



HAL
open science

Balance Talking and Doing! Using Google Design Sprint to Enhance an Intensive UCD Course

Marta Larusdottir, Virpi Roto, Jan Stage, Andrés Lucero, Ilja Šmorgun

► To cite this version:

Marta Larusdottir, Virpi Roto, Jan Stage, Andrés Lucero, Ilja Šmorgun. Balance Talking and Doing! Using Google Design Sprint to Enhance an Intensive UCD Course. 17th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2019, Paphos, Cyprus. pp.95-113, 10.1007/978-3-030-29384-0_6 . hal-02544605

HAL Id: hal-02544605

<https://inria.hal.science/hal-02544605>

Submitted on 16 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Balance Talking and Doing!

Using Google Design Sprint to Enhance an Intensive UCD Course

Marta Larusdottir¹, Virpi Roto², Jan Stage³, Andrés Lucero², Ilja Šmorgun⁴,

¹ Reykjavik University, 101 Reykjavik, Iceland

² Alto University, FI-00076 Aalto, Finland

³ Ålborg University, DK- 9100 Ålborg, Denmark

⁴ Tallinn University, 10120 Tallinn, Estonia

marta@ru.is

Abstract. Design, evaluation and enhancement of teaching activities in user-centred design (UCD) is characterized by limited research. This is particularly paradoxical as effective high-quality teaching is a key prerequisite for professional work in UCD. This paper reports the development of a two-week intensive UCD course for university-level students in an international setting. The first edition of the course ran during the summer of 2017. Based on both qualitative and quantitative data collected from students, the course was enhanced and a new edition that introduced Google Design Sprint (GDS) was conducted during the summer of 2018. Similar student feedback data was collected during both years (i.e., 2017 and 2018). In both editions, the course included lectures and hands-on use of UCD and interaction design methods in a design assignment. In this paper, we focus on the 2018 edition of the course and the students' qualitative and quantitative feedback on that edition. The results illustrate that students liked the intensive teamwork, clear structure, and the international setting of the course. The main concerns from the students were on inefficient time management and the lack of user involvement in GDS. However, GDS was preferred to the traditional teaching methods, as the students saw the rapid development cycle to provide a good balance of talking and doing.

Keywords: User-Centred Design Education, Students Feedback, Google Design Sprint.

1 Introduction

While research in Human-Computer Interaction (HCI) and User-Centred Design (UCD) is characterized by extensive emphasis on concepts, methods and their use in practice, much less attention has been devoted to the teaching of these topics, although well-trained practitioners is a prerequisite for successful use in practice. It is a paradox that HCI and UCD research with its strong emphasis on concepts and methods for evaluation has very little concern for the teaching of these topics, not to mention the assessment of such teaching activities.

Assessment has been a key topic for many years in general research on teaching. Cronbach was one of the early proponents of using assessment or evaluation as the basis for improving teaching activities. His aim was to use evidence gained through evaluation as the basis for developing and improving educational programs [4; 5]. A recent example of using quantitative data to improve teaching activities is provided by Aggarwal and Lynn [1] as they described how they improved a database management course through collection and interpretation of quantitative data. Quantitative data was also collected through questionnaires in a three week intensive course focusing on measuring the outcomes of learning outcomes [16].

Up to and during the 1990s, the field of teaching evaluation was re-shaped from this quantitative foundation by two parallel developments that were based on qualitative methods with a strong emphasis on context. First, the case study method was introduced to achieve a stronger focus on context of the teaching activities, and this was often combined with a participatory approach to teaching evaluation. Second, action research was introduced for “teachers [to] experiment in their own classrooms with how to improve teaching performance and enhance student learning” [17]. The effect of using practical exercises in a human-computer interaction course was studied recently and compared to using a more typical approach [20]. The results show that students involved in realistic projects are significantly more motivated and perceive that the HCI activities are more useful or important than students involved in a more general approach, with not as realistic projects to solve.

While there is significant research on the introduction of UCD in software organizations, few authors have discussed the design of courses on UCD. Seffah and Andreevskaia [15] describe how they developed a course in UCD for software practitioners. They developed the course over a period of eight years, and their approach embodies continuous improvement which is based on qualitative techniques such as class discussions combined with interviews and observation. In another study, students that had taken at least one master's level course in usability and user centered design (UCD) answered a survey to assess the value of a teaching philosophy that considered usability skills to be of value to future information professionals, even when they are not pursuing careers as usability engineers [2]. The survey results show that almost 95% of respondents regularly used the general principles of usability on the job, despite only 20% were hired to perform user centred design.

Over two years, we have developed a two-week course on UCD for university-level students. We gave the course for the first time in 2017 and evaluated it with 18 participants who had different nationalities. This has been reported with focus on the contents of the course, cf. [11]. Based on the evaluation, we redesigned the course, and in 2018 we gave it to a new group of 19 international students.

This paper reports from a case study of the evaluation of the second edition of the course that was given in 2018. The purpose of the paper is both to present the evaluation of the course and to provide a framework for evaluation of similar courses. In the following sections, we present the background for this paper by describing the overall rationale of the course, the first edition of the course briefly, the Google Design Sprint (GDS) process and the schedule of the course in 2018.

2 Background

In this section we describe the overall rationale for this UCD course and give an overview of the 2017 edition. We introduce the GDS process briefly, which was used in the 2018 edition of the course, and describe the structure of that edition of the course.

2.1 Overall Rationale for the Course

Over the past three decades we have witnessed shifts and re-framings in just about every area of interaction design: how it is done, who is doing it, for what goals, and what its results are. These changes show shift from designing things to designing interactions, first on micro-level and lately also on a macro level; and from designing for people to designing with people and very recently, to designing by people.

