ASRHE
ADVANCING SCHOLARSHIP & RESEARCH IN HIGHER EDUCATION

HERDSA

Advancing Scholarship and Research in Higher Education, 2021, 2(1) 1

Exploring course evaluations based on self-determination theory

Eva Heinrich
Massey University, New Zealand
e.heinrich@massey.ac.nz

Keywords
First year, course evaluation, self-determination theory

Key contributions/Pathways to collaboration

- The research demonstrates the use of self-determination theory (SDT) for course evaluations by combining scales for course design features, SDT, and students' perceptions of knowledge.
- Results from two first-year information technology courses indicate that the instrument provides valuable insights into course features and maturity.
- Invitations for collaboration are extended for comparison of results from other courses, fine-tuning of the SDT scale for contemporary technology-supported blended course contexts, large scale data collection to facilitate modelling and qualitative explorations to deepen understanding.

Abstract
This 'Research in Progress' article explores how to build on self-determination theory (SDT) in the context of first year university course evaluations. It suggests the use of an established SDT scale, Balanced Measure of Psychological Needs (BMPN), plus two course specific scales on course design features and perceived knowledge gains. The scales form a conceptual design linking course design, needs fulfilment and knowledge gains. Two information technology courses have been examined and findings indicate the usefulness of the conceptual design for course evaluation. Per course, data from the three scales combine to form a coherent picture; across courses differences show up that might be explained via course maturity. While approaches for a combined statistical analysis of the three scales are indicated, participant numbers

Please cite as: Heinrich, E. (2021). Exploring course evaluations based on self-determination theory, *Advancing Scholarship and Research in Higher Education*, *2*(1), 1-20.

were too small for testing of models. Several areas for follow up research are suggested and include refinement of the scales, application to other courses for comparison regarding course maturity or subject specifics, examination of SDT for application in modern technology-supported and blended course environments, as well as combination with qualitative data for a deeper understanding.

Introduction

Questions on how best to teach have been at the forefront of discussions in higher education for many years. Educators have looked for answers in 'blending', mixing on- and off-campus study, synchronous and asynchronous deliveries, pushing assessments online, using technologies for delivery, interaction and analysis. The 2020 global Covid-19 pandemic has amplified the challenges faced by teaching and learning, globally and across sectors, increasing the pressure to find solutions and to do so quickly.

My own teaching context lies in first year information technology. I am looking for ways to improve courses, to steer courses towards providing opportunities for diverse student populations, to create interest and enthusiasm for the subject area, and to facilitate development of intrinsic motivation and drive. Working towards this, I am investigating how we can measure if courses fulfil such goals and which aspects of course design and delivery contribute. The theoretical foundation for this work lies in self-determination theory (SDT). In brief, SDT posits that the satisfaction of needs leads to self-regulated motivation which in turn drives students to invest themselves in their learning. In first-year university courses, much importance lies in developing a love for the subject area, in learning to study, in pushing oneself instead of being driven by external factors. Therefore, I look at SDT to inform the design and delivery of courses and, in this research specifically, to facilitate their evaluation.

In the New Zealand context, a 'course' is a unit of study with a prescription, learning outcomes and assessment items. Typically, a full-time student studies eight undergraduate courses per year and 24 courses are required to complete a bachelor degree. For this 'Research in Progress' report I draw on data related to two courses I designed with SDT principles in mind. In 158.100 we ask students to conceptualise, specify and implement mobile apps using a graphical development environment (MIT App Inventor). As this is an introductory course without requirements for prerequisite

knowledge, our focus is on guiding students informally through a software engineering process and catering for students with different programming abilities. A central aspect is that each student decides on the focus of their app. With everyone working on their individual unique app, we can ask for sharing and peer feedback without compromising assessment. In 158.120 we explore web-based systems with the goals of gaining hands-on experiences with the technologies involved and understanding their interactions. We provide each student with their own Moodle site. Using a real and complex system for our explorations provides the interest factor and gives scope for beginners as well as more advanced students. My aim is that students leave these courses with foundational skills and a strong sense of confidence and enthusiasm, ready and keen to tackle subsequent learning.

Hattie (2015) describes the importance of teachers critically investigating their impact on students. My research aims to gain insight into the evaluation of courses, identifies what questions to ask and how best to learn from the answers. The research questions I am exploring in this article are:

- How can the use of a SDT survey for course evaluation help me understand if my SDT-based course design is successful?
- What measures might allow me to capture course design features and perceived knowledge gains in the course evaluations?
- How could the measures be used to compare courses while still capturing characteristics of individual courses?

