Kjetil Liestøl Nielsen

# Student Response Systems in Science and Engineering Education

Thesis for the degree of Philosophiae Doctor

Trondheim, July 2012

Norwegian University of Science and Technology
Faculty of Natural Science and Technology
Department of Physics

**NTNU – Trondheim**
Norwegian University of
Science and Technology

**NTNU**

Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Natural Science and Technology
Department of Physics



ISBN 978-82-471-3901-1 (printed version)
ISBN 978-82-471-3902-8 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2012:293

Printed by NTNU-trykk

# Abstract

This thesis explores aspects related to the use of a Student Response System (SRS) in preparatory engineering courses at Sør-Trøndelag University College (HiST), Norway. The work described here started as a part of a project in which one of the goals was to develop and test a new SRS for modern mobile devices. An important part of the project was the implementation of this system in lectures during the development stage in order to perform evaluation and conduct research on the use of SRS. Moreover, the goal was to use the results both to improve the design and functionality of the system as well as to increase the knowledge about SRS-use in science education. The main focus of this thesis is not on the actual development of the new SRS, but rather on addressing methodological choices in the main areas, in addition to the actual software, which are important for successful implementation and use of SRS in science education. These include the role of the teacher, the SRS-questions, and the different sequences during SRS-sessions. This thesis also tries to answer some of the unanswered questions that exist in the SRS-research scene and to increase teachers' awareness of how they use SRS in class.

The research focuses both on students' own experiences as well as observations of students engaged in peer discussion. Analyses of video clips of students discussing quiz problems during SRS-use provided insight into what students focus on in their argumentation as well as how different methodological choices can have a significant effect on the discussions. While such observations are important for understanding various aspects of SRS-use, the success of SRS-implementation is highly dependent on students' attitudes towards the system. Students are very aware of their own learning environment and it is important that they see the benefits of using SRS if it is to be positively received. While the use of written surveys provided an efficient way of assessing students' general attitudes in the classes using SRS, analyses of focus group interviews provided deeper insight into their experiences. Although most of the research was conducted in introductory physics classes, the majority of the results in this thesis with regard to implementing and teaching with SRS are applicable to science education in general.

# Preface

This thesis is submitted as part of the requirements for the degree Philosophiae Doctor at the Norwegian University of Science and Technology (NTNU). The work has been performed at Sør-Trøndelag University College (HiST), which also funded the PhD-position. My supervisors have been Jon Andreas Støvneng at NTNU and John Birger Stav at HiST.

My contributions to the papers, and the work described therein, are as follows: I taught one of the preparatory engineering classes in physics using SRS from the autumn semester of 2009 through the spring semester of 2011 and all SRS-quizzes for the autumn semester of 2009 through the spring semester of 2010 were adapted by me. I did not conduct the actual interviews, but the interview topics and questions were developed jointly, mainly together with Gabrielle Hansen-Nygård. The interview analysis in paper B was conducted by Gabrielle Hansen-Nygård while I conducted all of the interview analysis in paper C. Paper E contains interview analysis conducted by Gabrielle Hansen-Nygård as well as by me. However, the relevant aspects from these analyses were summarised by me to form the results of the paper. The written surveys from the different years of testing were also developed together with Gabrielle Hansen-Nygård. The surveys were analysed by me. All video observations were conducted by me with the assistance of my colleagues. All video material was analysed by me, but I had assistance in the initial transcription of the numerous video clips.

As for the development of the Student Response System, I took part in the planning of its design, functionality, and workflow. My contribution to the design planning included, among other elements, producing mock-up videos of how the SRS should function in a teaching environment. All papers, except for paper B, have been written by me.

Kjetil Liestøl Nielsen,
Trondheim, July 2012

# Acknowledgements

There are several people I have to thank for their help, patience and contribution in order for this thesis to become a reality. First of all, I have to thank Gabrielle Hansen-Nygård for fruitful collaboration during research, scientific discussions, feedback on the papers and on this thesis, and for making a stubborn physicist able to see the world through the eyes of a qualitative researcher. In addition, I want to give her a special thanks for keeping me sane during the most stressful parts of gathering the data material presented in this thesis.

After numerous academic discussions, and for enduring my constant nagging about reading my papers, I have to thank Trond Morten Thorseth for his feedback that has forced me to see my work with critical eyes. On the same note, I want to thank Magnus Strøm Mellingsæter for his valuable feedback on my research. He was also a tremendous help when gathering the video material, as was Knut Bjørkli, who also receives my warm thanks. For contributing to the improvement of the language in my papers, I have to give a huge thanks to Elizabeth Ashe-Hegbom.

The research presented in this thesis would not have been realised had it not been for the unconditional cooperation of the teachers of the preparatory engineering courses at HiST. In addition, I have to thank the students in the preparatory courses for their patience and willingness to participate in my research. This thesis would also not have been realised were it not for my supervisors, John Birger Stav and Jon Andreas Støvneng, and I thank them for giving me the opportunity to take on a PhD-degree.

Last, but not least, I have to thank my family for their support during the course of this study. In particular, I want to thank my wife for all her support, love and patience.

# List of papers

**Paper A.** Nielsen, K. L., Stav, J.B., Hansen-Nygård, G., and Thorseth, T.M. (2012) Designing and Developing a Student Response System for Mobile Internet Devices. In: *Learning with Mobile Technologies, Handheld Devices and Smart Phones: Innovative Methods.* Zhongyu Lu. IGI-Global (pp. 56-68).

**Paper B.** Hansen-Nygård, G., Nielsen, K.L., Thorseth, T.M., and Stav, J.B. (2012) Developing and Evaluating Practical Methodological Guidelines for use of Student Response System in Teaching. In: *Learning with Mobile Technologies, Handheld Devices and Smart Phones: Innovative Methods.* Zhongyu Lu. IGI-Global (pp. 90-104).

**Paper C.** Nielsen, K.L., Hansen-Nygård, G., and Stav, J.B. (2012) Investigating Peer Instruction: How the initial voting session affects students' experiences of group discussion. *ISRN Education.* Article ID 290157, 8 pages.

**Paper D.** Nielsen, K.L., Hansen-Nygård, G., and Stav, J.B. (Preprint 2012) How the Initial Thinking Period affects Student Argumentation during Peer Instruction: Students' Experiences Versus Observations. *Submitted to Studies in Higher Education*

**Paper E.** Nielsen, K.L., Hansen-Nygård, G., and Stav, J.B. (Preprint 2012) Teaching with Student Response Systems: Teacher-centric aspects that can negatively affect students' experiences of using SRS. *Submitted to Research in Learning Technology*

**Paper F.** Nielsen, K.L. and Stav, J.B. (Preprint 2012) Student Response Systems in Physics Lectures: Do Voting Results Represent a Correct Image of Student Understanding? *Submitted to Research in Science & Technological Education*

# Contents

# Chapter 1

# Introduction

Taking on an engineering degree can be a daunting task and there are many students that struggle during the course of their education. An evaluation of engineering education in Norway in 2008 found that 45% of engineering students complete their three-year bachelor's degree within the normal time [1]. By using additional time, this number increases to about 60-70%, which means that there is a significant number of students that never finish their degree. Engineering students often struggle with highly theoretical courses such as mathematics and physics, the latter often being regarded as particularly frustrating and demotivating [2]. One aspect of learning science that proves challenging for novice students is that they enter their education with deeply rooted conceptions about the world that are often in stark contrast to scientific views [3]. Unfortunately, the traditional blackboard style lecture format, consisting of a teacher monologue, has many limitations and often does not give students the necessary opportunities to detect problems early in their learning [4].

Students are rarely given time to reflect upon the subject matter or test their knowledge during class [5], and as the size of the classes increases the traditional lecture format also limits the communication between the teacher and students [5]. Not only does the physical distance between the teacher and students increase, but when surrounded by a large number of their peers, students often find it embarrassing asking or answering questions in fear of losing face [6, 7]. This makes it difficult for teachers to assess if students are following the lectures and understand what is being taught. Additionally, students are to a large extent passive participants during lectures, which is not conducive to facilitating cognitive skills such as critical thinking [8, 9], especially since students' ability to concentrate decreases significantly after 20 minutes [10, 11].

There are several ways a teacher can deal with these challenges. A popular strategy, especially in the US and the UK, is to utilise a Student Response System (SRS) during lectures. Such systems go by many names in the literature, such as 'Audience Response System', 'Personal Response System' or 'Interactive Voting Pads'. An SRS can in its simplest form be described as an electronic voting system where the teacher can present students with questions, often in the form of a multiple choice quiz, and students give their answer anonymously with a small voting device, commonly referred to as 'clickers' [12]. Students can be further challenged by engaging in peer discussions, where they try to convince each other of the correct answer and come to a consensus before voting [5, 13]. Students are given a break from the constant flow of information, but where the break has the additional benefit of giving students the opportunity to test whether they have understood the subject matter [14]. Using an SRS in class has been found to have many positive effects, such as students feeling an increase in motivation and engagement [15, 5] and increased student performance [16, 17] including better conceptual understanding [18].

