Encouraging Students Evaluative Judgement in Project Courses: Do's and Don't
Lessons Learned from Three Course Instances

Truong Ho-Quang

Chalmers and Gothenburg University, Sweden

Abstract

Evaluative judgement, which is defined as "the capability to make decisions about the quality of work of oneself and others", has long been recognised and considred in higher education literature as an important goal of study programs. A number of in-course practices (such as self-assessment, peer review, feedback, etc.) have been used to help students to develop evaluative judgement. However, only a few empirical studies has been conducted to actually prove the impacts of those practices to improvement of student's evaluative judgement. In this paper, we present our observation on student's judgement ability during a period of 3 years across 3 instances of a project course where different practices have been applied to help students to achieve evaluative judgement at each instance. From the observation, we present a lesson

1. Introduction

According to Cowan Cowan [1], evaluative judgement has recently been taken up in the higher education literature, as a high-level cognitive ability required for life-long learning. The impact of evaluative and critical judgement to production of effective feedback have been underlined by Sadler [2], Nicol [3, 4] and Boud and Molloy [5].

Little empirical work has so far been conducted within an explicit evaluative judgement framework. Firstly, Nicol et al. (2014) demonstrated the ability of peer learning activities to facilitate students judgement making in higher education settings. Secondly, Tai et al. (2016) also explored the role of informal peer learning in producing accurate evaluative judgements, which

impacted on students capacity to engage in feedback conversations, through a better understanding of standards. Thirdly, Barton et al. (2016) reframed formal feedback processes to develop evaluative judgement, including dialogic feedback, self-assessment, and a programmatic feedback journal.

This paper is structured as follow: Section 2 provides readers with basic knowledge of evaluative judgement, section 3 summarises the methods to collect and analyse data. Section 4, 5 and 6 highlight the settings of instances of the project course toward applying the practices. Section 7 concludes the paper with lesson learned and potential future works.

2. Background

This section aims to provide readers with basic components of evaluative judgement including its definition (Section 2.1), the practices that have been used to improve student's evaluative judgement (Section 2.2) and the threats to the development of student's evaluative judgement (Section 2.3). This knowledge is lately used as a basis to discuss the treatments/incentives for developing student's evaluative judgement in the three course-instances.

2.1. Definition of Evaluative Judgement

The definition of evaluative judgement has a root from a closely-related literature which is formative assessment. This supports students developing capability for evaluative judgement through applying practices of self-assessment [6, 7] and peer assessment [8, 9]. Tai. et al. proposed to take into account an extra aspect of sustainable assessment (introduced by Boud and Soler [10]) when it comes to definition of evaluative judgement [11].

Tai et al. given two definitions of evaluative judgement in two different contexts. The first definition was formed in the context of medical student's learning on clinical placements being: *" the ability to critically assess a performance in relation to a predefined but not necessarily explicit standard, which entails a complex process of reflection. It has an internal application, in the form of self-evaluation, and an external application, in making decisions about the quality of others work"* (pp.661 of [11]). The second definition is a simplified version of the first definition with exclusion of tautological components. The second definition is: ***"Evaluative judgement is the capability to make decisions about the quality of work of self and others"*** [12]. In this paper, we discuss student's evaluative judgement in

2

general project courses, thus we shall take the latter as the baseline to build up our discussion.

2.2. Practices to Improve Student's Evaluative Judgement

According to Tai. et al., student's evaluative adjudgement can be developed in higher educational courses by applying/employing a number of practices, including self-assessment, peer review/assessment, feedback, rubrics and examplars [12]. This section summarises the practices and its potential application in project courses.

2.2.1. Self-assessment

Self-assessment is the practice that encourages students to appraise their own work. This practice often implies a comparison of student-generated marks and teacher-generated marks. Despite long-established criticisms by Boud and Falchikov [13], this strategy remains common in self-assessment research (Panadero et al. 2016) because of a number of reasons. Firstly, such comparison is easy to conduct and analyse within a course - It requires little to no changes to curriculum and pedagogy. Secondly, there remains a legitimate interest in understanding students self-marking [12].