Self-Determination Theory (SDT) and its relationship to learning and teaching

Self-determination theory (SDT) draws on a range of theories and has been discussed by many authors. Ryan and Deci (2002) is one of the main references. Heinrich and McDonald (2018) contains a summary written to underpin work on first-year teaching. SDT states that humans have basic psychological needs that must be fulfilled for growth and development, just as physiological needs require fulfilment to nurture physical development. While every person possesses an inherent drive, external influences are also crucial. External influences can be positive, promoting self-motivation and well-being, or detrimental, limiting initiative and causing negative experiences. The three basic psychological needs are competence, autonomy and relatedness. Competence is about opportunity to act according to one's own abilities; it encourages taking on development opportunities. Autonomy is about being able to

follow one's own interests and values; where it is necessary to carry out actions as instructed by others, autonomy is about being able to endorse these actions, to see their value. Relatedness refers to being connected to others, being part of a community, being accepted.

SDT relates closely to motivation theories. Ryan and Deci (2000, 2020) and Niemic and Ryan (2009) are key references in this area. These authors establish a continuum from non-self-determined to self-determined behaviour, from amotivation to extrinsic and intrinsic motivation. Extrinsic motivation is divided into four forms, working from external rewards and drives to internalizing reasons for doing something and matching those with one's own values and needs. Self-determined behaviour, based on intrinsic motivation or the more self-regulated forms of extrinsic motivation, bring benefits for effectiveness, persistence, well-being and social integration. SDT states that the fulfilment of needs for competency, autonomy and relatedness leads to higher rates of self-regulated motivation and therefore self-determined behaviour. To summarize SDT in relationship to learning and teaching: students experience better learning outcomes when they are given tasks at individually suitable levels of difficulty (fulfilment of the need for competency), have some control over how, when and what they study (fulfilment of the need for autonomy), and experience a positive connection to peers and teaching staff (fulfilment of the need for relatedness).

The three factors of competency, autonomy and relatedness are trait factors, meaning that they can change yet do not change quickly (Sheldon and Hilpert, 2012). Appropriate course design and support facilitates higher levels of needs satisfaction and with that a move towards the more self-regulated forms of extrinsic motivation and towards intrinsic motivation, leading in turn to better learning outcomes. In their higher education study on academic achievement and dropout intentions, Jeno at al. (2018) recommend

> teachers to support students' need for autonomy, competence and relatedness, by providing choice and volition to facilitate autonomous motivation, and give students effectance-relevant feedback and optimal challenges to increase perceived competence. (p1163)

While the development of SDT occurred in the 1970s, there was a strong uptake of the theory in the 2000s and more recently studies based in online learning environments

have been published. The impact of needs satisfaction and dissatisfaction on motivation and study outcomes has been proven for traditional teaching contexts. Studies aimed at assessing the validity of SDT in online contexts apply statistical methods to ascertain relationships, working towards verification of SDT in these newer teaching contexts. Those studies draw on participants across a range of courses without analysing course characteristics. Wang et al. (2019) show validity of SDT in online learning contexts. Hsu et al. (2019) state 'satisfying the three basic needs proposed in SDT can promote learning outcomes in online learning contexts' (p2172), just as this has been proven in numerous studies in traditional contexts.

A challenge with the work by Wang et al. (2019) and Hsu et al. (2019) is that the authors provide no explanation of 'online' learning. Some comments made in the articles suggest less than ideal online learning settings or a limited understanding of online. We need further research on the validation of SDT in technology-supported learning contexts that carefully addresses characteristics of these contexts. I proceed with the building on SDT in researching my courses based on my understanding of SDT and technology-supported learning and teaching. I acknowledge that the validity of the SDT in 'online' contexts is yet to be formally established.

Studies like the ones reported utilize large numbers of participants drawn from many courses. The focus of such work is on model verification and theory building. My focus is on the evaluation and improvement of individual courses, with a smaller number of potential participants and close examination of course characteristics. The following two studies demonstrate course-specific research based on SDT comparable to my approach. Scogin et al. (2015) researched a first-year programme aimed at providing an inclusive environment. They draw tentative conclusions on the characteristics of their programme and make recommendations with regards to needs satisfaction for the development of inclusive STEM environments. Bombaerts and Spahn (2019) present research on an ethics and history course for engineering students which had been redesigned based on SDT principles. They report on the value of SDT for course design and research.

Methodology

At the centre of my research design is the Balanced Measure of Psychological Needs (BMPN) scale reported in Sheldon and Hilpert (2012). I have chosen this scale in

preference to other SDT scales based on its balanced set of questions (six questions each for competency, autonomy and relatedness with three questions each targeting need satisfaction and dissatisfaction). Sheldon and Hilpert say that BMPN can be used in a variety of settings, from the more general to the particular. This fits my application in the context of specific courses. Considering that the link from needs fulfilment to self-regulated motivation, self-determined behaviour and positive outcomes has been well established (Ryan and Deci, 2000; Levesque-Bristol et al., 2006), I focus on measuring the level of needs fulfilment via the SDT scale. I do not measure motivation directly but suggest that needs fulfilment by itself will be a valuable indicator for positive influence on learning outcomes.