The idea of an electronic voting system is not new. In fact, Student Response Systems have been used in education since the 1960s [19], but the old systems were expensive, did not function well and were difficult to use [19]. Moreover, adaptions of SRSs in class were very technical-centric, regarding the system as a catalyst for student achievement and attitudes [20]. It is therefore perhaps not surprising that the early studies on SRSs did not find any significant improvement to students' achievements [20]. History has shown that implementing new technology into education with a belief that technology automatically improves learning, is bound to result in failure [21]. As Chandler [22] put it: 'as soon as learning processes are not a core consideration and pure technological capabilities, functionality, and the "wow" factor are made central, then chaos inevitably ensues' (p.392).

As the focus on SRS-studies shifted towards pedagogy and methodology, SRSs began to show positive results in terms of increased performance [20]. While traditional lectures focus on a transmission model of learning (learning through demonstration and repetition), researchers have argued that SRS can, in addition, promote learning within multiple learning paradigms, such as the constructivist, social constructivist [5, 23], and the metacognitive learning paradigm [5]. Merely exposing students to new information is not always enough to effectively facilitate learning. Students also need to examine their own ideas and to be challenged to use their prior knowledge to try to understand new experiences [23]. They have to, in a way, *construct* their own knowledge. The students need to be active participants, and activating the students is one of the key motivations behind utilising SRS in class. With the presented quizzes, students have to use their knowledge about the subject matter to solve challenging tasks during the lectures. The social aspects of constructivism come into play through the social interaction in group discussions, scaffolding, and assisting in joint construction of meaning [5]. Through peer discussions students are often required to justify their answers to the group and evaluate justifications from other group members. In doing so, they have to develop strategies which enable them to monitor their own learning and understanding, which are important aspects of the metacognitive learning paradigm [5].

In other words, it is not the technology of SRS, but how SRS is used that gives benefits to classroom lectures. The study by Poulis et al. [24] is a clear example of this where the researchers showed an increased pass rate and reduced variability between the achievements of different students, using an SRS which was installed in the classroom in 1966! Another example is the study by Lasry [25] where the researchers compared a modern SRS to paper flashcards, which were often used in the 1990s whereby students voted on quizzes by holding up a card with their answers. The researchers found no statistical difference on students' performance between the two. These results raise an important question: if the technology is not an important aspect of the benefits of using an SRS, why should any attention be given to it? The answer may be obvious from the paragraph above: the old SRSs were expensive, did not function well, and were difficult to use. With systems in such a state, the threshold for using them is fairly high.

At the time of the research conducted in this thesis, most modern SRS technology was based on either infrared- or radio-transmitters and receivers. Teachers used a form of either stand-alone software or plugin embedded into standard presentation software such as PowerPoint. Although these types of SRSs are a vast improvement on the old hard-wired systems of the 1960s, they are still not without their problems. Despite the wireless technology, teachers wanting to use SRS still need hardware, such as receivers, to be installed in each classroom where the system is going to be used. In some instances, mobile receivers can be brought to the lecture when SRS is going to be used, but this is not an ideal solution. Receivers and clickers can also be very expensive and there are often expenses for technical support as well. Some researchers have even reported on SRS-software being unusable because of being heavily license-managed. Expenses with SRS can be reduced by having students buy their own clicker, but studies have reported on students' satisfaction with SRS decreasing because of the cost of the clickers [26, 11]. Having the system locked to particular software is also not an ideal solution as this decreases the flexibility of the system.

The research scene on SRS is also not without its issues. In recent years, research on SRS has seen a shift from comparing learning outcomes by using SRS versus not using it, towards comparing and investigating various aspects of using such systems, for instance, different SRS-sequences and how methodological choices can affect group dynamics. While this is an important change in order to understand the effects of using SRS in class, there are still many unanswered questions as well as unaddressed aspects of using SRS. Kay and LeSage [19] argue that the research scene on SRS needs more detailed studies on why specific benefits and challenges influence the use of such systems and that studies need to combine both qualitative and quantitative data. Despite the simple functionality of SRSs, implementing and using an SRS successfully in lectures is not a simple task and there are many issues that can affect both possible learning outcomes and students' experiences with the system. The unanswered questions which are addressed in this thesis will be discussed in the next chapter.

## 1.1 Research goals

The work described in this thesis started as a part of an EU co-funded project at HiST, the Edumecca project (www.histproject.no). One goal of this project was to develop a new Student Response System for modern mobile devices based on Wi-Fi. Although Wi-Fi based systems existed at the time of the work involved in this thesis, they were often limited to certain devices and the teacher software still had many of the same limitations as described for earlier systems. The goal was to develop a system which addressed many of the limitations of commercial SRSs, with regard to flexibility, effectiveness, ease of use, and the cost of implementation. Moreover, the goal was to use this software in class during development to provide evaluation of the system from teachers and students in order to improve the functionality and design. Another important aspect was to use the system to conduct research on SRS in order to improve SRS methodology and understanding of using such a system in classroom lectures.

While the developed system is described and discussed in this thesis, the development itself is not the main focus. Rather, the focus is on the research conducted to advance understanding of different aspects of SRS-use. Successful implementation and use of SRS is highly dependent on various factors including:

1. The system
2. The SRS-sequence (i.e. the different phases during an SRS-session)
3. The quizzes
4. The role of the teacher

Instead of focusing on one particular aspect, this thesis aims to address issues and highlight certain aspects on all of the areas listed above. Although having a broad focus introduces a risk of shallowness, the motivation for this choice is to increase the overall usefulness of this thesis for teachers wanting to use SRS in their lectures.

Despite most of the research in this thesis being conducted in introductory physics courses, it was not the intention to restrict the research to only be applicable to physics education. Thus, with a few exceptions, most of the results in this thesis are relevant to science in general. Additionally, it should be noted that it was not the intent to conduct research with a focus on how students learn, or investigate why SRS can facilitate learning in a context of different learning paradigms.

Rather, the thesis has a more practical approach and focuses on some of the common SRS-strategies in lectures while addressing methodological issues within such use. The thesis tackles some of the unanswered questions that exist within the SRS-research scene and how methodological choices can affect key features such as the group discussions and students' experiences of using the system. In addition, the thesis focuses on how the different aspects are related and affect each other. Some aspects might seem unimportant or trivial, but as will be shown in this thesis, can still have a significant impact on both the effectiveness of SRS and how it is perceived by the students. In this way, this thesis aims to increase teachers' awareness of how they use SRS in order to facilitate a more conscious and constructive use of such systems during lectures.

### 1.1.1 Research questions

As a result of having a broad focus, there were several different research questions that were asked. Still, there are also a few overall research questions that are addressed throughout the thesis:

- How can methodological choices affect students' experiences of using SRS?
- How can methodological choices affect the different phases of the SRS-sequence?

Some of the papers included in this thesis address these questions in a more general sense, such as paper B, which evaluates students' experiences of the methodological procedures when SRS was initially implemented. Another example is paper E, which focuses on the negative effects of methodological choices with the research questions: *What teacher-centric aspects can negatively affect students' experiences of using SRS* and *why do these teacher-centric aspects negatively affect the students?* Also in paper A, which describes and discusses the developed SRS used in this thesis and which is not a traditional research paper, the possible effects of methodological/design choices (with regard to the actual software) are an important topic.

Other research questions had a more narrow focus and addressed specific features of using SRS in lectures, such as the group discussions. An example of this is in paper C which addresses the effects of having an initial voting session before quiz discussions: *How does the initial thinking period and voting session affect students' experiences of the group discussions?* In a study by Nicol and Boyle [27], the researchers gave some insight into this particular question (more on this in the next chapter), but the researchers emphasise in their paper that the effects of the initial thinking period need further research. In their study, the initial thinking period was not the only focus as they compared two SRS-methods which had other differences apart from the initial voting session. Thus, there might be details of the effects of the initial voting session that were left unnoticed. An in-depth analysis of student interviews where the students could compare two methods with the only difference being the initial voting session, could give more insight into its effects. More insight could also be provided if student interviews are accompanied by video observations of students engaged in peer discussion. This leads to another important research question which was addressed in paper D: *Do students' own experiences of the initial thinking period correspond to what really happened?*

It should be noted that the different research questions addressed in this thesis were not all formulated from the start of these studies. Rather, several of the questions have arisen during both the testing periods and during the analytical processes. One example of this is the effects of students seeing the results of the initial voting session: *How can seeing the initial voting results affect students' experiences of the group discussions?* This question emerged from student comments during some of the initial interviews when investigating the effects of the initial voting session, and resulted in weaving this topic into the later interviews (paper C). Before the students addressed this in the interviews, we did not realise the possibility of seeing the voting results having a dramatic effect on the group discussions.

Another example of unforeseen research questions was during the video analysis where two similar quizzes were found to have very different voting results. This raised the question of why the students voted correctly on one quiz, but were more divided in the other. The following research question was then formulated: *What are the underlying causes for students voting differently on similar quizzes?* On the surface, it was obvious that one question was more difficult than the other.

However, the intent was to investigate the arguments presented by the students to see if this could shed light on, for instance, how the different quizzes were formulated or on less obvious aspects of the difficulty of the questions. The results were, as will be described later, surprising and demonstrated some of the challenges with using SRS as an assessment tool for student understanding.