Self-assessment can assist students to develop multiple criteria and qualitatively review their own work against them. Therefore self-assessment helps students to refine their judgements rather than to generate grades that might be distorted by their potential use as a substitute for teacher grades.

Self-assessment can only be effectively used with students thoroughly understanding the task they are performing, the requirements towards the task what is expected as 'standard'. In modern project course which often consists of highly-dynamic tasks (with the aim to capture realistic working environment), before applying self-assessment, teachers need to pay an extra attention to ensure all students have an adequate understanding of the above-mentioned components.

2.2.2. Peer Review/Assessment

Similar to self-assessment, peer review/assessment practice focus on the accuracy of marks generated and the potential for a number calculated from peer assessments to contribute to grades [14, 15]. Peer review/assessment is aimed at contributing to multiple aspects of teaching and learning, including formative, educative and pedagogical elements [16, 17, 8]. In addition, with peer review/assessment, students are required to give and receive peer

3

feedbacks. According to van den Berg et al. [18] and Nicol et al. [4], this significantly contributes to student learning and development of evaluative judgement.

Empirical works by McConlogue [19] and Tai. et al. [11] report that by engaging more closely with criteria and standards, and having to compare work to these, students have obtained a better understanding of quality work. In particular, by a more explicit focus on evaluative judgement, students can pay close attention to what constitutes "quality" in others work and how that may transfer to their own work. Moreover, understanding the notions of quality through evaluating the work of peer students may reduce the effect of cognitive biases directed towards the self when undertaking self-assessment [20]. Therefore, this is generally considered as a complementary option for self-assessment. However, it is important (for teachers) to keep in mind that conducting peer review/assessment is often more "expensive" than self-assessment.

2.2.3. Feedback

Feedback is the practice where dialogues are given to students about their work as the courses undergo. Recent studies about feedback practices in higher education has emphasised the effects of feedback on students subsequent work and on engaging the students as active agents in a forward looking dialogue about their studies [5, 21, 22]. These works suggest that feedback could help develop students evaluative judgement in different ways, including: i) clarifying applicable standards and criteria that define quality; ii) refining the judgement of students about their work; and iii) assisting students to formulate actions that arise from their appreciation of their work.

The consideration of feedback toward evaluative judgement requires an extra attention to the content of the feedback and the way it is given. Comments to students should not only about their work per se but about *the judgements they make about it.* According to Tai. et al., in order for students to refine their judgements, the following feedbacks can be given: i) whether the students have chosen suitable criteria/standard; ii) whether the students have reached justifiable decisions about the work; and iii) what the students need to do to develop these capabilities further [12].

In the context of project courses where the working topic is often broad and could be extensively developed, feedback can also be used to help students to scope their work. With a consideration to develop student's evaluative judgement in mind, teachers may give comments that trigger students

to judge their (own or group's) resource and capability to finish given tasks, thus better scope their project.

2.2.4. Rubrics

Rubric is defined as *"a document that articulates the expectations for an assignment by listing the criteria or what counts, and describing levels of quality from excellent to poor"* [23, 24, 25]. A rubric has three essential features: evaluation criteria, quality definitions and a scoring strategy [26]. Rubrics can be considered of as a scaffold or pedagogy to support the development of evaluative judgement. There is extensive diversity of rubric practices, some of which better supports evaluative judgement than others, according to Dawson [27].

Prior study (such as[5]) suggests that students could be trained to use rubrics to evaluate their own work and the work of others. However, the ability to apply a rubric does not necessarily imply that students have developed or achieved evaluative judgement without a rubric. Hendry and Anderson (2013) proposed that marking guides are only helpful for students to understand quality when the students themselves use them to critically evaluate work. Andrade, by observing how middle-school students use rubrics, concludes that students must engage deeply with rubrics, perhaps by co-creating them and using them for self- and peer assessments. Andrade lists an example of proper use of rubrics in Reitmeier et al.'s study [28].