I created two additional scales, one with a focus on course design features, the other on students' perceptions of their learning. All three survey instruments use a 7-value low to high agreement scale. I used random ordering of questions for the BMPN scale to vary the sequencing across factors and dis/satisfaction constructs. The three scales were combined into one survey and distributed via the Qualtrics survey management tool. The idea behind the scale on course design features is to target elements relevant across courses (course materials, assessment design, interactions with peers, interactions with staff) while also giving space to address aspects specific to the design of individual courses. For 158.100 this specific aspect is that students work on individual apps, chosen and designed according to their personal interests. The key feature of 158.120 is that we work with Moodle as a real-life complex system, with each student controlling their own site. While the tasks are common, students set their own contexts via the content and design of their Moodle sites. I split the questions for these course elements in three groups, asking students for value in context of gaining marks, assisting learning and enjoying the courses. While gaining marks is important for passing a course, it can be different from assisting learning. As I see a crucial role of first year courses is in creating enthusiasm for the discipline studied and for learning more generally, I included asking for the enjoyment factor (for details of this and the other scales please see the tables in the findings section).

Other researchers, such as Hsu et al. (2019) and Wang et al. (2019), use course-independent scales when asking for perceived knowledge gain (e.g., asking how knowledge gained can be transferred without identifying the knowledge in more detail). My approach targets the specific learning objectives for each course and asks students

to assess their knowledge at the end of the semester, and, retrospectively, at the start of the semester. For 158.100 this resulted in six question pairs relating to software development skills and understanding as well as the importance of communication and working with peers. For 158.120 there were eight question pairs, targeting knowledge of the technologies covered in the course and the value of working with peers. Multiple factors contributed to asking students for their perceptions and to doing so only at the end of the semesters. Two separate surveys would likely reduce the number of responses available for analysis as this would require participation in both surveys. At the start of the course students are not familiar with the concepts underpinning the learning outcomes and have little basis for assessing their knowledge. Drawing on assessment results requires a relatively high level of ethics clearance and involves gaining institutional approval. Full assessment results are typically not available when course evaluations are run and might not necessarily map directly to learning outcomes. The accuracy of student perceptions of their own learning is typically high (Hattie, 2015). Figure 1 illustrates the conceptual design for the research.

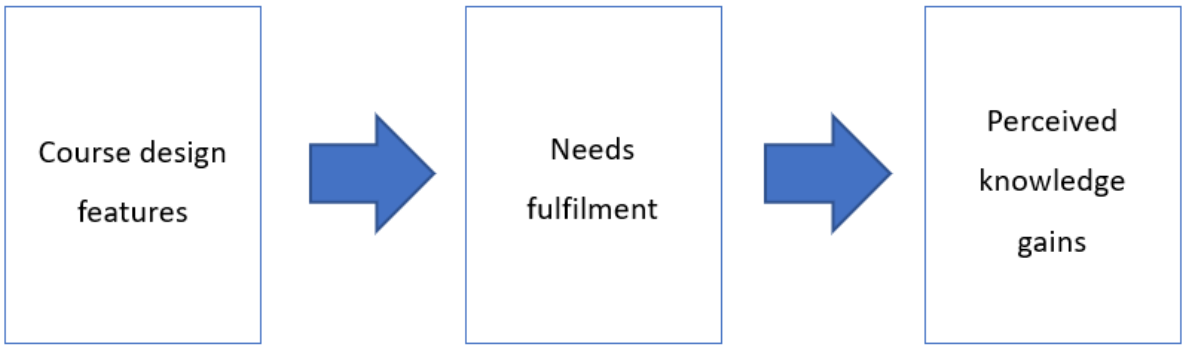

Figure 1: Conceptual research design

The major considerations for gaining ethics approval were the conflict of interest and with that the potential pressure on students due to my dual role as teacher and researcher. I completed a full ethics application (approval number SOA20/20) for the research on 158.100 as I included plans for student interviews (to be conducted by an external, independent researcher). For 158.120 I filed a ‘low risk notification’, as this research only consisted of an anonymous survey. For 158.120 I also ran my department’s standard course evaluation using the university’s procedures for handling the responses in anonymous form. My university’s ethics guidelines allow me to

analyse such course evaluation data for research purposes without further ethics approval (see Code of Ethical Conduct, Decision Chart). Appendix A shows a summary of the student responses on the standard course evaluation.