## 1.2 Structure of the thesis

This thesis consists of five chapters in addition to the included papers, the first being this ***Introduction***. The second chapter, ***SRS in Education***, describes SRS-use in more detail and introduces the reader to common methodological choices when using SRS. The chapter also presents recent research on SRS and discusses unanswered questions as well as areas where research is lacking. The third chapter, ***Methods***, describes the different kinds of data and analytical methods used and the motivation behind these. The fourth chapter, ***Results***, provides a summary of each paper included in this thesis. The first two papers function as an introduction to the research, describing and discussing the developed Student Response System as well as describing the first implementation in more detail and the initial feedback from the students. The last four papers consist of more specific research on various aspects of SRS. In addition to functioning as an introduction, the first two papers also provide a better context for the ***Discussion and Conclusions***, the fifth chapter. Here the relations between the different results are discussed as well as their implications for teaching with SRS and future SRS-design. At the end of this thesis before the included papers, there is a brief chapter with suggestions for ***Future Research*** on the use of SRS.

# Chapter 2

# SRS in Education

The aim of this chapter is to describe the use of Student Response Systems in more detail, such as different methodology and SRS as a tool for the teacher to activate the students and provide feedback. In addition, previous research on SRS will also be discussed. This chapter does not aim at being a complete guide to SRS and the research scene, but rather to focus on presenting important challenges when using SRS, and describing findings in the research scene that inspired the work described in this thesis.

## 2.1 Peer Instruction

The basic functionality of a Student Response System can almost be comically simple: the students can press a button on a small device and the teacher can see how many have pressed specific buttons. Some systems have different variations on this functionality [28], but the principle of a simple voting system stays mostly the same. The simplicity of an SRS is also its strength as it presents a flexible tool for the teacher with many possibilities on how to use it. However, herein lies also one of the main challenges with SRS: how should one use it? Teachers can, for instance, ask students if they can follow the lecture or ask for students' opinions on a specific matter. A common way of using SRS is to present students with a multiple choice quiz. However, there are several methodological questions that teachers need to answer when using SRS in such a manner: what type of questions should be asked, should students think alone or discuss with their peers, and what happens after the voting has ended?

Perhaps the most popular SRS-sequence or method is 'Peer Instruction', developed by Eric Mazur in the 1990s as a means of focusing students' attention on underlying concepts in physics without sacrificing students' ability to solve problems [13]. The method was actually developed without SRS in mind, as students would use the flashcards mentioned earlier, but SRS removes the practical challenge of having to manually count student votes. An important aspect of Peer Instruction is 'ConcepTests', conceptual multiple choice questions that revolve around a particular physical concept. The questions are designed so that students have to use their knowledge about physics instead of relying on memorisation. After students are introduced to the problem, they first think individually for about a minute without talking with their peers. Students give their answer before engaging in group discussion, where they try to convince their peers of their choice. The session ends with a revote and the teacher explaining the solution.

## 2.2 Group discussions

Engaging students in group discussions can be an important factor for increasing students' understanding of the subject matter [13]. Not only do they have to figure out an answer for themselves, but they also have to try to generate plausible arguments to convince their fellow group members. In order for students to develop their scientific language and critical thinking skills, it is important that students are challenged to put their thoughts into words [29]. It can be difficult, however, to motivate students to engage in group discussions with the traditional lecture format. Students often do not see the point in discussing as there is no obvious payoff for their engagement [14]. The voting session with SRS provides this payoff and gives the discussions a meaning because it functions as a clear conclusion for the discussions. Additionally, students can state their opinions in the voting session and provide feedback to the teacher as well as see how they fared compared to their peers [14]. The voting session can also create competitive feelings among students [30], wanting their group to 'win' or at least be the correct party when disagreeing during discussion. As a consequence, students have been reported to be more alert during lectures with SRS [31]. Students also value the group discussions because they can hear how other students think, and students have emphasised that it is easier to understand explanations from their peers as they often 'speak the same language' [27].

Thus, the group discussions can facilitate learning both with regard to students trying to convince other group members as well as for those being explained to by their peers. In a study by Smith et al. [32] the researchers found that group discussions during SRS indeed do facilitate learning. On the other hand, there are studies that show that there are several factors that can influence the dynamics of group discussions, such as grading SRS-answers. In many universities, it is common for students to receive points towards their grades for SRS-participation [10], which has been shown to be an effective strategy of increasing attendance during lectures [33]. A step further from giving points for participation, is to give extra points for correct quiz answers. Even though the motivation for doing so is to encourage students to participate during discussions, a study by James [34] showed that the effect can in fact be the opposite. One could argue that students are not punished for incorrect answers, but rather rewarded for answering correctly. Still, the absence of the additional points towards their final grades, can be seen as a punishment. The study found that students relied heavily on the opinions of more skilled students when the correct answer gave more points, while they were more inclined to explore different solutions and ideas when only participation was rewarded.

The notion of more skilled students influencing group discussions is also found in the study by Nicol and Boyle [27], where the researchers compared Peer Instruction to a sequence initially described in Dufrense et al. [23]. In this 'class-wide' sequence students engage in discussion directly after the quiz is presented without having the initial thinking period. Also, instead of ending the SRS-session with a teacher explanation, students (and teacher) engage in a class-wide discussion. The study found that the students preferred having this initial thinking period because they use it to gather their thoughts and think about what to say in the discussions. Students also emphasised that they were more likely to participate during discussions if they had the time to think about the solution. Further, students felt they would be more likely to be dominated by 'stronger' (i.e. more skilled) students if they did not have the time to think about the questions before discussion. The researchers argued that without the initial thinking period, students are not given the opportunity to generate their own mental model of the quiz, resulting in less 'conceptual conflict' at the start of group discussions. As a consequence, students are more likely to accept dominant explanations [27].

Nevertheless, there have not been any studies that have investigated students' claims of more fruitful discussions when this initial thinking period is included. Although students' own experiences and attitudes can be vital for succeeding in implementing SRS [35], they do not necessarily correspond to measurements. An example of this is the early studies on SRS where students were very positive towards the system and felt that it helped them learn the subject matter even though no significant increase in performance was found [20]. Thus, it is important for research on SRS to include students' own experiences and measurements/observations, both as stand-alone studies as well as comparisons between the two. As the inclusion of the initial thinking period increases the time used on each SRS-session (as well as increasing the pedagogical requirements of the teacher) one has to ask the question: is the initial thinking period and voting session really needed?

Recent studies on Peer Instruction have also found other aspects that can influence group discussions. When the initial voting session is used it is common to display the voting results, usually in the form of histograms, while students are discussing [10]. Perez et al. [36] published a study where the researchers found that if students are shown the initial voting results, they are 30% more likely to choose the alternative with the majority of votes. Although the results did not explain *why* this happens, the researchers provided some interpretations for their results. One explanation was that the alternative in the majority provided a talking point or stimulus to which students try to find out why the answer was so common. Another interpretation was that seeing the results would make students realise their own mistakes and find flaws in their own reasoning. Note that this interpretation assumes that the alternative in the majority is the correct one, which may not always be the case. Another quantitative study by Brooks and Koretsky [37] also showed that students' confidence in their own argument increased if they could see that their answer was in the majority. Still, there are no studies as of the writing of this thesis, to my knowledge, that have investigated (qualitatively) the effects of seeing the initial voting results on group discussions.

## 2.3 SRS as a tool for feedback

Student Response Systems have long been regarded as a valuable tool for improving communication between the teacher and students [12] by providing feedback through the voting results [38]. If students vote incorrectly, SRS can give the teacher an indication of gaps in students' knowledge and reveal misunderstandings [39, 40]. The students can also see how they performed compared to the rest of the class and are given an indication of areas where they need to study harder [41]. By realising gaps in students' knowledge, the teacher can tailor the lecture accordingly by, for instance, providing an extra example on the blackboard [13]. On the other hand, if the vast majority of students vote correctly in quizzes, it indicates that students understand the subject matter and the lectures can progress on to other topics [42].

However, as has been described in this chapter, there are several factors that can influence students' decision-making during discussions, which raises an interesting question: how reliable are the voting results as an assessment of students' understanding of the subject matter? It is likely that a wrong answer suggests gaps in knowledge or the presence of misunderstandings, but does a correct answer translate to correct knowledge? Students can choose the same alternative as more skilled students without understanding the solution, or just follow the majority of the class. Both examples challenge the interpretation of voting results as a correct representation of students' understanding. The reliability of SRS as an assessment tool is likely dependent on several factors, such as the nature of the questions presented to the students. For instance, if the questions are too easy, students might think that they understand the subject matter better than they really do [43]. The reliability of voting results as a measurement of students' understanding is an aspect that is rarely addressed in SRS-research.

## 2.4 The role of the teacher

These challenges with SRS emphasise the importance of the teacher and methodological choices with the system. The success of SRS is highly dependent on the pedagogical skills of the teacher [41], and while SRS can act as a tool for improving the lectures, it can also be detrimental to them if used poorly [28]. Using too much time setting up the system [5], not managing discussion time correctly [23], and teachers having a negative attitude towards SRS [44], are examples of aspects that can lower student satisfaction with using such systems in their lectures.