In addition, to support the development of evaluative judgement, rubrics should be designed and used in ways that reflect the sound understandings of quality that experts use when making judgement about a particular type of task while acknowledging the complexity and fluidity of the practices, and hence standards, they represent [12].

2.2.5. Exemplars

Sadler defines exemplars as *"key examples of products or processes chosen so as to be typical of designated levels of quality or competence"* [29]. Exemplars can also be simply understood as examples of best or worst practice in the educational environment. Exemplars provide a form of scaffolding for students. Scaffolding refers to support provided during the learning process, focused on helping students achieve learning goals and gradually reduced or faded according to the progress of the learner. The intention is that the support not only assists students in accomplishing tasks but also enables them to learn from the experience to improve performance in future tasks [30]. Wise

and O'Neill mention an important social aspect to scaffolding in which learning occurs through interactions with more capable others, primarily peers or teachers [31].

A challenge in using exemplars is to ensure that they are sustainable, meaning that, they represent features not just of a given task but of a broader discipline or profession. In project courses, which often aim at preparing students to future working tasks, this challenge can be resolved by recruiting actual industrial tasks and involving industrial assessment in the examplars.

2.3. Threats to the Development of Student's Evaluative Judgement

Joughin et al. in their paper [32] draw an attention to a number of heuristics that unconsciously hold the ability of making evaluative judgement and decision-making. This section summarises the threats listed in the paper, including attribute substitution, overconfidence, the endowment effect, anchoring & addjustment and the law of least effort. We would argue it is very important to keep in mind and to manage these threats when designing a course with the objective of developing student's evaluative judgement.

2.3.1. Attribute Substitution

According to this heuristic, when confronted with a difficult, complex problem, we human may unconsciously substitute a simpler problem for which we can find an answer. Kahneman refers to this as "answering an easier question" [33]. The main concern is whether students, asked to evaluate their work by applying a complex rubric or a criterion they do not understand, might respond by unconsciously simplifying the task. For example, instead of answering the question, How well did I do against this criterion?, the students may unconsciously substitute the question, How much effort did I invest in relation to this criterion?.

2.3.2. Framing

Framing refers to ones conception of a task, its intended outcomes and the implications of those outcomes. Framing arises from both how a task is presented and how it is perceived and is thus influenced by both the task per se and its context. The main concern is that any given task can be presented in different ways, and any given presentation of a task can be perceived differently. The impact of framing has been powerfully demonstrated in relation to complex judgements in medical [34] and financial contexts [35] amongst others. For example, a task being "Write the formula on the white board"

6

can be interpreted as writing an actual chemical formula or writing the words "the formula" on the whiteboard, resulting in two different outputs. To reduce this type of threat, it is important that teachers and students have to achieve a mutual agreement on their "framing" of the given task. Techniques such as *replicating*, *clarification*, *requesting student's interpretation* might be used by teachers to ensure check student's framing of the task.

2.3.3. Overconfidence

Overconfidence plays a major role in misjudgement and flawed decisions and is linked to numerous other factors affecting judgement [36, 37]. Overconfidence concerns human's tendency to overrate the quality of own work, decisions or judgement and is based on how coherent a view of the judgement we have. More information might challenge that sense so that, we may be more confident if we have less rather than more information to guide our judgement. Accordingly, the overconfidence heuristic should be expected to lead students to overvalue their work, with less knowledgeable students more prone to this than their better informed and hence more uncertain peers. In fact, as the well-documented Dunning-Kruger effect shows, the effect on students is more nuanced, with lower performing students tending to exaggerate their grades while higher performing students tend to do the opposite [38]. This phenomenon provides a rare instance of research into an unconscious heuristic in higher education.