I have taught 158.120 since 2014. For the 2020 delivery I had redeveloped all material to refresh the technologies used (e.g., updating to the most current version of Moodle), making minor content changes but staying with the proven course design. 158.100 had undergone a major revision before the 2020 delivery. The core aspect of allowing students to choose their own app content was new and with that a range of assessment activities that asked students to peer review app designs and implementations. Both courses are delivered in parallel to on-campus and distance students via a shared learning management system site and are designed for delivery in blended format with heavy reliance on learning technologies. All lecture and demonstration materials are pre-recorded and available several weeks in advance of the study schedule. Distance students are supported via online tools (fora, chat systems, online meetings, emails, private messaging), for on-campus students we hold on-campus tutorials and labs. For 158.100 the ratio of distance to on-campus enrolments was about 2:1, for 158.120 about 4:1. Both courses were taught in 2020 during the Covid-19 pandemic. As most of the on-campus events were cancelled, effectively all students studied in distance mode. The pandemic affected the lives of many students, as was apparent in communication with students, a larger than usual number of extension requests and a higher than usual number of withdrawals and non-completions. From my perspective, the courses worked largely as normal for students who saw them through (the exception being a relatively small number of on-campus students who would have benefitted strongly from the on-campus sessions). Table 1 displays student and response numbers for the surveys. I have provided two student numbers per course: the official number of students enrolled at course conclusion; the number of students I judge as still involved in the course at the end of the semester based on submitting the last assessment item.

Table 1: Student and response numbers

Course	Official enrolments at end of course	Judged active at end of semester	Survey responses*	Course evaluation*
158.100	81	59	15 (25%)	*(not done)*

158.120	155	120	17 (14%) **	47 (39%)

* Percentages calculated against 'Judged active at end of semester' numbers
** 4 additional students filled out course design part of the survey

Findings

Tables 2 and 3 provide the responses to the course design feature questions. It stands out how much students value working with a real, complex system, which received the highest rating of the averages across the sub-questions (5.9, Table 3 for 158.120). Working in this real context (instead of building small system components for teaching purposes) sits at the centre of 158.120. It is pleasing to see that the students' evaluation highlights this central course design feature. The comparable aspect of 158.100, choosing one's own app, is also seen as valuable (4.9, Table 2 for 158.100) even if it does not reach the same height as the key feature of 158.120. On average, the ratings for the course design features of 158.100 are 0.7 points lower than for those of 158.120. This could reflect the more mature course design of 158.120, where materials, assessment items and interaction features have undergone several iterations. The design of 158.100 emphasizes students studying each other's work, creating learning opportunities and interactions. To achieve this, students have to provide feedback to their peers and reflect on this feedback.

The lower value for interactions with peers (3.7) matches comments made during the semester, with students having to wait for others and not always receiving full cooperation. The higher value for interaction with peers in 158.120 (4.3) might reflect the voluntarily exchange on an external chat system (Discord), which was heavily used by a substantial number of students, supporting each other in constructive and polite exchanges. My participation in this chat system might have contributed to the high rating (5.1) on the value of interaction with staff. The internal consistency across groups of questions is high, especially looking at course aspects (e.g., value of course material) across the sub questions that focus on marks, learning and enjoyment (0.86 to 0.97). This could indicate that the sub questions do not address different facets and could be combined. The questions on the value of being able to choose one's own app (0.51, Table 2 for 158.100) are an exception, with answers indicating differentiation.

Table 2: Course design features 158.100; scale from 1 to 7 (low to high agreement); means/standard deviations; 15 respondents

How valuable were the following 158.100 course aspects for …	**… gaining marks?**	**… assisting your learning?**	**… enjoying the course?**	***Average***	***Cronbach Alpha***
Choosing your own app	4.7/1.3	4.9/1.7	5.1/1.8	*4.9/1.1*	*0.51*
Course materials	4.7/1.8	4.7/1.8	4.1/2.1	*4.6/1.7*	*0.87*
Assessment design	4.9/1.3	4.3/1.5	4.1/2.1	*4.4/1.5*	*0.92*
Interaction with peers	3.9/1.9	3.7/2.1	3.5/1.8	*3.7/1.7*	*0.86*
Interaction with staff	5.1/1.6	5.1/1.7	4.71/8	*5.0/1.7*	*0.96*
Average	*4.7/1.1*	*4.6/1.4*	*4.3/1.6*		
Cronbach Alpha	*0.74*	*0.84*	*0.89*		

Table 3: Course design features 158.120; scale from 1 to 7 (low to high agreement); means/standard deviations; 21 respondents