The university-level course targets higher education students in various fields and provides interaction design understanding and skills to new, but highly interested audiences. Additionally, we targeted non-ICT (Information and Communications Technologies) professionals. Upon completion of the course, higher education students and professionals should be able to conceptualize and prototype digital artefacts ranging from simple web-based services and small applications to wearable computing solutions and public space installations. The course was given in two weeks as a 4 ECTS intensive course, brought together, delivered, and hosted on rotation, by four partner universities: Aalto University, Ålborg University, Reykjavik University and Tallinn University.

2.2 Overview of the 2017 Edition of the Course

The first edition of the User-Centred Design Course was given at Tallinn University in 2017. It lasted for two weeks, Monday to Friday, between July 24 and August 4. The main learning objective of the course was that students would gain the ability to apply common user-centred design methods and interaction design tools in practice during a two-week intensive course. A total of 18 international students worked on designing and evaluating a software system, and used altogether 15 UCD methods along the way. The students worked in groups of three to four students, which were formed by the lecturers. Students brainstormed ideas for five different systems using similar methods for analyzing, designing and evaluating the system prototypes.

During the first three days the students were introduced to the following user-centred design methods: Visioning, Contextual interviews [7], Affinity diagram [12], walking the wall, personas and scenarios [7]. After an introduction of the method, the students used each of the methods with supervision from the lecturers. The next two days the students were introduced to user experience (UX) goals [9], they made low-fidelity paper prototypes of the user interface and evaluated those through heuristic evaluations [14]. They also used the System Usability Scale (SUS) questionnaire for evaluation [3] and additionally evaluated the interface according to their UX goals. During the second week the students were introduced to formal usability evaluations.

Students then prototyped the interface using the Just-in mind prototyping tool¹ and did an informal think aloud evaluation of that prototype. After redesigning the prototype, the students stated measurable usability and UX goals and made a summative user evaluation to check the measurable goals. At the end of the course all students gave a 15-minutes presentation to the class where they presented their work to each other. The methods introduced to the students were chosen partly based on results on what methods Information Technology (IT) professionals rate as good methods for UCD [8].

The feedback from the students was largely positive. Students liked being active during the class, since they used various methods during the class hours and could get guidance from the teachers. Students especially liked making the hi-fi prototypes and gave that method the highest rating. Students also liked working with international students with various backgrounds. Students disliked a mismatch between the description and the actual content during the course. The hi-fi prototyping started quite late, day seven, and it was difficult to get the prototype ready in time. The students also commented that they met real users quite late in the course, which was on the ninth day, and would have liked that to happen earlier in the course. They had used many UCD methods, but in those, they had “played” users for each other, so students from other groups participated in UCD activities.

With this feedback it was decided to structure the next edition of the course according to the Google Design Sprint (GDS) process. GDS offers a well-structured interaction design process, with one activity feeding into the next. Such a framework is especially important for the main target audience of the course - those encountering the interaction design process for the first time. As a result, the students would have a clearly defined process to follow that would allow them to reach tangible results fairly quickly, while also leaving the time for user evaluation.

2.3 Introduction to the Google Design Sprint

Created as a means to better balance his time on the job and with his family, Jake Knapp optimized the different activities of a design process by introducing a process called the Google Design Sprint (GDS) [10]. Knapp noticed that despite the large piles of sticky notes and the collective excitement generated during team brainstorming workshops, the best ideas were often generated by individuals who had a big challenge and not too much time to work on them. Another key ingredient was to have people involved in a project all working together in a room solving their own part of the problem and ready to answer questions. Combining a focus on individual work, time to prototype, and an inescapable deadline Knapp called these focused design efforts “sprints”.

The GDS is a process to solve problems and test new ideas by building and testing a prototype in five days. The main premise for the process is seeing how customers react before committing to building a real product. It is a “smarter, more respectful, and more effective way of solving problems”, one that brings the best contributions of

¹ Just-in Mind. <https://www.justinmind.com/>

everyone on the team by helping them spend their time on what really matters [10]. A series of support materials such as checklists, slide decks, and tools can be found on a dedicated website².

An important challenge is defined, small teams of about seven people with diverse skills are recruited, and then the right room and materials are found. These teams clear their schedules and move through a focused design process by spending one day at each of its five stages (i.e., map, sketch, decide, prototype, test). On Monday, a map of the problem is made by defining key questions, a long-term goal, and a target, thus building a foundation for the sprint week. On Tuesday, individuals follow a four-step process (i.e., notes, ideas, crazy 8s, and solution sketch) to sketch out their own detailed, opinionated, and competing solutions. On Wednesday, the strongest solutions are selected using a structured five-step “Sticky Decision” method (i.e., art museum, heat map, speed critique, straw poll, and supervote) and fleshed out into a storyboard. On Thursday, between one and three realistic-looking prototypes of the solutions proposed in the storyboard are built, using tools like Keynote to create the facade for apps and websites, a 3D printer to quickly prototype hardware, or just build marketing materials. Finally on Friday, the prototype is tested with five target customers in one on one interviews or think-aloud sessions. While only some of the resulting solutions will work, going through such sprints provides clarity on what to do next after spending only five days tackling that big important challenge.

While GDS is oriented towards quickly achieving tangible results and experimenting with a number of potential design solutions, it has some limitations that one needs to be aware of. The first one being that GDS assumes prior knowledge and familiarity with the context and user requirements. Thus, all background research should be done before GDS starts, as the process is more focused on putting together a team of experts, who can quickly propose workable solutions to the formulated problem. This also means that core UCD activities, such as the creation of personas and user scenarios, need to be carried out before hand, and later be used as input and guidance for the subsequent design activities. The second one being that although GDS includes basic user testing on the fifth day, more thorough evaluation might be necessary for achieving better results. Subsequent iterations could be planned to then incorporate the concerns the users voice during evaluation.

The GDS on our course differed a bit from the above in the initial preparations, as there were four to five students in one team. Each team was asked to create a small software application, designed either for a mobile device or a bigger screen. Each team formulated their own specific topic for the design exercise.