2.3.4. The Endowment Effect and the IKEA Effect

These effects concern our tendency to overvalue what we have come to own or what we have created ourselves. The *endowment effect*, so named by Thaler [39], posits that we attach more value to an object if we have a sense of ownership than if we do not. The *IKEA effect* as its name says-it-all: the increased valuing of products we have constructed ourselves compared to those constructed by others. These effects should be expected to inflate students judgements of work in which they have a strong personal investment and/ or which represent a significant effort or act of creation. The effects may apply not only to how highly a piece of work is valued, but also to the students judgement itself once that judgement has been made, so that students may not only over-rate their work but also hold more strongly to their evaluation of it.

7

2.3.5. Anchoring and Adjustment

This heuristic refers to the tendency to use a given value as a starting point and then adjust our judgement from that point. The effect of this heuristic on second marker judgement has been documented by Brooks (2012). When students are evaluating their own work, four sets of apparently useful points of reference may act as anchors: verbal descriptors of standards; exemplars; the work of peers; and their own previous work. Thus, for example, students advised of an expected class average of 80% would be more likely to rate their work as closer to 80% than students advised of an expected class average of 50% or students not advised of any average. A similar dynamic may occur in relation to student's evaluations with respect to individual criteria and standards. In other words, the presence of particular exemplars, criteria and standards statements may ironically work to distort rather than promote accurate judgement.

2.3.6. The Law of Least Effort

This heuristic assumes that we are working in the right direction, but that our attention and energy are finite and that we therefore tend to choose the means of achieving a goal that requires the least effort. This heuristic has several possible implications for students evaluative judgement. For example, if judgement requires attention to too many factors either simultaneously or sequentially, students may find their capacity to attend to these diminishing as they work through them. Moreover, if evaluative judgement is required at the end of a complex assessment task or is seen as peripheral to the main task, it may be done reluctantly and with limited energy and attention. If such judgement can be undertaken through several more or less intensive means, students may tend to opt for the most straightforward and avoid the more demanding.

3. Study Method

This study is constructed as an action research on a 15 HEC project course (DIT598 Software Evolution Project) over 3 course instances between September 2015 and December 2017. During the three iterations teachers and students lenses [40] were used to analyze problems and change the course design.

TBD: Course is about ...

Data for assessing the students lens was collected throughout the three courses with the help of course surveys and course evaluation meetings. Course evaluation meetings were held twice for each course instance, once around mid-term and once after the course was over. Between one and five students ("course representatives") participated in these meetings together with the teachers and, during the respective final meetings, the program responsible. The course survey includes standard questions, which were asked in all three years. While the course survey includes quantitative measures, assessed with Likert-scale questions, it is necessary to be careful when using these quantitative measures. This is due to the existence of changes that are not related to the research questions, which biases the quantitative measures. Instead, the focus is on the qualitative data, i.e. free-text answers in the course survey that can be related to the changes performed to address the researched issues. The teachers perspective will be reconstructed based on personal notes of the main teacher, who is also the author of this paper.

4. Course Instance Fall 2015/2016

4.1. Course Design

4.1.1. Team Setup

Teams had 6 to 8 students and were assigned semi-randomly. The assignment process was: Students could submit project proposals together in small teams of 1 to 4. Then students could vote for the projects they wanted to participate in. The 7 most popular projects were instantiated. Each of these projects was then filled first with the students who proposed it and afterwards with additional students who voted for it. Due to differences in popularity of the project proposals not all students could participate in the projects of their first choice.

4.1.2. In-course Components to Develop Student's Evaluative Judgement

Self-assessment. Allow students to choose their own projects as a way to motivate them to work and to reduce the learning curve and can better choose a project that fit their capability (e.g. the proramming language that fits, the topics that they like...).

