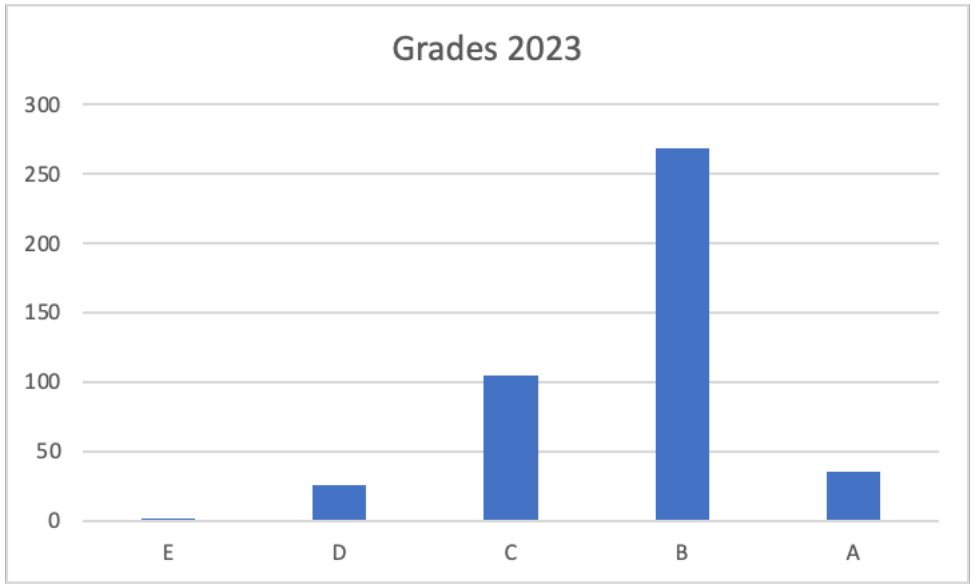
- Team supported by
    - Supervisor (TA)
    - Product owner (TA)
- Deliverables evaluated using rubrics by
    - Two independent TAs
    - A third TA summarised feedback which was given in written
- Example rubric for “Sprint review”
    - 10 points to be considered
    - and nine topics where evaluators were to give feedback on points which was positive and negative
    - In addition, the evaluator was asked to write between 100 and 200 words as feedback to the team.

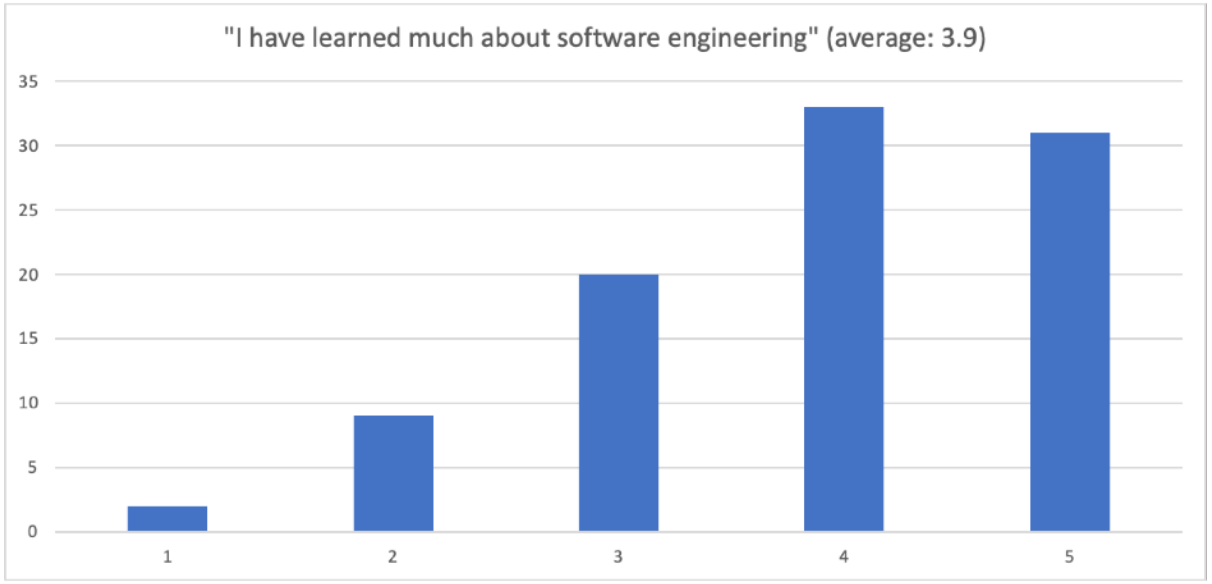
Actions taken last three years

Reorganising to provide project teams with more productive feedback at scale:

- Rubrics and assignments revised to reduce number of evaluation criteria and increase clarity. First oral then written feedback
- Teaching staff reorganised according to principles of large-scale agile development, four “villages” consisting of 6 TAs and one TA “village facilitator”. Each village supported by an external agile coach. TAs work in pairs
- Increased training of TAs - in total 25 hours for training, five four-hour workshops and time for re-reading course curriculum
- Introduced peer assessment (individual reflection report, group reflection report), groups observe sprint reviews of a “sister group”
- Feedback discussed in “village meetings” (2 hours) prior to giving oral and written feedback

Learning outcome





What teachers think works well

- Project work motivates for effort
- Frequent feedback on deliverables, formative assessment
- Weaker students learn from stronger students in team
- Peer assessment provides more feedback than staff is able to
- Teaching assistants appreciate training and organisation

What teachers think should be improved

- Challenge of timely, fair and high quality feedback to 72 teams
- Too much effort on programming during project work (14.5 hours per week, aim: 10)
- And too little effort spent on reading curriculum
- Many teams tend to focus more on grades than feedback

Ideas for further development

Managing challenges in large project courses

- We expect that more study programs will take the course, maybe 600 students in a few years from now

- Discussion in small groups: *Measures to manage scale in large project courses*

Feedback in the course: Student perceptions

“*we have received good feedback in the course”*

“it was good to have more of the same type of deliverables, so that we can improve”

“useful and fast feedback on the deliverables”

But also:

“make evaluation criteria clearer – we got a low score on something we did not know was evaluated”

“some teaching assistants seemed to be much more strict than others, which makes the assessment unfair”

“the teaching assistant give strange feedback and does not want to discuss the feedback. The feedback depends a lot on which teaching assistant you get”.

Lessons learned: Productive feedback at scale

- Sacrifice breadth and volume in feedback to increase timeliness and fairness
- Exploit lessons learned in industry on how to manage scale
- Invest in training of teaching assistants

A step forward towards educating reflective systems developers who "*must bring to bear something more than a repertoire of methods and tools. They must engage in reflections and dialogues to get the necessary insights into the situation at hand*" (Mathiassen and Purao, 2002)

Mathiassen, L. and Purao, S., "Educating reflective systems developers," Information Systems Journal, vol. 12, pp. 81-102, 2002.