While these examples describe some of the more unintentional aspects of teaching with SRS, deliberate implementation choices can also have a significant impact on students' attitudes. Using SRS, for example, to only record attendance is looked down upon by students [45] as the system is not used as a pedagogical tool to facilitate learning and/or engagement. Although grading of SRS-participation has been shown to increase attendance, it has also been shown to have a negative effect on students' attitudes [46, 47]. The increase in attendance becomes somewhat artificial because it becomes mandatory to be in the lectures as students risk losing points on their grade [26].

There are several best practice guides that give tips to teachers wanting to use SRS (such as the paper by Caldwell [10]) and several publications also mention aspects of SRS-use that were found to cause irritation or dissatisfaction among the students, such as those mentioned above. However, these issues are often sparsely described and, understandably, not the main focus of the publications. There is a need to focus on the challenges teachers face when using an SRS. In order to appreciate what to do, it is important to know what not to do, and more importantly, understand *why* and *how* such aspects have a negative impact. In order to gain a deeper understanding of aspects that can negatively affect students' satisfaction with SRS, it is not enough to simply generate a list of what not to do. Instead, there is a need for deeper analysis of students' own experiences and how the different aspects of SRS-use relate to each other.

# Chapter 3

# Methods

The aim of this chapter is to describe the type of data and analytical methods used to examine the research questions. The results are based on both qualitative and quantitative data analyses and describe the use of SRS from the perspective of the students as well as from observations. Students' experiences of using SRS in their lectures were examined using interviews and surveys while video/audio clips of students engaged in group discussions represent the observations. Most emphasis is placed on the interviews and video analyses, with written surveys in the form of questionnaires being used mostly as complementation. The specific analytical steps are described in the relevant papers. Therefore, this chapter will not focus specifically on all of the analytical processes, but rather on the motivation behind them and on a more general description of the data type and analytical methods. Before the presentation of the research methods, however, the chapter will provide a brief description of the classes and the extent of SRS-use during the course of the research in this thesis.

## 3.1 Background

This thesis contains results from data collected in the course of three years of SRS-testing in preparatory engineering courses at HiST, from 2009 to 2012. Preparatory courses consist of students, usually about 50-70 per class, who wish to undertake an engineering degree, but who do not have the required courses from senior high school. The students often come from a vocational background with an age average in the early to mid-20s, many with several years of previous work experience. The courses last for a year and provide the necessary curriculum, corresponding to about the two last years of senior high school. The lecture format is usually divided between the traditionally teacher-centric blackboard lecturing (although with digital smart-boards) and problem solving. In the 2009-2012 period, there were four classes per year. The first testing of the SRS in the preparatory courses was in a single class in physics for the duration of four weeks in the autumn semester of 2009. This was followed by the SRS being used for eight weeks in all four classes in the spring semester of 2010, still only in physics. During the 2010-2011 semesters, the SRS was utilised in two classes for the whole year in three courses: physics, mathematics, and social science. In the last year, 2011-2012, the SRS was made available to all teachers in all four classes. Most of the research presented in this thesis was conducted during the first two years of testing.

## 3.2 Research software

An important aspect of qualitative research is the notion of coding in order to extract meanings from, and describe, the data. As the data-set becomes larger, organisation of codes becomes more challenging, which underpins the need for a dedicated software. All coding, both of the interviews (which were analysed by me) and video material, was performed using 'TAMS Analyzer', an open source qualitative research tool developed by Matthew Weinstein (tamsys.sourceforge.net). A small exception is the addition of time codes to student speech time in the video clips, which was done in 'Inqscribe'. In addition, SPSS (version 19) was used for all statistical analyses, which were relevant both for the video and survey analyses.

## 3.3 Capturing student experiences

### 3.3.1 Focus group interviews

Focus group interviews are a popular and reliable method in qualitative research for examining aspects such as experiences, opinions and views [48]. The word 'interview' can be slightly misleading in this context as it is usually associated with a formal setting with an interviewer asking a series of questions to which informants give their answer. In focus groups the interviewer acts more as a moderator, trying to encourage discussion among the informants regarding specific themes. The tone is often informal and focus groups therefore become more 'naturalistic', i.e., closer resembling the nature of actual conversations [49]. Consequently, focus group interviews can be very dynamic with synergy effects where one informant builds on aspects brought forth by another, for instance, by stating points the other would not have thought of him/herself [50].

The interviews described in this thesis were conducted in a semi-structured manner, i.e., using an interview guide based on a list of themes and general questions that the researchers wish to examine. Semi-structured interviews provide a good balance between standardisation and flexibility [48, 49] as they maintain the openness encouraged in focus group interviews, but also maintain structure. In order to find participants for the interviews, students could sign up as volunteers, from which four students, both male and female, were chosen from each class using SRS. In total, there were eight focus groups: two from the spring semester of 2009, four from the autumn semester of 2010, and two from the spring semester of 2011. Some groups were interviewed twice (see paper C for more details) resulting in a total of 13 interviews. All interviews were conducted by Gabrielle Hansen-Nygård, a PhD-student (at the time of writing this thesis) with a background in social psychology and qualitative research methods. Moderating focus groups can be very challenging and best results are seen when conducted by a trained interviewer [49].

### 3.3.2 Grounded Theory

Focus groups owe much of their popularity within qualitative research to their flexibility, not only as an interview form, but also because focus group interviews are not tied to a specific theoretical framework [49]. Grounded Theory is regarded as a rigorous analytical method for extracting aspects from qualitative data such as experiences, feelings and attitudes [51], and was therefore chosen as the main tool for analysing the interviews. Grounded Theory as a method was originally developed in the 1960s by Barney Glaser and Anselm Strauss as a response to the monopoly of positivistic research design in social science [51]. Glaser and Strauss criticised researchers for disregarding human problems that did not fit positivistic research design, as well as for interpreting empirical data solely based on already established theories. Instead, they thought that the social research scene should strive to develop procedures which enabled theories to be generated from empirical data [48]. At its core, Grounded Theory consists of a loop where data is collected and analysed, and where the analysis further motivates additional data collection or 'theoretical sampling'. In the end, the researcher reaches a 'theoretical saturation', where ideally no additional information can be extracted. The result is a set of empirical categories describing the areas of interest or simply, as the name of the method implies, a 'grounded theory'.

It should be noted that the research described in this thesis has not been conducted through a true Grounded Theory methodology, i.e., with theoretical sampling to achieve theoretical saturation, as this would have been extremely time-consuming. Besides, it would have proven to be a severe practical challenge as the interviewed students might not have been easily available after they had finished their preparatory course. However, Grounded Theory includes very concrete and systematic procedures for analysing data material and these procedures are suitable as a self-sufficient analytical method [48]. The analytical tools include a three-step coding scheme (line-by-line coding, focused coding, and categorisation) which in this thesis was adapted from Charmaz [52, 51]. The analytical process starts by systematically coding each line or sentence in the interview transcripts based on their content. Through the three-step procedure (see paper C for more details), the analysis ends up with a small set of categories describing students' most significant experiences. However, only paper C in this thesis describes categories in the traditional Grounded Theory sense, which has strict rules of what constitutes a category. All papers, nonetheless, which include interview analyses have used the coding scheme as described above even though the end product may not necessarily have fulfilled the requirements to be defined as categories.

### 3.3.3 Surveys

Focus interviews, accompanied with a robust analytical procedure such as that of Grounded Theory, provide an effective method for examining experiences, feelings, and opinions, and are also very dynamic as unforeseen aspects can arise both during interviews and analysis. However, this method also has a weakness of being very time-consuming, which restricts the number of informants and focus groups which one can include in such studies. This can be seen as a limitation for the method's ability to form generalisable findings [49]. To address this possible limitation, written surveys (questionnaires) were conducted in all classes using SRS. Surveys are one of the most widely used methods of gathering quantitative data in social research [48]. It is a time-efficient way of providing information from a large sample of a population [53], which makes it suitable for statistical analyses [54].

Notwithstanding, surveys do also have their share of weaknesses. Questions on the written surveys can be open or with predetermined answers. Although open questions can be more flexible, they significantly increase the time needed to analyse the answers [48]. Predetermined answers, while being very time-efficient to analyse, lack flexibility as the answers have to be made beforehand by the researchers. This limits their ability to deeper explore experiences and attitudes [53]. Additionally, written survey questions may be misinterpreted and bias or inaccuracies in the responses may occur. Respondents often have difficulty assessing their own experiences and accurately fitting them into, for instance, a Likert scale [53], and some researchers have even argued that surveys do not necessarily give a better impression of 'reality' than qualitative methods [48]. Thus, surveys were mostly used as complementary to the interviews, and survey questions were to a large extent formed based on information that emerged from the interviews. Surveys were conducted during all three years of testing.