Peer review. Peer reviews have done at several levels: i) between team members: teachers motivated the teams to apply industrial working practices such as work-in-pair (for coding/reviewing code); ii) between teams: in the

9

Scrums-of-scrum meetings, representatives from all groups meet and present their progress, other groups are allowed to ask questions and give recommendations

Feedback. Similar to peer-review/assessment, feedback was given in several levels: i) between team members: ...; ii) between groups: ; iii) between students and teachers: via supervision sessions or

Rubrics. The final grade included the teams grade to 50%. An additional part of the grade that was bound to the team performance was a bonus of up to 3 points (worth 3% of the grade) that could be reached by teams. The other 50% were based on: - the individual contribution to the project (10%): This differs a bit from the strategies of personal reporting and peer assessment listed by Hayes et al. [14], as students report themselves and others - but as a team. Thus, they have to agree on the reporting. In addition, the reporting should include details about the actual tasks performed. - the 4 discussions (20%) and an individual report at the end of the project (20%): Both, discussions and the individual report functioned as cross-validation as known from Hayes et al. [14]. We aligned the report to the projects by asking students to describe and measure quality improvements achieved on the projects during the team work.

Exemplars.

4.1.3. Data Collection

4.2. Result

4.2.1. Teacher's Perspective

We observed problems in multiple teams. Several teams split up the work into small chunks they worked on individually or in pairs (an effect that needs to be prevented for learning based on social constructivism). Some of these pairs worked in a way that made us teachers suspect that one student covered for lacking contribution of the other. There were single students who had little to no contribution to the teams work. In one team, some of the students discussed their progress in team meetings, but refused to actually share their contributions with the other team members over weeks. Furthermore, we saw a first case of a weaker student who was not supported by the team.

The topic choice of individual reports caused problems, as students competed for the topics that seemed easiest to address. This might ultimately

also have contributed to a less fair grading. Students would often also contribute very little to the team work and spend much more time working on their individual reports.

Many components of the course, e.g. poster sessions or discussion, had no positive effect on the teams. It was difficult to compare contributions between teams, due to the self-selection of the topics.

4.2.2. Student's Perspective

The following issues were mentioned by students in the course evaluation.

Teaching team does not have enough insight. Multiple feedback from students mentioned the lack of understanding and proper check between teachers and students. This feedback included partly very harsh comments on supervisors being useless and the teaching team not having enough insight into what the students really did. A student commented: "the teachers (have) got no deep insight about the code we were writing, our report's contents, our experiments, etc.". Surprisingly, the same student lately mentioned that he/she used a superficial communication strategy to always say "well" to the supervisors when being asked about the project progress. This can be considered as part of the anchoring effect where a student *with the expectation of getting no useful help from supervisor* in mind might overlook the opportunity to discuss and get feedback from teaching teams, thus lose the chance to develop his/her evaluative judgement.

Lack of communication between teachers and students. Some students mentioned lack of feedback and communication between their teams and teachers. This, according to the students, caused some serious issues as the course progresses, including: i) the teachers do not understand student's background (e.g. a student mentioned "...*the teachers thought that we had taken Software Quality (which wasn't taught last year), so they kind of assumed that we knew what we were doing, e.g. about the metrics...*"); ii) the students do not have the same views as teachers when it comes to project outcomes (e.g. a comment being "*... because they (the students) hadn't understood what kind of contributions she (the teacher) had expected*"). It is clear that this poor communication channel could leave a huge room for student's self interpretation of requirements and tasks, thus being a serious threat of *framing* and *attribute substitution* to development of student's evaluative judgement.

Lack of communication and peer review/assessment within team. Two students also reported lack of understanding and communication within their team, which lately resulted in stress taking extra work to cover others work, giving free-ride to other team members (so called free-rider/slacker [41]). As an example, a student reported: "*The majority of my group members were incompetent and I had to help them with coding problems, etc, while I gained nothing from them*". The comment might sound a bit extreme, but it clearly shows that the reporting student was expecting more feedback from his/her teammates. This might have a root cause from the team not being able to define/choose a task that every team members can participate or the other team-members intently expecting a free-ride. Either ways could potentially harm the student's experience with development of his/her own evaluative judgement.