2.4 The Course Schedule in 2018

The User-Centred Design course given in 2018 at Reykjavik University, lasted for two weeks, Monday to Friday. Similar to the 2017 edition, the main learning objective of the course was the ability to apply common UCD methods and interaction design tools in practice during a two-week intensive interaction design sprint.

² Sprint book website. <https://www.thesprintbook.com>

A total of 19 international students worked on designing and evaluating a software system in four groups of four to five students, which were formed by the lecturers before the beginning of the course. We applied a similar strategy while forming groups with varying backgrounds, gender, and nationalities in each group as in 2017. Five potential project ideas were suggested to the students but they were told that they could also brainstorm ideas themselves for the systems to be designed and evaluated during the course. Students worked on four different software systems ideas and used altogether 13 methods for analyzing, designing and evaluating the systems prototypes. The course schedule is illustrated in Figure 1. The lectures are shown in bold text and the methods that the students practised are shown in italics. The results on the student evaluations of the methods are shown in Table 3, in the result section of this paper.

| Week 1 | Monday | Tuesday | Wednesday | Thursday | Friday |
|------------------|--|--|---|---|--|
| Morning | Introduction lectures <i>Discuss project ideas</i> | <i>Ask the experts</i> Interaction design lecture | <i>Voting on design</i> <i>Speed critique</i> Storyboard lecture | <i>Hi-fi prototyping</i> | <i>User evaluations</i> |
| Afternoon | <i>Making a map</i> | Sketching lecture <i>Lightning demos</i> <i>Sketching-crazy 8</i> | <i>Making a storyboard</i> Prototyping lecture | <i>Hi-fi prototyping</i> User evaluation lecture | <i>Interpret evaluation results</i> <i>Prepare students presentations</i> |
| Week 2 | Monday | Tuesday | Wednesday | Thursday | Friday |
| Morning | <i>Student presentations</i> Setting UX goal lecture | <i>Prototyping for evaluation of UX goals</i> | Evaluation lecture <i>Choosing a target for the last evaluation</i> | Agile UCD lecture <i>Preparing summative UX evaluations</i> | <i>Preparing presentations</i> |
| Afternoon | <i>Setting UX goals</i> | <i>Evaluating against the UX goals</i> | <i>Prototyping for the last evaluation</i> | <i>Summative UX evaluations</i> | <i>Final presentations & Wrap up</i> |

Fig. 1. The schedule of the UCD course in 2018. The text in italics explains hands-on activities.

The course schedule focused on running the GDS during the first week by conducting all the GDS methods in that process “by the book”, following the checklists and descriptions in the process. Typically, there was a short lecture explaining how to use each method and right afterwards the students got one or two hours to practice that method under the supervision of the lecturers. During the second week user experience aspects were added to the design and the prototype was redesigned and evaluated with users to understand the user experience better of the prototype and the system idea. So in the second week, more emphasis was on lecturing and practicing user-centred design methods.

Upon completion of the course students should have acquired an understanding of what design is and should grasp the full cycle of the design process including the stages of discovering, defining, developing and delivering concepts targeting areas of their interest. There was a continuous assessment of the learning outcomes by observing the students while they practised the methods introduced in the course.

3 Method

In this section we describe the students taking part in the course in 2018 briefly, the course evaluation methods and the data analysis methods.

3.1 Students

Nineteen students participated in the course, 16 females and 3 males. Students living in Denmark, Estonia and Finland were selected and received grants to come to Iceland to participate in the course. Some of those students were originally from other countries, so we had participants from: Iceland (4), Estonia (4), Denmark (3), Poland (2) and one from each of the following countries: Belarus, Greece, Spain, Mexico, Russia and Vietnam. The participants were between 23 and 37 years old.

The participants had various backgrounds. Information is missing from three students, so the background information is based on 16 answers. Three students had a high school degree and were studying for BSc degree; eight had a Bachelor degree and were all studying on Master level, four had a Master degree, where three were studying further and one had a PhD degree. Several of the students had humanities and social sciences as their discipline, some had design sciences, and other computer science, technical science and engineering. Three students were not studying at the time of the course. Eleven students had some experience from working at a software company/organisation and the time varied from two to 36 months.

When the students registered to the course, they responded to question “What do you expect to learn on this course?”. Out of 15 responses, six spontaneously mentioned Hands-on experience, and four specifically mentioned (Google Design) Sprint, four Prototyping and two Programming. Experiencing the full design process came up in six responses. The concepts mentioned were Interaction Design (5), UCD (3), User Interface (3), UX (2) and Usability (1). Evaluation was mentioned twice. In addition to learning these hard skills, collaborating with students from other cultures and disciplines was mentioned by five participants.

3.2 Course evaluation methods

Two data gathering methods were used to gather feedback from students on their opinions on the course. The Retrospective Hand technique was used as a weekly evaluation for collecting open-ended feedback from the students and a questionnaire form on the methods taught was used in the last session of the course. Both the questionnaire and the Retrospective Hand were distributed on paper. The data gathering methods will be described in more detail below.

The Retrospective Hand Technique. Students were asked to draw their right hand on an empty A4 sheet of paper as the last thing in the afternoon during both Fridays, so data was gathered twice with this technique during the course. In the space near the thumb, they were asked to write what they thought was good during the current week

(i.e., thumb raised), in the space for the index finger, things they wanted to point out (i.e., indicate), in the third finger space what was not good (i.e., middle finger), in the space for the fourth finger what they will take home with them (i.e., ring finger), and the fifth finger what they wanted more of (i.e., pinky finger). The students wrote sentences in free text, so this was a qualitative technique. To keep anonymity, students handed in their feedback paper by putting it in a box that was placed at the back of the room, so that lecturers could not see who were returning the evaluation forms. When all the students had handed in their evaluations, we asked if there was something that they wanted to share with the group. There were open discussions for about 15 minutes of improvements that could be made to the course, but these discussions were not analysed for this publication.