How valuable were the following 158.120 course aspects for …	**… gaining marks?**	**… assisting your learning?**	**… enjoying the course?**	***Average***	***Cronbach Alpha***
Working on a real, complex system	5.7/1.3	6.1/1.4	6.0/1.3	*5.9/1.2*	*0.9*
Course materials	5.3/1.2	5.4/1.2	5.0/1.4	*5.3/1.2*	*0.9*
Assessment design	5.6/1.1	5.4/1.5	5.4/1.2	*5.5/1.2*	*0.87*
Interaction with peers	4.3/2.0	4.3/2.2	4.4/2.2	*4.3/2.1*	*0.97*
Interaction with staff	5.2/1.7	5.3/1.6	4.9/1.9	*5.1/1.7*	*0.96*
Average Mean	*5.2/1.1*	*5.3/1.1*	*5.1/1.1*		
Cronbach Alpha	*0.80*	*0.68*	*0.67*		

For the BMPN scale I followed Sheldon and Hilpert's (2012) suggestion to calculate six individual and three aggregate scores (based on reversing the scale for the dissatisfaction items; Table 4). I interpret the aggregated needs satisfaction scores as an indication that both courses fulfil students' needs. The scores for 158.120 are higher (relatedness 4.9, competence 5.4, autonomy 5.5) than for 158.100 (relatedness 4.7, competence 4.6, autonomy 4.8), potentially again reflecting the higher maturity of 158.120. As BMPN is a verified instrument it should show internal reliability. With Cronbach alpha values between 0.7 and 0.9 this is the case for most question

groupings. Exceptions are the dissatisfaction groupings for relatedness for both courses and some groupings for autonomy, with an extremely low Cronbach alpha value (0.05) for the aggregated need satisfaction score for 158.100. Dropping the first autonomy need satisfaction question (Q13) for 158.100 raises the Cronbach alpha value from 0.48 to 0.86 and the aggregated autonomy Cronbach alpha value to 0.39 (which is still low and does not improve much by dropping a further question). For the relatedness dissatisfaction the internal reliability improves by dropping one question each (for 158.100, Q4, 0.53 to 0.63; for 158.120, Q6, 0.59 to 0.70). Overall, both courses provide similar results for relatedness and competence but considerable differences for autonomy. The low reliability values for relatedness dissatisfaction for both courses could indicate that the questions need to be reviewed for suitability in online teaching settings. As the internal reliability for autonomy is strong overall for 158.120 but weak for 158.100 there could be issues with 158.100 and not with the BMPN items.

Table 4: BMPN, Balanced Measure of Psychological Needs scale

BMPN Items Scale from 1 to 7 (low to high agreement), means, standard deviations and Cronbach Alpha							
		158.100, n=15			158.120, n=17		
Relatedness							
1	I felt a sense of contact with people who care for me, and whom I care for.	3.6 1.8	3.2 1.5 0.77	4.7 1.0 0.70	4.2 2.1	3.5 1.8 0.90	4.9 1.2 0.79
2	I felt close and connected with other people who are important to me.	3.4 2.1			3.5 2.1		
3	I felt a strong sense of intimacy with the people I spent time with.	2.6 1.5			2.9 1.8		
4	I was lonely.	2.4 1.6	1.8 1.0 0.53		2.2 1.9	1.7 1.0 0.59	
5	I felt unappreciated by one or more important people.	1.5 1.1			1.6 1.1		
6	I had disagreements or conflicts with people I usually get along with.	1.6 1.4			1.4 0.7		
Competence							
7	I was successful completing difficult tasks and projects.	4.6 1.7	4.6 1.3	4.6 1.4	5.6 1.5	5.5 1.3	5.4 1.2

8	I took on and mastered hard challenges.	4.5 1.4	0.76	0.83	5.1 1.5	0.87	0.88
9	I did well even at the hard things.	4.9 1.5			5.6 1.3		
10	I experienced some kind of failure, or was unable to do well at something.	4.5 2.2	3.5 2.0 0.86		2.6 1.5	2.6 1.4 0.79	
11	I did something stupid, that made me feel incompetent.	2.8 2.1			2.4 1.9		
12	I struggled doing something I should be good at.	3.1 2.3			2.6 1.5		
Autonomy							
13	I was free to do things my own way.	5.3 1.8	4.6 1.2 0.48	4.8 0.7 0.05	5.2 1.3	5.3 1.1 0.75	5.5 0.9 0.76
14	My choices expressed my “true self.”	4.7 1.4			4.9 1.7		
15	I was really doing what interests me.	3.9 2.0			5.8 1.1		
16	I had a lot of pressures I could do without.	3.4 1.8	3.0 1.4 0.70		3.0 1.6	2.2 1.1 0.55	
17	There were people telling me what I had to do.	3.0 1.8			2.1 1.2		
18	I had to do things against my will.	2.8 1.7			1.6 1.2		