Disconnection between individual components of the grading. A disconnection between individual components of the grading (discussions and individual reports) form the project was mentioned negatively. Students mentioned the lack of frequent checks and indicated that they did not work much on the course for a long time. A student also mentioned their decision to follow the law of least effort: "*many group members ended up changing characteristics later completely and we did, for instance, an analysis of the issue tracker and mailing list for the sake of the report in December*". Allowing students to be able to adjust the goals/requirements of their task without an effective guidance/monitor from the supervision team could therefore be a clear threat to the development of the students, including development of evaluative judgement.

4.2.3. Lessons Learned

5. Course Instance Fall 2016/2017

5.1. Course Design

5.1.1. Team Setup

The creation of teams changed only little compared to the year before. While the topics were no longer selected by students, they were still asked to indicate sub-teams up to 4 students, which were then mixed randomly by the teacher to for teams of 6-8.

6. Course Instance Fall 2017/2018

6.1. Course Design

6.1.1. Team Setup

To find another approach towards the problems in teams, the team size was reduced to 4-5 student per team. Furthermore, students got the chance to select their teams fully on their own.

6.1.2. In-course Components to Develop Student's Evaluative Judgement

6.1.3. Data Collection

6.2. Result

6.2.1. Teacher's Perspective

6.2.2. Student's Perspective

7. Conclusion

7.1. Do's and Don't

7.2. Threats to Validity

7.3. Future Work

[1] J. Cowan, Developing the ability for making evaluative judgements, Teaching in Higher Education 15 (2010) 323–334.

[2] D. R. Sadler, Beyond feedback: Developing student capability in complex appraisal, Assessment & Evaluation in Higher Education 35 (2010) 535–550.

[3] D. Nicol, Resituating feedback from the reactive to the proactive, in: Feedback in higher and professional education, Routledge, 2012, pp. 44–59.

[4] D. Nicol, Guiding principles for peer review: unlocking learners evaluative skills, Advances and innovations in university assessment and feedback (2014) 197–224.

[5] D. Boud, E. Molloy, Rethinking models of feedback for learning: the challenge of design, Assessment & Evaluation in Higher Education 38 (2013) 698–712.

[6] D. Boud, The use of self-assessment schedules in negotiated learning, Studies in higher education 17 (1992) 185–200.

[7] E. Panadero, G. T. Brown, J.-W. Strijbos, The future of student self-assessment: A review of known unknowns and potential directions, Educational Psychology Review 28 (2016) 803–830.

[8] E. Panadero, Is it safe? social, interpersonal and human effects of peer assessment, Handbook of human and social conditions in assessment (2016) 247.

[9] K. J. Topping, Methodological quandaries in studying process and outcomes in peer assessment, Learning and Instruction 20 (2010) 339–343.

[10] D. Boud, R. Soler, Sustainable assessment revisited, Assessment & Evaluation in Higher Education 41 (2016) 400–413.

[11] J. H.-M. Tai, B. J. Canny, T. P. Haines, E. K. Molloy, The role of peer-assisted learning in building evaluative judgement: opportunities in clinical medical education, Advances in Health Sciences Education 21 (2016) 659–676.

[12] J. Tai, R. Ajjawi, D. Boud, P. Dawson, E. Panadero, Developing evaluative judgement: enabling students to make decisions about the quality of work, Higher Education (2017) 1–15.

[13] D. Boud, N. Falchikov, Quantitative studies of student self-assessment in higher education: A critical analysis of findings, Higher education 18 (1989) 529–549.

[14] N. Falchikov, J. Goldfinch, Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks, Review of educational research 70 (2000) 287–322.

[15] R. Speyer, W. Pilz, J. Van Der Kruis, J. W. Brunings, Reliability and validity of student peer assessment in medical education: a systematic review, Medical Teacher 33 (2011) e572–e585.