The idea of this technique comes from industry and has been used by the first author on four different courses. The students like the method, since it has an open and somewhat creative format, so they can comment on various issues and it takes them around 10 minutes to complete. When used in the middle of the course, the instructors of the course have the possibility to respond to the comments of the students, and make enhancements accordingly, which the students appreciate.

The Method Questionnaire. The questionnaire was on paper and contained three pages. On the first page there were: four questions on the student's background, three questions on their currently highest achieved degree, one question on whether they were studying currently or not, and three on their current education field (if applicable). Also, on the first page they were asked if they had worked in a company/organisation developing software systems. If so, they were asked to fill in four more questions about the work role and company.

On the second page of the questionnaire the students were asked to rate their opinion of the 13 GDS/UCD methods used in the course. The first nine methods were all from the GDS process and four methods were more typical UCD methods. For each method they were asked to rate:

- a) If the method was thought provoking;
- b) If the method was useful for the course;
- c) If they thought that the method would be useful for their future job/education.

For each item we provided a 7-point scale from 1 = not at all to 7 = extremely so. The 13 GDS/UCD methods they evaluated were: *Making a map*, *Ask the experts*, *Lightning demos*, *Sketching (including crazy 8)*, *Voting on design solutions*, *Speed critique of the designs*, *Storyboard making*, *Hi-fi prototyping*, *User testing (of the hi-fi prototypes)*, *Setting UX goals*, *Evaluation of UX goals*, *Prototyping for the last evaluation*, *Summative UX evaluation*. Furthermore, students were asked to rate the whole GDS process (used during the first week) and the inclusion of the user aspects (the focus of the second week).

On the third page, there was just one open question for any other comments that they would like to share with us. They had a full A4 page to freely share their comments. Some student used the whole page to write detailed comments.

The questionnaire was filled in right after the retrospective hand evaluation during the last session of the class. The students typically used 20 minutes to fill in the questionnaire. When all the students had filled in this questionnaire, a group discussion was facilitated on the overall evaluation on the course and notes were taken by the lecturers on their comments.

3.3 Data Analysis Methods

The data from 2018 using the Retrospective Hand technique was analysed according to theme analysis [6]. The themes we used are shown in Table 1.

Table 1 The categorization themes from Steyn et al. used in this study.

| # | Theme | Explanation according to Steyn et al. |
|---|-----------------------|---|
| 1 | Assessment | Assessment standards, structure, schedule, criteria and feedback. |
| 2 | Staff quality | Tutor and lecturer availability, teaching skills, quality and frequency of communication with students and the number of lecturers on the module |
| 3 | Learning environment | Quality of lecture hall equipment, the size and comfort of lecture and tutorial venues and the quality of the learning environment created through group project work. |
| 4 | Learning support | Adequate preparation for assessments and the provision of additional learning support, such as workshops and guest lecturers. |
| 5 | Learning resources | The provision of additional resources and the quality, timeliness and affordability of resources provided. |
| 6 | Teaching methods | <i>Steyn et al. definition:</i> Suggestions to make the lectures more interactive and application-based. <i>Our definition:</i> HOW the students learn, i.e., the format of activities on the course. The outcome of the design exercise. |
| 7 | Course content | <i>Steyn et al. definition:</i> Relevance of the curriculum, and the workload. <i>Our definition:</i> WHAT the students learn. |
| 8 | Course administration | <i>Steyn et al. definition:</i> Class scheduling. <i>Our definition:</i> Practical arrangements such as team formation and availability of practical course information. |

We based our themes on themes suggested by Steyn et al. [18] as shown in Table 1, but we adjusted some of the definitions to the characteristics of our course. Soon after starting to analyse the data, we noticed many data items did not fall to any of the

themes in Table 1, and we needed more themes for the thematic analysis. The new themes are shown in Table 2.

Table 2 Additional themes added by the authors.

| # | Theme | Explanation |
|----|------------------|---|
| 9 | Course structure | Structure and scheduling of the activities, days and the course. WHEN the learning activities take place . |
| 10 | Soft skills | Critical thinking, problem solving, leadership and responsibility, communication, and collaboration (e.g., team work) |
| 11 | People | Personal relationships, selection of people on the course. |
| 12 | Experience | Overall course experience, level of motivation, atmosphere, freetime activities, lunch and snacks, accommodation |

We had to add the themes because our course was very different from Steyn et al. This intensive course ran for two weeks and many of the students comments were on the course schedule, so we decided to include a separate theme on Course structure. Our course included many problem-solving activities in teams, so we created a new theme for Soft skills. An intensive course with international students is a memorable experience, therefore themes for the comments regarding Experience and People were added.

The first two authors of the paper individually analysed the data according to the above themes. When both authors had analysed all the data, the inter-evaluator agreement was calculated to be 58%. Then each author re-evaluated their theme analysis in light of the other's category suggestion. After this, the evaluators discussed the disagreements until reaching consensus.

4 Results

In this section we first describe the results on the quantitative ratings from students and then the results on the qualitative feedback gathered.

4.1 Quantitative Ratings from Students

Rating of Methods: The average numbers on how students rated the methods and the focus of each week are summarized in Table 3 on a scale from 1 to 7, where 1 was “not at all” and 7 was “extremely so”.

The students really liked the whole process of the GDS. It got the highest rating of all the course content in the questionnaire, for being thought provoking, useful in the course and useful in the future. Out of the GDS methods used during the first week,

the students gave the highest ratings to the *Sketching* method, but they also rated the GDS methods, *Making a map* and *Making a storyboard* high.