## 3.4 Observations

Observations of students engaged in group discussions were an important part of the research in this thesis, both in order to reveal factors that students might not have realised themselves as well as comparing observations to students' own experiences. During the first two years of testing at HiST, there were often observers present during SRS-use, consisting for the most part either of myself or Gabrielle Hansen-Nygård. This had several functions, such as acting as technical support, seeing how students and teachers behaved when using SRS and observing how they responded to SRS-quizzes, histograms and so on. Observations could reveal possible issues with SRS-use and would often affect later survey-questions or interview topics. However, personal observations have obvious limitations as a research method which is why students engaged in SRS-discussions were also captured on audio and video.

### 3.4.1 Video and audio

The use of video has become more common in social research in recent years as it 'offers means of close documentation and observations and presents unprecedented analytical, collaborative, and archival possibilities' [55] (p.5). Although video has the ability to capture fine details such as body language, appearance, and facial expressions, which are regarded as one of the main strengths of using video in research [56], this thesis focuses on the utterances by the students during discussions.

The main purpose of using video was to simplify transcription of audio from small recorders placed at different student groups. The audio would often suffer from a large amount of ambient noise (i.e., other students discussing) which would have made transcription very difficult and unreliable without the ability to see students' faces while speaking.

Students engaged in SRS-discussions were captured on video during the spring semester of 2010 in four preparatory classes in physics over a total of four weeks. The motivation for using four weeks was to have students become accustomed to having video cameras in their classroom to minimise cameras' influence on student behaviour during discussion as well as to minimise the effects of random or daily based factors that could have had an affect on the students. Prior to filming, students signed permission waivers which also explained that the video was being used for research purposes. We gave a small presentation explaining our wishes to film during class lectures, but emphasised strongly to the students that signing the permission waivers was voluntary. Only two students did not sign. A blind spot in the camera angles was created so that these students would not appear in the video. In the lectures being filmed, students were encouraged to form groups of three, but were otherwise not influenced to how they formed their groups in order to keep the discussions as close to real life situations as possible.

### 3.4.2 Analysing the video

This thesis contains both qualitative and quantitative analyses of the video material. Group discussions of six different quizzes in all of the four classes were analysed and the results were compared to those of the interview analysis (see paper D for more details). Two of these quizzes were subjected to further analysis which goes deeper into examining the actual conceptual content of the arguments presented by the students (see paper F for more details). The quizzes were subjected to a coding process adapted from the discourse analysis method developed by Kaartinen and Kumpulainen [57]. The method was developed in order to investigate the mechanism of explanation-building in small-group discourse. Their coding scheme is based on placing each utterance by the students into predetermined categories by focusing on four parallel analytical frames: discourse moves, logical processes, nature of explaining and cognitive strategies. James [34] used a simplified set of these categories to study students discussing during SRS-questions. The categories used in this thesis were an adaptation of the categories from Kaartinen and Kumpulainen [57] and James [34], resulting in three-level categories and subcategories describing the arguments presented by students. The categories included: 1) discourse action, 2) the extent of the argument being new, or being a rephrasing or extension of previously presented ideas, and 3) the language used in the argument. In addition, there were categories describing various utterances which did not qualify as arguments, such as asking for opinions, stating uncertainty and so on. For more information, see papers D and F.

Besides being placed in categories, each utterance also included time-codes to indicate how long each student was speaking. This was used to test whether there was any statistically significant difference in argumentation time between the two different SRS-methods (paper D). In paper F, the categories described above were used as a starting point for categorising the content of students' argumentation. For more information on the analytical steps, see paper F.

# Chapter 4

# Results

The aim of this chapter is to summarise the work and research that is presented in the included papers. The first paper, A, introduces and discusses the Student Response System used in the classes, while the second paper, B, describes how it was implemented into the lectures. Paper B also evaluates methodological guidelines and presents the initial feedback from the students. These two papers provide an overview of the SRS-implementation and function as an introduction to the rest of the research in this thesis. The last four papers contain more specific and in-depth research on different aspects of using SRS, such as how the initial voting session and thinking period with Peer Instruction affect student argumentation in groups, both from students' own experiences (paper C) and from video observations (Paper D). Paper E focuses on the teacher and describes aspects of how the teacher uses or implements SRS that can negatively affect students' attitudes and experiences of the system. The last paper, F, focuses on the arguments presented by students during quiz discussions, and addresses challenges with using SRS as an assessment tool for student understanding.

## Paper A

**Designing and Developing a Student Response System for Mobile Internet Devices**

This first paper presents the developed Student Response System used in the preparatory engineering classes. As well as presenting the actual software, it also describes and discusses the design philosophy for developing an effective SRS for teaching. Although not a traditional research paper, but rather a design paper, it nonetheless addresses an important aspect of SRS, the actual software, which is not addressed in any of the other papers. Additionally, the paper provides a context for the rest of the research as well as the discussion and conclusions, in particular the implications for future SRS-design.

The design philosophy is centred around three key features: *speed*, *ease of use*, and *flexibility*. The teacher should be able to make questions 'on-the-fly' and start votes within seconds (speed). The software should be intuitive, require minimum of system preparation and also have a layout that makes it easy to use from a computer or digital blackboard (ease of use). The teacher should be able to present the questions in the format he/she sees fit as well as not be restricted by a particular lecture format or method (flexibility).

The paper discusses how the presented Student Response System fits into the design philosophy as well as how various design choices can affect the lectures. Several commercial SRSs, for instance, show the voting results on the voting devices. Although this might seem like a logical and beneficial feature it may provide undesirable effects, such as student attention being on the voting device rather than on the teacher during the explanation of the solution.

Another important feature of the SRS, which relates to the flexibility philosophy, is that students should be able to use their own mobile device. The SRS was made with modern mobile devices in mind, i.e., devices with high resolution screen with, preferably, a touch interface. Although students borrowed Apple iPods as voting devices during the testing period, the prediction was that most students would have a capable device within a few years. The paper presents a survey result from 2009 where 68% of a total of 57 students answered that they had a device suitable for SRS. The question was asked again in a survey at the end of the 2011 autumn semester, where the number of students with a suitable device had increased to 81% (of a total of 160 students) within two years.

# Paper B

**Developing and Evaluating Practical Methodological Guidelines for Use of Student Response System in Teaching**

This paper describes the initial implementation of the Student Response System in the preparatory engineering courses at HiST. Methodological choices for implementing SRS into lectures are described and the paper presents students' experiences with using the system in their class. There are six methodological procedures described: 1) *Introducing SRS to the students*, 2) *Start up*, 3) *When and how to present the questions*, 4) *Small group discussions*, 5) *Polling with a timer and a clock*, and 6) *Teacher's explanation after the quiz*. Two focus group interviews, consisting of four students per group, were conducted to examine students' experience with the system and the different procedures.

Previous research on SRS has emphasised the importance of students 'buying into' why SRS should be used in class [35], and this was also apparent from the student interviews. The introduction gave students a clear view of how to use the system and, more importantly, *why* SRS was used. This piqued their interest in the system and probably increased the acceptance of introducing a new element into their lectures. The interview analysis also showed students' awareness of their own learning environment and that implementation choices that might seem trivial, can have unforeseen consequences. For instance, in each lecture where SRS was used, students could pick up iPods from a suitcase placed at the front of the classroom. The position of the suitcase might seem unimportant, but it resulted in students who came late to the class not picking up an iPod because it was embarrassing going up in front of the class.

Some of the methodological choices were based on recommendations from previous research on SRS while others were based on intuition. For instance, Beatty [12] recommended that SRS-questions should not be read out loud by the teacher as this can render the SRS-session too teacher-centric. This recommendation was not followed, a choice that was favourably valued by the students. Reading the questions out loud clarified the problem and decreased the probability of students misinterpreting the question. It was also considered time-efficient by the students as they felt it would take more time to read the questions themselves. Also, some students suffered from reading and writing difficulties which resulted in their finding it very difficult to participate if they had to read the question on the blackboard themselves.

The students also showed awareness about the motivation behind using SRS in class. There are several reasons for using SRS and students valued it as a means of increasing their learning by being presented with challenging questions. However, the students emphasised that if SRS were to be used for *assessing* learning, there was room for methodological improvements. Lectures with SRS usually consisted of new theory being presented, followed by a quiz after approximately 20 minutes into the new topic. Some students felt that the quizzes came too early, i.e., that the students did not have enough time to digest this new theory and thus not have the prerequisites for answering the quiz. The quiz would indeed present a challenge to the students, but it would not assess if learning had taken place. The students also emphasised their wish to be presented with a quiz at the end of each lecture after they had been given time to solve problems in the textbook. This was to give them a way to verify that they had understood the subject matter before the end of the lecture.

## Paper C

**Investigating Peer Instruction: How the Initial Voting Session Affects Students' Experiences of Group Discussion**

This paper focuses on two different SRS-methods/sequences, Peer Instruction, and a sequence we named 'Classic' which is similar to the Peer Instruction method with the omission of the initial thinking period and voting session. The aim was to investigate how this voting session affects students' experiences of using SRS, in particular, the group discussions. The experiment was conducted during the eight-week testing period of the spring semester of 2010 in all four classes. For the first four weeks, two classes used the Peer Instruction method while two used the Classic method. After four weeks, all classes switched methods. The data consists of seven focus group interviews. Four student groups, one from each class, which consisted of four students per group, were interviewed twice, once after Classic and once after Peer Instruction (with the exception of one group which was only interviewed once).