14

[16] N.-F. Liu, D. Carless, Peer feedback: the learning element of peer assessment, Teaching in Higher education 11 (2006) 279–290.

[17] F. Dochy, M. Segers, D. Sluijsmans, The use of self-, peer and co-assessment in higher education: A review, Studies in Higher education 24 (1999) 331–350.

[18] I. Van den Berg, W. Admiraal, A. Pilot, Peer assessment in university teaching: evaluating seven course designs, Assessment & Evaluation in Higher Education 31 (2006) 19–36.

[19] T. McConlogue, But is it fair? developing students understanding of grading complex written work through peer assessment, Assessment & Evaluation in Higher Education 37 (2012) 113–123.

[20] D. Dunning, C. Heath, J. M. Suls, Flawed self-assessment: Implications for health, education, and the workplace, Psychological science in the public interest 5 (2004) 69–106.

[21] D. Carless, D. Salter, M. Yang, J. Lam, Developing sustainable feedback practices, Studies in higher Education 36 (2011) 395–407.

[22] S. Merry, M. Price, D. Carless, M. Taras, Reconceptualising feedback in higher education: Developing dialogue with students, Routledge, 2013.

[23] H. G. Andrade, Using rubrics to promote thinking and learning, Educational leadership 57 (2000) 13–19.

[24] R. J. Stiggins, Student-involved classroom assessment, Prentice Hall, 2001.

[25] J. A. Arter, J. Chappuis, Creating and recognizing quality rubrics, Prentice Hall, 2006.

[26] W. J. Popham, What's wrong-and what's right-with rubrics, Educational leadership 55 (1997) 72–75.

[27] P. Dawson, Assessment rubrics: towards clearer and more replicable design, research and practice, Assessment & Evaluation in Higher Education 42 (2017) 347–360.

[28] C. Reitmeier, L. Svendsen, D. Vrchota, Improving oral communication skills of students in food science courses, Journal of Food Science Education 3 (2004) 15–20.

[29] D. R. Sadler*, Interpretations of criteria-based assessment and grading in higher education, Assessment & Evaluation in Higher Education 30 (2005) 175–194.

[30] B. J. Reiser, Scaffolding complex learning: The mechanisms of structuring and problematizing student work, The Journal of the Learning sciences 13 (2004) 273–304.

[31] A. F. Wise, K. ONEILL, Beyond more versus less: A reframing of the debate on instructional guidance, in: Constructivist Instruction, Routledge, 2009, pp. 94–117.

[32] G. Joughin, D. Boud, P. Dawson, Threats to student evaluative judgement and their management, Higher Education Research & Development (2018) 1–13.

[33] D. Kahneman, P. Egan, Thinking, fast and slow, volume 1, Farrar, Straus and Giroux New York, 2011.

[34] G. B. Chapman, The psychology of medical decision making, Blackwell handbook of judgment and decision making (2004) 585.

[35] A. Tversky, D. Kahneman, The framing of decisions and the psychology of choice, science 211 (1981) 453–458.

[36] D. Lovallo, D. Kahneman, Delusions of success: How optimism undermines executives decisions. by harvard business school publishing corporation, 2003.

[37] M. J. Saks, R. F. Kidd, Human information processing and adjudication: Trial by heuristics, Law and Society Review (1980) 123–160.

[38] J. Kruger, D. Dunning, Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments., Journal of personality and social psychology 77 (1999) 1121.

[39] R. Thaler, Toward a positive theory of consumer choice, Journal of Economic Behavior & Organization 1 (1980) 39–60.

[40] S. D. Brookfield, Becoming a critically reflective teacher, John Wiley & Sons, 2017.

[41] J. Wolfe, E. Powell, Strategies for dealing with slacker and underperforming teammates in class projects, in: Professional Communication Conference (IPCC), 2014 IEEE International, IEEE, pp. 1–8.