Table 3. The average quantitative rating of the methods from students in 2018.

| GDS/UCD Methods and the focus each week | Thought provoking | Useful in the course | Useful in the future |
|---|-------------------|----------------------|----------------------|
| Making a map | 5,94 | 6,06 | 6,00 |
| Ask the experts | 4,81 | 4,44 | 5,63 |
| Lightning demos | 5,31 | 5,00 | 5,00 |
| Sketching (incl. Crazy 8) | 6,38 | 6,63 | 6,25 |
| Voting on design solutions | 5,00 | 6,13 | 5,56 |
| Speed critique of the designs | 5,63 | 6,00 | 5,44 |
| Making a Storyboard | 6,00 | 6,31 | 6,38 |
| Hi-fi prototyping | 5,50 | 6,25 | 6,50 |
| User testing – 1 st week | 5,79 | 6,07 | 6,06 |
| Setting UX goals | 4,56 | 4,31 | 5,00 |
| Evaluation against the UX goals | 4,63 | 4,50 | 5,13 |
| Prototyping for the last evaluation | 4,81 | 5,50 | 5,31 |
| Summative UX evaluation | 5,19 | 5,06 | 5,25 |
| Whole Google Design Sprint (1st week) | 6,69 | 6,75 | 6,63 |
| Including the user aspects 2 nd week) | 5,06 | 5,13 | 5,40 |

The GDS/UCD method with the lowest numerical scores was *Setting UX goals* in the second week. Students were asked to do that during Monday morning in the second week, after being very productive during the first week. The structure of this session was maybe not as clear as for the sessions in the first week, so the students got all very tired. Many students commented that because of the looser structure they felt disconnected and not as motivated. There is clearly a challenge there, how to keep the motivation for the students from the first to the second week, and how user-centered design can best be included in the GDS process.

After the two weeks many students reflected that maybe it would have been good to have some team building activities and user research activities before beginning the GDS process. One student commented that because there were no user research activities before starting the GDS process, she felt like cheating. Actually the group the student was in, did describe one persona and one scenario, even though they were not instructed to do so. Some students also commented that the course could have been shorter, so the activities during the second week felt not as important as during the first week.

Comparing the results from 2018 to the results from 2017: In table 4, the ratings from the students on GDS/UCD methods used both during the course in 2017 and 2018 are compared. In 2017 the UCD methods used during the first days of the course were the methods suggested by Holtzblatt et al. [7] in the Rapid Contextual Design process. In 2018, we followed the GDS process for the first week [10]. Some of the

methods in these two processes are similar and can be compared. *Visioning* (in 2017) and *Making a map* (in 2018) have the same objective of giving an overview of the vision for the whole system. *Low-fi prototyping* on paper (in 2017) and *Sketching (including the Crazy 8)* according to GDS (in 2018) are also very similar. *Hi-fi prototyping* in 2018, was highly similar as in 2017, but in 2018 we followed predefined roles in the process according to the instructions in the GDS process. *User testing* of the prototypes was done in the same way in 2017 and 2018, by conducting think-aloud sessions with five users. *Setting UX goals* was also done in the same manner and taught by the same person both years, the *Evaluation of the UX goals* was also conducted in the same way. Additionally, the *Summative UX evaluation* at the end of the course was also conducted similarly.

Table 4. Average Quantitative Rating from Students in 2017 and 2018 (standard deviation).

| GDS/ UCD method | Thought provoking | | Useful in the course | | Useful in the future | |
|---------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| | 2017 | 2018 | 2017 | 2018 | 2017 | 2018 |
| Visioning/Making a map | 4,6 (1,68) | 5,9 (1,06) | 5,4 (1,69) | 6,1 (1,12) | 5,1 (1,86) | 6,0 (0,89) |
| Low-fi prototyping/Sketching | 4,9 (1,76) | 6,4 (1,02) | 5,9 (1,61) | 6,6 (0,62) | 5,5 (1,92) | 6,3 (1,06) |
| Hi-fi prototyping | 5,8 (1,42) | 5,5 (1,15) | 6,8 (0,53) | 6,3 (1,13) | 6,8 (0,71) | 6,5 (0,97) |
| User testing (hi-fi prototypes) | 5,6 (1,11) | 5,8 (0,81) | 6,2 (1,12) | 6,1 (0,63) | 6,2 (1,13) | 6,2 (0,50) |
| Setting UX goals | 5,7 (0,87) | 4,6 (1,21) | 6,2 (1,12) | 4,3 (1,01) | 6,3 (1,11) | 5,0 (1,21) |
| Evaluation of the UX goals | 5,0 (1,29) | 4,6 (1,02) | 5,4 (1,64) | 4,5 (1,32) | 5,1 (1,66) | 5,1 (1,09) |
| Summative UX evaluation | 5,7 (1,20) | 5,1 (1,22) | 6,3 (1,00) | 5,1 (1,24) | 6,4 (0,90) | 5,4 (1,48) |

In Table 4, it is interesting to see how the ratings of the methods differ between the two years. There is a significant difference between how the students rated how thought-provoking the two methods *Visioning* (used in 2017) and *Making a map* (used in 2018) are (t-test, $p=0.01$) (shown in bold in the table). The reason for the higher rating could be that students received more precise instructions on the GDS method *Making the map*, than on *Visioning* the year before. But, there was not a statistical difference in the ratings of the other two aspects, how useful the methods were in the course and how useful the methods will be in the future.

Sketching was done more individually in 2018, as instructed by the GDS, than in 2017 when students did the *Low-fi prototyping* in groups. There is a significant difference between how the students rated how thought-provoking these two methods are (t-test, $p=0.01$). There was not a statistical difference in the ratings of the two other

aspects, how useful the methods were in the course and how useful the methods will be in the future.