The third scale focuses on students’ perceptions of their knowledge (see Table 5 for 158.100 and Table 6 for 158.120). I have used the Welch t-test to determine if changes between the start and end of the semester are by chance or indicate significant improvements. The Welch t-test adjusts the degrees of freedom in response to unequal variances (see https://statistics.berkeley.edu/computing/r-t-tests). To cater for different levels of incoming knowledge I have calculated the degree to which students achieved the potential improvement. For example, for 158.100 students started with fairly low programming skills perceptions (2.4) leaving wide scope for improvements (4.6 point on the scale to reach 7). The difference between end and start of semester values (2.1) indicates a 46% achievement of the potential improvement (2.1 of 4.6). For 158.120 students have, in their own judgement, achieved 65% of possible knowledge gains. The t-test is significant for all knowledge areas (at 1% and 5% confidence levels). This compares to 25% perceived knowledge gain for 158.100, with only two significant t-tests. I see as possible explanations the more mature nature of 158.120, and the differences in the questions asked. While the questions for 158.120 ask for knowledge,

several of the 158.100 questions ask for confidence and understanding. A key design goal of 158.100 was to create learning opportunities via student interactions and peer feedback. The low value gain reported by students suggests the need to reconsider this design. Alarmingly, there is a weak indication saying that students have less confidence in their ability to improve their programming skills after completing the course. While statistically not significant, reducing confidence would be an unwanted course outcome and further investigation is required to determine if this measurement is a true reflection of student perceptions.

Table 5: Perceived knowledge gain, course 158.100

Perceived knowledge gain, 158.100	**… at the start of the semester**	**… at the end of the semester**	**Difference**	**Improvement of potential**	**T-test**
I have programming skills …	2.4	4.5	+2.1	46%	t=3.13 df=27.09 **p<0.01**
I am confident in my ability to improve my programming skills …	4.9	4.7	-0.2	-10%	t=-0.33 df=27.35 p=0.745
I understand the complexities inherent in developing software	3.3	4.8	+1.5	41%	t=1.99 df=26.30 p=0.057
I regard communication as an important factor in developing software …	5.2	5.7	+0.5	28%	t=0.75 df=27.50 p=0.46
I am confident in my ability to communicate with others in relation to software development …	3.1	4.3	+1.2	31%	t=2.21 df=22.60 **p<0.05**
I value working together with peers …	4.1	4.5	+0.4	14%	t=0.71 df=25.58 p=0.484
Average	3.8	4.8	0.9	25%	

Table 6: Perceived knowledge gain, course 158.120

Perceived knowledge gain, 158.120	**… at the start of**	**… at the end of the semester**	**Difference**	**Improvement of potential**	**T test**

	the semester				
I have system administration knowledge …	2.8	5.7	2.9	69%	t=5.40 df=29.08 **p<0.01**
I have database knowledge …	2.6	5.8	3.2	73%	t=6.12 df=30.7 **p<0.01**
I have web technologies knowledge …	3.3	6.1	2.8	76%	t=4.90 df=19.64 **p<0.01**
I have programming knowledge …	2.4	5.4	3.0	65%	t=5.00 df=28.74 **p<0.01**
I have networking knowledge …	1.9	5.3	3.4	67%	t=8.55 df=32.00 **p<0.01**
I have knowledge of working with virtual machines …	2.2	5.3	3.1	65%	t=0.709 df=25.57 **p<0.01**
I understand how web-based systems work …	2.8	5.8	3.0	71%	t=6.97 df=28.80 **p<0.01**
I value working together with peers …	3.8	4.9	1.1	34%	t=2.11 df= 29.30 **p<0.05**
Average	2.7	5.5	2.8	65%	

Discussion

My goal is to build on SDT in evaluating courses and establish a connection between course design, needs fulfilment and perceived knowledge gain, as shown in the conceptual diagram (Figure 1). I have evaluated data from the three scales separately and drawn comparisons for the two courses. As a teacher I want students to value the course components, have their needs for relatedness, competence and autonomy fulfilled, and achieve their learning potential. I suggest that the data for both courses indicate successful course design, albeit to a different degree. 158.120 rates higher across all measures than 158.100. I did not anticipate this, but it makes sense in hindsight. 158.120 is a mature course that has undergone improvements over seven years. From my perspective the course works very well, and this is confirmed by

student responses to my surveys and my department's standard course evaluation (Appendix A). My own reflections after teaching 158.100 told me that adjustments are required. The survey results mirror this, e.g., in the lack of internal reliability in the autonomy scale. While finetuning is required, I see promise in each of the three scales.