The analysis resulted in three categories which reflected students' most significant experiences of SRS and the two different methods: 1) *Argumentation and explanation*, 2) *Peer Instruction, Opportunity for individual thinking*, and 3) *Seeing the results: Authority of the majority.* The aspect of argumentation is one of the most important differences between the two methods. Similar to the results in the study by Nicol and Boyle [27], the students valued the initial thinking period as a means to reflect on the questions and generate an explanation which they could use in the following discussions. If they did not have this thinking period, they had to use the discussion time for generating arguments and explanations, but this could prove very difficult as the students would often find it hard to concentrate when other students started speaking. By hearing other students it would 'block' their own thought process. This would often result in skilled students heavily influencing the discussions and arguments presented, and increase the probability of the other group members accepting explanations even without fully understanding them.

Students strongly emphasised the desire to have an opinion they could call their own during quizzes, i.e., to be able to generate their own answer without the influence of others. This would prove to be very difficult during the Classic method as they would be 'coloured' by other students when these started speaking before explanations had been generated. With an explanation ready, students would also be more inclined to participate during quiz discussions and everyone in the group would be more likely to contribute since they could provide more thought-out and convincing arguments.

Despite being heavily in favour of the Peer Instruction method, the students reported on weaknesses in the way the method was used. As is usual when using Peer Instruction and SRS, students could see the voting results from the initial voting session while discussing. Students emphasised that seeing these results could easily influence their decision making and even be detrimental to the discussions. If there were an alternative which had a vast majority of the votes, the discussions could be very focused on this alternative. This was also presented in the quantitative study by Perez et al. [36] as a possible explanation for why the students were more inclined to choose the alternative in the majority. The focus was, however, not to explore *if* this specific alternative was correct, as predicted by Perez et al., but rather find out *why* it was correct, often forcing an explanation upon the alternative irrespective of its correctness. Students felt that they would often lose focus during the discussions when the initial voting results were shown. These aspects, however, did not overshadow the benefits of using the Peer Instruction method.

# Paper D

**How the Initial Thinking Period Affects Student Argumentation during Peer Instruction: Students' Experiences versus Observations**

This paper, as paper C, addresses the effect of the initial thinking period on group discussions, but now with a focus on video observations and how they compare to students' own experiences. Students were filmed two weeks before and two weeks after switching SRS-methods. The video analysis showed that the time spent on arguments and presenting ideas during discussions increased by 91% (based on medians) from Classic to Peer Instruction. In addition to logging the time students spent on argumentation, the distribution of argumentation time across different group members, defined as 'bias', was calculated. A bias of 0 would translate to an even share of argumentation time in a particular discussion, while a bias of

100 would translate to all arguments being presented by a single student. The analysis showed a decrease in bias from Classic to Peer Instruction which makes for the results from the video analysis seeming to correspond exactly to students' own experiences. On closer examination of the bias, however, there were some discrepancies. Each group member was separated by how much they dominated group discussions, i.e., the one with the most argumentation time, D1, the second most, D2, and the least argumentation time, D3. Although there was significantly less bias during group discussions, the results suggest that this was mostly due to D2 students becoming more active rather than D3, who showed no statistically significant difference between the two SRS-methods.

These results might be a consequence of how the students were positioned during group discussions. All students in a group sat in a straight line, facing the front of the classroom. D1 students were found to be more inclined to sit in the middle position, while D3 students were found to be more likely to sit in one of the outer positions. When the middle positioned student started discussing with one of the outer group members, it is possible that this created a communication barrier towards the opposite positioned student, which could have made it more difficult to effectively communicate with the rest of the group. This does not, however, answer the most interesting question: if there is no improvement in argumentation for all group members, why do the students still feel that this is the case? A possible answer could be that there might have been the presence of an 'illusion of group productivity'-effect similar to that which often occurs in brainstorming groups. Research has found that group members during brainstorming sessions often fail to correctly estimate their own contribution, which can result in an overestimation of their own performance. A similar effect could have been present where the students overestimate their own contribution because they have thought more on the answer and thus can more confidently evaluate presented arguments. By more confidently agreeing or disagreeing on an argument, they might feel that they have contributed more even though they might not have actually contributed with more ideas or explanations themselves. The same effect could also be applied to D1 and D2 students who see an improved discussion between themselves, and therefore feel that the group as a whole (i.e., all group members) have more productive discussions.

## Paper E

**Teaching with Student Response Systems: Teacher-Centric Aspects that Can Negatively Affect Students' Experiences of Such Systems**

This paper focuses on the role of the teacher and describes different aspects that can negatively affect students' experiences of using SRS in their lectures. The general feedback from the students at HiST shows a positive attitude towards using SRS as they see it as a tool for both the teacher and the students. Still, the various evaluations during the testing periods have also revealed aspects that can have a negative impact on the students. The most obvious were technical difficulties which were a major issue during the first year of testing. However, this is an obvious pitfall and one that often does not relate to how the teacher uses the system. This paper focuses on aspects that are related to how the teacher uses the SRS during class and his/her implementation choices.

The paper consists of data from all three years of testing, comprising interview analysis and survey results. Although the various interviews had different focuses, they all included a broad spectre of topics, for example, good and bad aspects of SRS, how SRS was used by different teachers, how students experienced group discussions, different methodology and so on. Consequently, both negative aspects as well as students' positive experiences would emerge from the interviews. The results presented in this paper are an accumulation of the most significant teacher-centric issues that can have a negative impact on students' experiences and attitudes towards SRS. The paper covers areas such as consistency when using SRS, time usage, teacher commitment and attitude, the experience level of the teachers with respect to SRS, preparation of the questions, and how students fear the voting results can mislead the teacher.

The paper shows that students are very wary of how time is spent during their lectures. The students have one year to learn approximately two years of senior high school curriculum and thus they react negatively to all uses of SRS that do not benefit learning or student motivation. Some issues can become a significant source of irritation, such as teachers not having a clear goal of SRS-use or not having fully thought through the wording of the questions. One of the consequences of issues such as having repeated mistakes in the questions, inconsistent use of SRS, not having a clear goal or lack of preparation in general, is that it gives students an impression of low teacher commitment towards SRS. Teacher commitment is described by the students as a 'two-way-street' in that the students will lose interest in participating if they get the impression of lack of interest by the teacher. A possible consequence is that this further decreases teachers' motivation towards using SRS in their class, resulting in a 'vicious circle'.

Perhaps one of the more interesting results in this paper is students' concern with how the voting results might affect the teachers. As described earlier, there are several aspects that can affect students' decision making during group discussions, such as influence of stronger students. In addition, students might be very uncertain, but still *guess* the correct answer. The consequence is that the voting results do not necessarily represent the actual level of understanding among the students, who are very aware of this fact. The fear is that the voting results might mislead the teacher into thinking that only a short explanation of the correct alternative is necessary. Consequently, students emphasised the importance of including an alternative named 'don't know'. Such an alternative is to make sure that students can express their uncertainty to the teacher so that (s)he is not misled by the voting results. In other words, they want to use 'don't know' as a tool to 'push' the teacher into giving a more thorough explanation. This tool was so important that students would sometimes refrain from voting if they were uncertain and 'don't know' was not present among the alternatives.

# Paper F

**Student Response Systems in Physics Lectures: Do Voting Results Represent a Correct Image of Student Understanding?**

The last paper focuses on the group discussions of two SRS-quizzes in optics. While the solutions of both quizzes were dependent on very similar principles (even in the end being dependent on the same aspects of the same equation), they had different contexts and required different levels of knowledge of physics. Despite the similar principles, voting results showed that most students (90%) found the correct answer on the first quiz, while being more divided on the second. This motivated further investigation of the arguments used by the students, especially since both quizzes were presented during the same lecture. Both quizzes dealt with interference patterns either through a grid or double slit. In the first quiz, both red and yellow light enter a diffraction grid and the students have to determine how the different wavelengths affect the interference patterns on the screen behind the grid. The second quiz describes a double slit experiment, and the students have to determine what happens to the interference pattern if the whole apparatus, including the light source, is submerged in water.

The analytical process revealed that although students could find the correct answer on the first quiz, many arguments indicated conceptual misunderstanding of the physical principles relevant to the solution. The results indicated that the students used a geometrical optics model to explain diffraction, a physical optics phenomenon. More precisely, many students explained diffraction in terms of using the logic of refraction, and the results indicated that this was more than just a simple word-confusion of scientific terms. Refraction was also an important part of students' argumentation in the second quiz, in which they argued that there was no difference to the interference pattern since the light did not cross a surface between air and water, resulting in no refraction. With refraction established as an important part of the solution in the first quiz, this might have increased the impact of arguments focused on there being no refraction in the second. The voting results showed that students became more convinced of this alternative after discussion. The second quiz also revealed an additional misunderstanding among the students: the relationship between refraction and the refractive index. With no refraction, many students also argued for the irrelevance of the refractive index.