The difference between ratings of the UCD method *Setting UX goals* in 2018 and 2017 was significant for the three aspects, if it was thought provoking, useful in the course and useful in the future (t-test, $p=0.01$). The method was rated lower in 2018 than in 2017, when it was one of the most preferred methods. There was not a big difference in how the method was taught on the two courses, so the reason for the difference may be in the scheduling and structure of this session. The students had been very active during the first week in 2018 and managed to make hi-fi prototypes of their brand new idea and evaluate it in one week. Then on Monday morning they were back to a traditional lecture + team exercise structure. Setting UX goals require in-depth consideration and discussion before they can be taken into use. That activity was not as straightforward and fast-paced as the methods in the GDS. It seemed to be hard for the students to switch from the pace of the GDS to a more open and less guided structure.

The method *Summative UX evaluation* got higher rating in 2017 than in 2018 and the difference was statistically different for the two aspects: the usefulness in the course and the usefulness in the future (t-test, $p=0.01$). The reason could be that in 2018, the students felt it was a bit unnecessary to do the summative UX evaluation with users, because they had already done user testing with real users twice earlier in the course. So the last user testing sessions did not give that much additional information. On the contrary, in 2017, this was the first occasion where the students did evaluate with real users.

There was not a statistical difference in the ratings of the other methods and aspects that are compared in table 4.

Based on this comparison it seems that the context of using a particular GDS/UCD method plays a very important role for how useful the methods are rated by students. Moreover, which GDS/UCD methods the students have used previously in the course and in which contexts seems to play an important role in how valuable the students think the methods are. In other words, evaluations are always relational to the students' previous experiences.

4.2 Qualitative Feedback During the Course

The data collected via the Retrospective Hand technique yielded 207 data items from 18 students (16 on the 2nd week). In total there were 52 comments on "What was good", 41 on "What was not so good", 39 on "I want to point out", 34 on "What I will take home" and 41 on "I want more of". This feedback method seemed engaging for the students, as only 4 (2.4%) of the comment points were left empty. Often, one student indicated several different data items in one finger feedback, e.g., a list of items considered good on the course. Each data item was assigned an ID in format 2018-w1 Good 16-1, i.e., starting with the year of the course, week number (w1), the finger where the comment was reported (Good), followed by a running number for the student (16) and a running number in case several data items were mentioned (1).

The items collected were categorised against the extended framework of Steyn et al. [18], see section 3.3 for further explanation of the themes. The frequency of comments in each theme is shown in Table 5.

As can be seen from table 5, more than half of the students' comments were about Course content (29%) and Course structure (27%).

Table 5. The Frequency of data items collected analysed by themes.

| Theme | Week 1 | Week 2 | Total |
|-----------------------|--------|--------|-------|
| Course structure | 25 | 30 | 55 |
| Course content | 31 | 30 | 61 |
| Teaching methods | 14 | 9 | 23 |
| Learning support | 10 | 7 | 17 |
| Learning resources | 4 | 0 | 4 |
| People | 6 | 4 | 10 |
| Soft skills | 3 | 4 | 7 |
| Staff quality | 2 | 9 | 11 |
| Learning environment | 2 | 0 | 2 |
| Experience | 9 | 3 | 12 |
| Course administration | 4 | 1 | 5 |
| Total | 110 | 97 | 207 |

We analysed these two themes further by calculating the percentage of comments in each category that the students were commenting in (on each finger). The results are shown in table 6.

Table 6. Percentage of comments on each finger by week for the most frequently used themes.

| Category/finger | Course structure | | | Course content | | |
|-----------------------|------------------|--------|-------|----------------|--------|-------|
| | Week1 | Week 2 | Total | Week 1 | Week 2 | Total |
| What was good | 11% | 7% | 18% | 13% | 15% | 28% |
| I'd like to point out | 9% | 20% | 29% | 5% | 3% | 8% |
| What was not so good | 9% | 15% | 24% | 10% | 3% | 13% |
| I will take this home | 11% | 0% | 11% | 7% | 18% | 25% |
| I would like more of | 5% | 13% | 18% | 16% | 10% | 26% |
| Total | 45% | 55% | 100% | 51% | 49% | 100% |

Comments related to course structure were most prominent in the negative feedback responses of 'I'd like to point out' (29%) and 'What was not so good' (24%) fingers, which indicates that the course structure was still not optimal from the students' perspective. After the precisely structured first week of the GDS, the students missed a similar style for the second week, which was more focused on lectures and user evaluations, as these comments show: "was missing the structure from the first week" (2018-w2 NotGood 3) and "there was a bit of confusion on how to keep 1st week's pace" (2018-w2 NotGood 7).

The comments on course content were especially prominent in the ‘What was good’ (28%), ‘I would like more of’ (26%), and ‘I will take this home’ (25%) fingers. Typical comments in ‘I will take this home’ were methodological, such as “Setting experience goals and testing according to them” (2018-w2 Take 10) or “Quick decision making. Short tasks - test, implementation!” (2018-w1 Take 9). Course content that students would have liked to hear more of related to user involvement, which GDS did not include: “I would like more of UCD” (2018-w1 More 10).

Teaching methods was the third most frequent category with 11% of all comments in this data set. Most of the comments were positive. The students clearly liked hands-on work in a team, like one student commented: “very good and efficient teamwork” (2018-w1 Good 5-3) and “not much theory :)” (2018-w1 Good 5-5). Another one commented that he/she would take home “the way a project like this should be organized and managed and how to find the right balance in talking and doing” (2018-w1 Take 13). The students also liked lectures that provided information to apply in the following teamwork, such as the UCD methods. “The things that I liked this week were lectures. I think it was nice to involve UX in the design process.” (2018-w2 Good 5). The negative comments were mainly about the difficulty of doing high-fidelity prototyping as a team: “There was one person working no sharing” (2018-w1 NotGood 7-1).

On this intensive, international course, several students commented on the fellow students on the course. Comments in the ‘What was good’ finger report about the excitement to meet people from different countries, with different backgrounds, like: “Variety of participants” (2018-w1 Good 4-4) and “The people - students and teachers were very interesting” (2018-w1 Good 12).