In response to my first research question, I suggest that the SDT scale has provided me with valuable insights into both courses. The data confirm some high-level course design goals and show where such goals are not met. Using the scale for more courses, ideally with a higher number of participants, will provide more insights on course design characteristics and comparison across courses. My scales for course design features focussed on four course-generic features (course material, assessment designs, interactions with peers/staff) and one course-specific feature each (choosing own app; working with real, complex system). These scales provided me with insights on the courses (e.g., about the value of working with a real system, about the challenges is designing peer-interactions into a first-year course) that match my own reflections on the courses. I found that I can use the course design related data and the data on perceived knowledge gain to compare between courses. These are indications that the scales chosen are suitable for answering my second and third research questions. Yet, again, more course settings and data need to be analysed to gain firm insights. For example, I cannot be sure if the sub-questions for the course design value questions (on gaining marks, assisting learning, enjoying the course) are useful in eliciting distinctions. An important area for further investigation is how to target the perceived knowledge gain questions. The questions I used addressed skills, understanding and confidence. The results varied, yet at this stage I cannot determine if this is a fair reflection of the course designs and where expectations should lie. Gaining skill might be an easier target that gaining understanding and confidence. A small improvement in understanding and confidence might be a substantial achievement for a first-year course.

I have focused on two courses with relatively low participant numbers. This has limited my analysis to descriptive statistics, applied separately to the three scales. Looking ahead to larger studies, I would like to build and test statistical models to see how course design features impact on needs fulfilment and this in turn on perceived achievement. A first approach could look at multivariate regression to see how predictors (the course design features) impact on responses (the needs factors). The

conceptual model (Figure 1) could be transformed into a structural equation model to understand connections across the three scales. These approaches require much higher participant numbers than I had (as a rough estimate, ten responses per variable). Such numbers might be achievable in large first year courses. An alternative approach could be to address a larger cohort, e.g., all first-year students in a college or wider discipline area. The course design scale could consist of the four generic measures, the perceived knowledge gain scale could target gains aspired in the transition to university education.

Conclusions and future work

In conclusion I suggest that my approach of building on SDT and combining this with scales on course design and perceived knowledge gain has promise for course evaluation. I focussed on a quantitative approach in this research in order to build on a verified SDT scale and to keep the students' time required for filling out the survey short. To gain a fuller understanding, this approach needs to be complemented with other data, gathered via freeform comments or focus groups. Attention also needs to be paid to future insights on the validity of SDT in technology-supported learning contexts. I was concerned about the wording of some of the statements included in BMPN (e.g., 'I felt a strong sense of intimacy with the people I spent time with.') and their appropriateness in a teaching context, yet decided against making modifications and used the instrument in the form statistically verified in previous research (despite Sheldon and Hilpert's, 2012, suggestion that item wordings can be modified slightly). Velde et al. (2020) is an example for a study that changed or omitted questions in an SDT scale the researchers regarded as not suitable for a teaching context. Other directions could include looking at SDT student profiles such as 'globally satisfied and highly connected' or 'globally dissatisfied, highly connected, and competence deficient' as identified by Gillet et al. (2020). Course designs and interventions might need to be tailored differently for various student groups. Henderlong Corpus et al. (2020) look at motivational change over a study period (semester or year), something one could build on in context of first-year transition.

I would welcome cooperation with other researchers on the following areas for future work:

- Application of the three scales: Application of the three scales to other courses might lead to the establishment of indicators (e.g., for course maturity) or

discovery of discipline characteristics. We could learn how to best formulate the context specific aspects of the course design and perceived knowledge gain questions.

- Modification of the SDT questions: I would like to work with others on adjusting the BMPN questions for teaching contexts. I could see us splitting a cohort answering the original and a modified version. We could pay particular attention to relatedness in technology-supported study environments.
- Large scale studies: Studies with a larger number of participants could lead to the development of statistical models, driving the understanding of the application of SDT and the other two instruments forward.
- Combination with qualitative approaches: For a deeper understanding we need to look at qualitative data. We could map student freeform comments to SDT needs to underpin the quantitative measures. We could cooperate in guiding focus groups with each other's students to overcome the challenges around ethics and power relationships.

References

Bombaerts, G., & Spahn, A. (2019). Simplify! using self-determination theory to prioritise the redesign of an ethics and history of technology course. *European Journal of Engineering Education*, 46(2). https://doi.org/10.1080/03043797.2019.1702924

Gillet, N., Morin, A.J.S., Huyghebaert-Zouaghi, T., Alibran, E., Barrault, S., & Vanhove-Meriaux, C. (2020). Students' Need Satisfaction Profiles: Similarity and Change over the Course of a University Semester. *Applied Psychology: An International Review*, 69(4), 1396-1437. https://doi.org/10.1111/apps.12227

Hattie, J. (2015). The Applicability of Visible Learning to Higher Education. *Scholarship of Teaching and Learning in Psychology*, 1(1), 79–91.