These results highlight important aspects of teaching with SRS. First of all, they directly show that voting results do not necessarily represent students' understanding of the subject matter. Although most students could find the correct alternative in the first quiz, their arguments showed that many had a flawed understanding of the concept of diffraction. Thus, the voting results can give both the teacher and students a wrong impression of students' actual understanding of the subject matter. The results also show the importance of presenting students with several quizzes in the same subject matter, but with different and unfamiliar settings or contexts as this increases the probability of identifying misunderstandings and misconceptions.

# Chapter 5

# Discussion and Conclusions

This thesis has presented research results from using a new Student Response System in the preparatory engineering courses at Sør-Trøndelag University College in Norway. In the included papers, the results are discussed in more detail, for example, with regard to how they relate to previous research. Therefore, this will not be the focus in this chapter. Instead, this chapter focuses on discussing the results as a whole, how the different results relate to each other as well as discussing their implications for teaching with SRS and on future SRS-design.

## 5.1 Connections between the results

This thesis has explored important aspects of implementing and using SRS in lectures, such as the software, role of the teacher, time usage, effectiveness of the questions, positioning of students, and different SRS-sequences. At first glance, the various aspects might seem to cover separate non-overlapping areas with SRS-use, but the results show that there is often a strong connection. For instance, teacher commitment and engagement (paper E) was regarded among the students as one of the most important aspects of SRS. A likely first interpretation of a teacher being committed and engaged, is to regard him/her as enthusiastic and energetic. Although enthusiasm can be an important part of teacher commitment, there are several aspects that can give students an impression of low commitment and engagement, such as being unprepared, fumbling with the software, ineffective use of SRS or simply not showing that SRS-use is to benefit student learning.

Students emphasise that if SRS is to be beneficial to their learning, it is vital that they are given a thorough explanation after the quiz discussion. Students' decision making during discussion is affected by several factors, such as influence of skilled students and seeing the results from the initial voting session with Peer Instruction (paper C). Consequently, the voting results do not necessarily give an actual representation of the level of understanding. While papers C and E describe students' own experiences of this, that they might not have understood the solution even though they choose the correct alternative, paper F shows that voting results should also be handled with care even if the students feel they *have* understood the solution. Despite the majority of students being able to find the correct solution, their arguments showed that many had flawed understanding of the physical principles of the subject matter.

A Student Response System as an assessment tool is also addressed by the students in the interviews in paper B, where they emphasise the difference between SRS as a means to provide challenges and as a means to assess their learning and understanding. Furthermore, the students emphasise that if SRS is to be used as an assessment tool, the quizzes should not be presented too early as students need time to digest the subject matter. This could have been a factor for the results in paper F, where students were used to geometrical optics from the previous weeks and did not have much experience with physical optics.

Papers B, C, and E all show that students are very aware of how SRS is used in their lectures and do not fear criticising inefficient use of the system. Even seemingly trivial aspects such as having a few extra seconds of dead time during SRS-voting can cause irritation (paper E). Not only do these results show the importance of being prepared before using SRS and having a thought-out methodology, but they also show the importance of having a fast and effective SRS-software as described in paper A. Unintuitive and/or cumbersome SRS-software increase the risk of dead time. Additionally, a cumbersome or complex interface is more likely to intimidate novice SRS-users. This can both raise the threshold for teachers utilising SRS as well as increase fumbling with the software, which, in turn, can give the students an impression of low teacher commitment.

## 5.2 Implications on teaching with SRS

One of the motivations behind the broad research focus in this thesis, was to increase its usefulness for teachers wanting to use SRS in their lectures. Although there are good best practice guides on SRS, the results in this thesis show aspects of SRS that are not included in summary papers such as Caldwell [10]. Further, some of the results are even contrary to best practice tips given in the literature, such as the advice by Beatty [12] not to read the questions out loud to the students. The students emphasise that they read together with the teacher as the questions and alternatives are read out loud. Reading out loud has two important roles: 1) it decreases the probability of confusion and misinterpretations of the premise/context of the quiz and what is actually being asked; 2) students will more quickly understand the question, which is important as students react negatively to all aspects that take up more time than necessary.

The quizzes with SRS can be a significant challenge for the teachers. They are the main focus of the SRS-session and can be a major source of irritation among the students if not done properly. The students are not able to 'skip' the questions in the same way that they can skip textbook problems. This is likely to increase the requirements of SRS-quizzes being positively received by the students, compared to other assignments/problems. As one student in paper E put it: 'The questions have to be done properly; it is much easier to regard an SRS-question as ridiculous in comparison to other problems'. The quizzes have to be well prepared without errors and fit naturally into the lectures, and the motivation behind the quiz has to be clear. Does the teacher want to check if students remember yesterday's lecture, by presenting quick repetition questions, or does the teacher want to challenge students' knowledge about the subject matter with conceptual questions? Lack of a clear motivation and consistency behind the questions (and SRS-use in general) can cause irritation if this is repeatedly demonstrated by the teacher.

Although SRS can be a powerful tool to check if students are following the lecture, the results from this thesis have also emphasised issues regarding voting results as a measurement of student understanding. Teachers new to SRS might find these results discouraging and interpret them as voting results having little value, but it should be stressed that this is not the case as the feedback that the voting session provides is highly valued among students. Rather, the research in this thesis highlights the caution that has to be used when interpreting the voting results and how to use them to tailor the lectures. Voting results should be regarded as an indication and not a measurement as students can choose or find the correct alternative without having fully understood the subject matter. Thus, one could argue that SRS is rather a powerful tool to check if students are *not* following the lecture, as voting incorrectly is a strong indication of misunderstandings or gaps in knowledge. However, it does not necessarily have to be the case. If the questions include errors or are poorly phrased, students might not understand what is being asked or the context of the problem, and consequently vote incorrectly. These results emphasise the fact that despite the very simple core functionality of SRS, it is not straightforward to put it to use in lectures. Nevertheless, there are certain factors that can increase the correctness of voting results as a representation of students' understanding. In addition to presenting clear questions, the results in this thesis stress three factors:

1) Students are more susceptible to influence by other group members when not given the time to reflect on the questions on their own before discussion. By including a thinking period and voting session before discussion (Peer Instruction), one can decrease the amount of influence from other group members. The individual voting session prior to discussion also acts as a motivation and it gives the students the opportunity to express their own 'true' opinion before discussion. It should be noted, however, that using Peer Instruction is not optimal in all situations, and depends on the nature of the question. Quick repetition questions based on memorisation, for example, do not benefit from having an additional thinking period as there are no deep cognitive processes involved. On the other hand, it is very valuable with challenging conceptual questions.

2) If Peer Instruction is used, teachers should not display the initial voting results until after the second voting session. According to the students, seeing the initial voting results can have a significant effect on the quality of the discussions and students feel they often become misguided by seeing a clear majority. However, teachers should show and compare both voting results after the second voting session as students find it interesting to see how the class changed its opinions during discussion.

3) Voting results from conceptual quizzes can identify misunderstandings and misconceptions, but can also mislead both the teacher and students if the quiz does

not challenge their misconception. In the first quiz in paper F, the focus of the quiz was on identifying the wavelength and applying the correct equation. It did not test if students understood the concept of diffraction, which the video analysis indicated that many of the students did not fully understand. Teachers have to have a critical awareness of the limitations of the quizzes and what knowledge they test. If teachers want to use SRS as a tool to identify misunderstandings, it is important to include several quizzes on the same subject matter, but where the quizzes have significantly different contexts or settings. The consequence of not doing so is that the voting results can give both the teacher and students an illusion that the subject matter is understood, when there still might be many misunderstandings and misconceptions. Nonetheless, teachers should always try to find a balance between the number of quizzes presented as an exaggerated amount of quizzes can compromise the time needed to go through the rest of the curriculum [40].

All of the three factors above emphasise the importance of teachers thoroughly explaining the solution after the discussion and voting session. This includes explaining the incorrect alternatives as students who believed these to be correct often need to know why they were not correct in order to fully understand the solution. Even if there is a vast majority that have chosen the correct alternative, students still stress that an explanation of the solution is necessary. The explanation functions as a final conclusion, which should remove any doubt about the solution. The students emphasise that voting correctly does not necessarily mean that they have understood the solution.

The results in this thesis also suggest that how students are positioned during group discussion can influence the dynamics of the discussion (paper D). The paper suggests that a communication barrier can arise in groups of three because of students being positioned in a linear manner. If it is not practical to have students seated in a way that physical communication barriers do not occur, for instance, by sitting in a circle, it could be more beneficial to have students discuss in pairs rather than small groups. Although groups of three might be more productive than sitting in pairs, it is important to remember that peer discussion with SRS is not about the productivity of the group, but the learning gains of the individual student.