Looking at the shares of ‘What was good’ comments in the different categories we see the excitement of the fast-paced course structure during the GDS on the first week. Staff quality was praised after the second week. The students commented also about the general course experience, such as the location, free scholarship, and the food arrangements.

In ‘What was not so good’ section, many commented on the dramatic change in the course structure after the first week, and 31% of the comments after the 2nd week were about the Course structure. The negative comments on Course content were many but mild: “maybe also what would be alternatives to Google sprint” (2018-w1 NotGood 5-2), “Some of the lectures could be more advanced or go into more detail” (2018-w1 NotGood 10-1), and “I think there is a need in intro to UX research methods and tools” (2018-w2 NotGood 14).

5 Discussion

In this section we discuss the possible reasons behind the main results of student feedback presented in the previous section. We also discuss the challenges of integrating user-centred design into the GDS process.

5.1 Reflection on the Results

In this section, we discuss the main findings according to our observations and compare our experiences in 2018 with our experiences in 2017, which were reported in another paper [11].

The course content with GDS was preferred by the students. They all gave it a high numerical rating and commented that they liked the process and were highly motivated during the intensive week of the GDS. They liked detailed instructions, the timeboxing of activities and that the outcome of one activity was used while conducting the next method. They also felt that they had achieved much during the first week, having both done hi-fi prototypes and evaluated those with five real users just in one week. But they also got a bit tired after following the intense schedule, as one of the students explained: “The schedule was too intense, a day off to work on your own could be nice.” (2018-w1 NotGood 8).

The change from the intense schedule of the first week to a more traditional and less structured UCD course schedule on the second week seemed too dramatic for most students. The nature of Iceland also attracted the international students, so some of them used the time until late Sunday evening to experience Iceland. We believe that communicating the pace and style of the course work right in the beginning of the course will help the students to prepare and stay motivated also outside the design sprint.

In the course description 2018, we promised to teach user-centred design and hands-on interaction design. Some comments from the students showed that during the first week when using the GDS process, students missed the user-centredness. Users were involved only on the last day of the sprint to get feedback on the hi-fi prototypes through user testing with real users. The next week, there were two more rounds of evaluations with users, which the students positively commented on. In 2017 the students did user testing sessions with real users once during the course, on the ninth day. But they did evaluations with fellow students and expert evaluation on previous days of the course. Some students commented that they would have wanted to meet real users earlier in the course. It was clear that the students were eager to meet real users and show the prototypes to them already in the first week, so it was very important for them to have the design process as realistic as possible.

Also during 2018, the students expressed interest in solving real-life problems. Students wanted to study people to find interesting problems before deciding the design focus. On the course the teachers provided six examples of suitable design topics, but only one group did choose a project idea from that pool. The other three groups came up with their own ideas to work on. Especially one of the groups seemed to have lost interest in the idea after evaluating with users on the fifth day, so the motivation to work further on the idea in the second week was not as high as during the first week. Choosing project ideas is a big issue for the students, not only to learn to use particular methods.

A clear limitation of this work is that the two courses were not run identically. They took place in different universities, with different students, and with different course structure. Three lecturers gave similar lectures during 2017 and 2018, but both

years there were two lecturers who did not attend the other edition of the course. The courses in 2017 and 2018 used similar teaching methods with short lectures and hands-on training right after the lectures. Therefore, the comparison results between the two courses are indicative, and we can explain some differences in quantitative results by contextual changes only. However, the main contribution of this work is not based on the quantitative comparison but on the qualitative feedback on how the changed course structure with GDS was seen from the students' perspective.

5.2 Thoughts on Integrating UCD into Google Design sprint

User-centred design requires that designers first envision how people are likely to use a product, and then to validate their user behaviour assumptions by conducting tests in the real world. With its main premise of seeing how customers react before committing to building a real product, GDS certainly shares the latter aspect of evaluation with UCD. However, people seem to be only indirectly involved in informing the envisioned designs.

Conducting user studies such as probes [21] or contextual inquiries [7] to gather rich information about people's needs, dreams, and aspirations can be time consuming. Due to GDS's focus on optimizing the different activities of a design process, people's input only comes through asking experts who know most about a given customer. Perhaps reacting to this lack of user input, it was earlier mentioned that one group felt they needed to create a persona and a scenario before starting with the GDS process. Based on what these students suggested, we feel there is an opportunity to extend people's involvement in the design sprint.

One simple way to involve users more in GDS is to invite one potential customer on the first day of GDS (i.e., Monday afternoon) to the activity called Ask the Experts. On top of having a person who knows most about technology, marketing, and business, the person who knows most about the customer could be further complemented by having one potential customer directly involved in these discussions. Involving a potential customer should, however, be carefully planned.

It may be the first time that some experts (e.g., technology) meet potential customers, and thus there could be a risk to turn the Ask the Experts activity into an ad hoc user study. Instead, the potential customer should act as 'one more expert', and the facilitator(s) should keep this in mind to keep the activity on track and within time. Such a setup would allow keeping GDS's focus on optimizing the design process and bringing the best contributions of everyone on the team, while increasing its user-centredness.

Additionally, the students could be asked to think more about the users while using the GDS process. Shortly after the mapping, students could be asked to describe better the user groups that are listed on the map, but either analysing the main characteristics of the user groups or making a persona for each user group. Furthermore, after making the paper prototypes and before making the storyboard, the students could be asked to set the user experience (UX) goals that they would like to enable with their prototype. This way, students would have the UX goals in mind when making the storyboard.

6 Conclusion

This paper has reported the development of a two-week intensive UCD course for university-level students in an international setting. The first edition of the course ran in the summer of 2017 and the second edition in the summer of 2018. We have presented and interpreted both qualitative and quantitative data collected from the students during the two editions of the course.