Heinrich, E., & McDonald, J. (2018). A proposal for student-centred first year teaching. Project Report, Ako Aotearoa. https://ako.ac.nz/knowledge-centre/learner-guidance-in-flexible-teaching/a-proposal-for-student-centred-first-year-teaching

Henderlong Corpus, J., Robinson, K.A., & Wormington, S.V. (2020). Trajectories of motivation and their academic correlates over the first year of college.

Contemporary Educational Psychology, 63. https://doi.org/10.1016/j.cedpsych.2020.101907

Hsu H-C.K., Wang C.V., & Levesque-Bristol, C. (2019). Reexamining the impact of self-determination theory on learning outcomes in the online learning environment. Education and Information technologies, Vol 24, 2159-2174. https://doi.org/10.1007/s10639-019-09863-w

Jeno, L.M., Danielsen, A.G., & Raaheim, A. (2018). A prospective investigation of students' academic achievement and dropout in higher education: A Self-Determination Theory approach. *Educational Psychology*, 28(9), 1163-1184. https://doi.org/10.1080/01443410.2018.1502412

Levesque-Bristol, C., Sell, G. R., & Zimmerman, J. A. (2006). A theory-based integrative model for learning and motivation in higher education. In S. Chadwick-Blossey & D. R. Robertson (Eds.), *To improve the academy* 24, 86-103. Boston, MA: Anker.

Niemic, C.P. & Ryan, R.M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*. 7(2), 133-144.

Ryan, R.M., & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*. 55(1), 68-77.

Ryan, R.M., & Deci, E.L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In *Handbook of Self-Determination Research*. E.L. Deci and R.M. Ryan (Eds.). The University of Rochester Press, Rochester, USA, 3-33.

Ryan, R.M., & Deci, E.L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contempory Educational Psychology*, Vol 61. https://doi.org/10.1016/j.cedpsych.2020.101860

Scogin, S.C., Austin, M., Alexander, C., Mader, C., Best, A., Rietberg, K., Supanich, K., Iaderosa, L., Sommavilla, & C., Barber, J. (2015). Supporting Students from Day 1 of College: The Importance of Relatedness to Inclusivity. *Journal of Microbiology and Biology Education*, 21(2), 1-10. https://doi.org/10.1128/jmbe.v21i2.2015

Sheldon, K.M., & Hilpert, J.C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. Motivation and Emotion, 36, p439-451. https://doi.org/10.1007/s11031-012-9279-4

Velde, R. van der, Blignaut – van Westrhenen, N., Labrie, N.H.M., & Zweekhorst, M.B.M. (2020). 'The idea is nice… but not for me': First-year students' readiness for large-scale 'flipped lectures'—what(de)motivates them? *Higher Education*. https://doi.org/10.1007/s10734-020-00604-4

Wang, C., Hsu H-C.K., Bonem, E.M., Moss, J.D., Shi, Y., Nelson, D.B., & Levesque-Bristol, C. (2019). Need satisfaction and need dissatisfaction: A comparative study of online and face-to-face learning contexts. *Computers in Human Behaviour*, 95, 114-125. https://doi.org/10.1016/j.chb.2019.01.034

Appendix A – Course evaluation results for 158.120

Course evaluation results 158.120 n=47/N=166 Distributed to 166 students Graded: 155 students Completed last assignment: 120 students Scale 1 to 6	Strongly disagree (1)	Disagree (2)	Tend to disagree (3)	Tend to agree (4)	Agree (5)	Strongly agree (6)	Average	# comments	# words
I am satisfied with the quality of this course.	3	0	2	10	16	16	4.8	6	146
I am satisfied with the course content.	2	3	1	9	16	16	4.7	7	223
I am satisfied with the course delivery.	2	3	3	9	15	15	4.6	10	247
I am satisfied with the assessment in this course.	0	0	3	8	12	23	5.2	10	460
The workload for this course is reasonable	0	0	2	3	22	20	5.3	7	144
The teaching team comes across as knowledgeable.	1	0	3	5	12	26	5.2	4	112
The teaching team comes across as approachable and respectful.	1	1	2	5	12	26	5.2	5	135
The teaching team comes across as enthusiastic and engaging.	2	0	6	5	17	17	4.8	6	79
What do you like best about this course?								42	1049
What do you like least about this course?								34	1053
Total								131	3648