This paper contributes to the limited academic literature on teaching UCD methods, and this paper seems to be the first scientific publication discussing use of GDS in an academic interaction design class. There naturally remains much further work to do. Based on our experiences of the two-week intensive course, where the aim was to teach both user-centred design and hands-on interaction design methods, we propose the following improvements.

First, it seems important to spend a day for setting the scene for the teamwork and the design assignment before starting the GDS. Unlike in companies using GDS, the students on this class did not know each other and had varied educational and cultural backgrounds, therefore the need for team building was higher than normal. Second, UCD starts from understanding the users and the context, but this phase was missing in the 2018 class. In 2019, we are planning to go to the field to learn user research in a real context. Third, most students had just arrived to a foreign city without knowing anyone. Jumping to the intensive five-day GDS seemed to mostly go well, but may have influenced the tiredness of the students the second week. In 2019, the GDS could start the first Wednesday and continue after the weekend break. Fourth, students considered user evaluations highly useful for improving the design, so we need to find a way to integrate the evaluations to the design process. We also need to plan how to find external representatives of the target user group for each team.

Finally, this paper provides a methodological contribution for evaluating the students' course experience. We used the so-called 'Retrospective Hand' as a qualitative evaluation method for our course. In our case, it provided a significant amount of rich and relevant feedback on the design and content of the course. Since the results were promising, further validation of this student feedback collection method would be welcome. We utilized the framework by Steyn et al. [18] for the analysis of the Retrospective Hand data. Due to the different type of our course, the types of students' comments were differing from those in Steyn et al [18] and we needed to create new categories to classify all comments. However, further work is required to test the categorization framework in other educational contexts. We expect both the categories and definitions of each category to become more comprehensive through future studies. While we find the open answers highly interesting on a smallish class like this, the emerged categories may help develop a quantitative questionnaire for collecting quick feedback on larger classes.

7 Acknowledgements

We would like to thank Nordplus for granting money to the project and to make it possible to invite students and lecturers to take part in the course. We would like to thank all the students that have participated in the two occasions of the course and given us access to the data which this paper is based on.

References

1. Aggarwal, A. K., Lynn, S.: Using Continuous Improvement to Enhance an Online Course. *Decision Sciences Journal of Innovative Education*, 10(1): 25-48, (2012).
2. Bias, R. G., Marty, P. F., Douglas, I.: Usability/User-centered design in the iSchools: Justifying a teaching philosophy. *Journal of Education for Library and Information Science*, 53(4), 274-289, (2012).
3. Brooke, J., SUS: A quick and dirty usability scale. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 189-194). London, UK: Taylor & Francis, (199,19).
4. Cronbach, L. J.: Course improvement through evaluation. *Teachers College Record*, 64:672-683, (1963).
5. Cronbach, L. J.: Evaluation for course improvement. In R. W. Heath (Ed.), *New Curricula*. New York: Harper & Row, (1963).
6. Ezzy, D., *Qualitative analysis: Practice and innovation*. Psychology Press (2002).
7. Holtzblatt, K., Wendell, J. B., & Wood, S.. *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier, (2004).
8. Jia, Y., Larusdottir, M.K., and Cajander, A., The usage of usability techniques in Scrum projects. In: M. Winckler, P. Forbrig, R. Bernhaupt, eds. *Proceedings of the HCSE 2012*. LNCS, vol. 7623. Heidelberg: Springer, 331–341, (2012).
9. Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J. P., Karvonen, H., ... & Tokkonen, H., Defining user experience goals to guide the design of industrial systems. *Behaviour & Information Technology*, 34(10), 976-991 (2015).
10. Knapp, J., Zeratsky, J., Kowitz, B.: *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*. Simon and Schuster, New York (2016).
11. Larusdottir, M. K., Roto, V., Stage, J., Lucero, A. (2019). *Get Realistic! - UCD Course Design and Evaluation*. In Bogdan, C., Kuusinen, K., Larusdottir, M. K., Palanque, P., Wincler, M. (Eds.): *7th International Working Conference on Human-Centered Software Engineering (HCSE 2018)*, Springer International Publishing.
12. Lucero, A. Using affinity diagrams to evaluate interactive prototypes. In *Human-Computer Interaction*, 231-248. Springer, Cham (2015).
13. Mattelmäki, T., Lucero, A., Lee, J.-J.: *Probing—Two Perspectives to Participation*. In *Collaboration in Creative Design*, pp. 33-51. Springer, Cham, (2016).
14. Nielsen, J., Molich, R. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256). ACM. (1990).
15. Seffah, A., Andreevskaja, A.: Empowering software engineers in human-centered design. *Proceedings of the 25th International conference on Software engineering*, pp. 653-658. IEEE Computer Society, (2003).
16. Sigurgeirsson, D. B., Daniels, M., Larusdottir, M., Jonsson, B. Þ., Hamdaga, M.: Learning Outcome Outcomes: An Evaluation of Quality. In *Proceedings of the 48th Annual Frontiers in Education Conference (FIE)*. IEEE, (2018).
17. Simons, H.: Utilizing Evaluation Evidence to Enhance Professional Practice. *Evaluation*, 10(4): 410-429, (2004).

18. Steyn, C., Davies, C., Sambo, A.: Eliciting student feedback for course development: the application of a qualitative course evaluation tool among business research students. *Assessment & Evaluation in Higher Education* 44(1), 11-24 (2019).
19. Vincent, J., A. Waters and Sinclair, J.: *Software Quality Assurance, Volume I: Practice and Implementation*. Prentice Hall, Englewood Cliffs, New Jersey (1988).
20. Urquiza-Fuentes, J., Maximiliano, P.-V.: Investigating the effect of realistic projects on students' motivation, the case of Human-Computer interaction course. *Computers in Human Behavior* 72 692-700, (2017).