## 5.3 Implications on future SRS-design

One of the design philosophies behind the Student Response System described in this thesis, was to rely on feedback from teachers and students to constantly improve the workflow, presentation, and functionality of the software. The aim was also that the research performed with the new system could, in addition to increasing knowledge about using SRS in class, further improve the software. The results highlight certain aspects of the system that either could be improved or where new functionality is needed. One such area is the use of Peer Instruction where students value being able to compare the voting results from the first and second vote. This was soon realised during the testing period, and quick changes to the software were made in order for the teacher to be able to compare the two voting results. Due to the nature of the time available to make changes during the testing period, however, this functionality suffered with regard to its intuitiveness which increased the probability of teachers fumbling with the software. As a consequence, steps have been taken to improve this functionality, for instance, by having easy access to earlier votes during the same session. Each voting session during a lecture is listed and teachers can click on several voting sessions to compare results.

In papers E and C, students describe their concerns of teachers being misguided by the voting results not representing a correct image of student understanding. The students want to be able to express their uncertainty to the teacher by, for instance, including an alternative called 'don't know', and some students might even refrain from voting if this option is not included. However, there are both possible benefits and disadvantages with the inclusion of such an alternative (more on these in the next section), and while some teachers want to include such an alternative, others might not because of the possible disadvantages. Instead of teachers having to rely on modifying quizzes if such an alternative should be included, a better solution might be to include it directly into the software as an added option. By having the option on, the software would then send out the alternatives in addition to 'don't know' to the students' voting devices.

Still, students should be able to state their uncertainty even if teachers do not include such an alternative. Some commercial SRSs allow for every quiz having two votes where students first vote on their answer, followed by grading how certain they are. Apart from research purposes, however, the benefits of such a grading system would be limited during lectures where there is little time for evaluation. A possibly better solution could be to add an option which allows students to register their vote as 'uncertain'. This could be a small button visible in the top corner of students' voting devices which can be highlighted when pressed. Students' uncertainty could easily be visualised to the teacher by the top of each column in the voting results' histogram consisting of, for instance, a darker colour representing uncertain votes. Thus, the teacher can quickly and easily visually assess the level of uncertainty among the different alternatives. Adding new options and functionality, however, should be done with caution as the SRS-interface can quickly become complex and, consequently, intimidating and cumbersome to use. Keeping the SRS-interface clean and intuitive should be a constant priority.

# Chapter 6

# Future Research

This chapter gives suggestions for future research on the use of SRS in classroom lectures. This thesis has addressed methodological issues with the use of SRS in lectures and the consequences of implementation choices on, for instance, students' own experiences of the system. However, there are still areas in the results that are in need of more attention. The first paper, A, discusses aspects of SRS-software that might have undesirable effects, such as displaying voting results on the hand-held devices causing students' attention being on the devices during the teacher explanation. This might have the consequence of students missing vital points and thus lowering the learning outcome from the teacher explanations. However, this hypothesis was not tested in this thesis, and might therefore be an area of interest for future studies.

As stated in the previous section, and described in paper E, there are possible disadvantages with including a 'don't know' alternative. Although the students valued having such an alternative, there is a possibility that it could give the students an 'easy way out'. An important aspect of SRS is that students are challenged to deeply reflect upon the problem at hand, and use their knowledge about the curriculum in unfamiliar settings. This can be daunting for students who mostly rely on memorisation when solving problems. Thus, the inclusion of 'don't know' might result in students not making an effort to challenge themselves to find the solution since the quiz presents them with an easier option, i.e., to 'give up'. Another possible scenario, however, could be that, when presented with a difficult quiz, the omission of 'don't know' decreases students' motivation to challenge themselves. The more difficult the question, the higher the probability of students not finding the correct solution or at least not understanding it. Since the students do not wish to guess, they might refrain from trying to find the solution because of the high probability of their not voting at that session and thereby seeing the effort of working out a solution as a waste. Both of these hypotheses are likely scenarios, and future research could entail comparison of two groups where the only difference in the questions is the inclusion or omission of 'don't know'.

An important aspect of SRS that has been addressed in this thesis, is the effects of the initial thinking period on group dynamics. Despite having investigated this topic both from the point of students' own experiences as well as analysing video clips of students discussing, there is still room for more research. While the video analysis provided insight into the effects of the initial thinking period, it did not provide intricate, qualitative details of, for instance, how students are affected by stronger individuals or how the student roles can change when they have the time to reflect on the question before discussion. An analysis based on Positioning Theory, an analytical method that focuses on the roles or social positions that are established between participants in a social context [58], could give more in-depth insight into the subject matter.

The effects that the displaying of the initial voting results have on group discussions are also in need of more research. In the case of a clear majority, students emphasise that they are more inclined to either not bother discussing or to focus all their attention on the alternative in the majority. Although the majority may be correct or incorrect, a solution is often 'forced' on this alternative. It is possible that this might result in a 'groupthink' effect where the group, although consisting of possibly very competent students, ends up with a poor solution to the problem. One student in the interviews stated that it is not even certain that people will admit having chosen an alternative in the minority, but rather try to argue for the alternative in the majority. This suggests that the voting distribution might result in 'conformity pressure', i.e., where the group is unwilling to express ideas that do not match the group norm (or the class norm in this case) in fear of being ridiculed [59]. More research is required to determine the level of conformity pressure, or how the group dynamics change in general, in groups which have seen the initial voting result with Peer Instruction.

## 6.1 Improving SRS-quizzes with video and animations

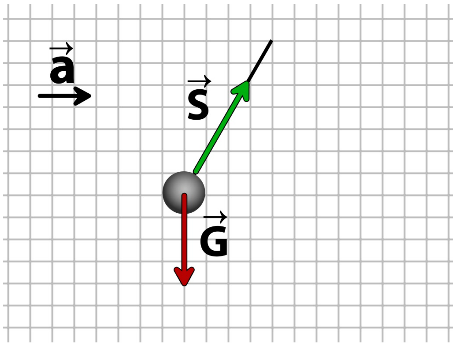
This section describes research in the preparatory physics courses, which is still in an early phase, that symbolises the natural progression of the methodological focus in this thesis – from having the main attention being solely on the *consequences* of methodological choices on the different parts of the SRS-sequences and students' own experiences, towards finding ways of increasing both the learning outcome and motivation of using SRS. The description of this research is included as it addresses important aspects of SRS that were stressed during the different studies in this thesis.

As highlighted in the included research papers, the SRS-questions themselves are a very important part of using SRS in class. The questions can challenge students' knowledge of the subject matter and increase their conceptual understanding [13]. Still, as stressed by the students in this thesis, they can also be confusing and frustrating. During the testing at HiST there were several occasions where students either misunderstood the question or the setting of the quiz. This could lead to irritation among the students. In addition, novice students often suffer from low analytical skills [60], which can likely increase confusion over SRS-questions. In my experience of teaching physics, I have also observed students regarding abstract questions or subject matter as uninspiring. Students' vocational background can be one possible factor for their lack of motivation for abstract subject matter.

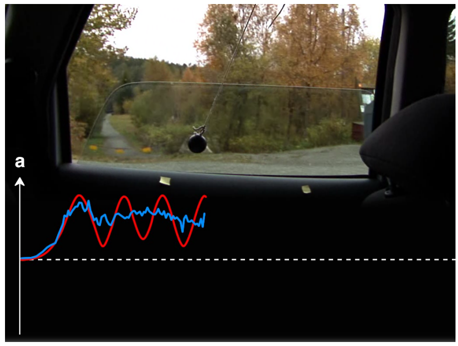
In order to improve the way SRS-questions are presented, in this case in physics, a set of teacher-controlled animations and/or videos were developed to describe the setting of the quiz as well as aid the teacher with explaining the solution. For example, students were presented with a time-lapse sequence of ice in a bowl of water and asked what happens with the water level when the ice melts. The quiz was followed up with a similar video and quiz, but where a small weight was put on top of the ice. A simple animation, which includes a control scheme so that the teacher can progress the animation at his/her own pace, aids the teacher in explaining the solution. The use of the video increases the impact of the results, and students can see that the abstract model in the textbook applies to the 'real' world, which can help them become more convinced of the solution.

Another example is the classic pendulum in an accelerating car. Most introductory physics textbooks have this example where the students are asked to calculate the acceleration of the car based on the angle of the pendulum. The textbooks often illustrate this example with a pendulum hanging at a constant angle. The quiz showed a video of a pendulum in a car starting at rest, accelerating up to a certain speed and then slowing down to a halt. The car included an accelerometer and students were shown the graph from this accelerometer compared to acceleration calculated from the angle. The video shows the graphs being generated in real time, and students were then asked why the two graphs did not look the same. The example not only gives the students an opportunity to see the abstract model in their textbook applied to a real life situation, but also to reflect and discuss (through the quiz) the effects of simplifications in the physics model described in their textbook. Figure 1 shows snapshots from this quiz.

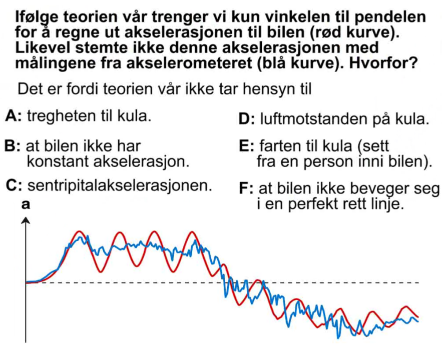
The use of video/animations can make it easier for students to picture and understand what the teacher explains as well as function as a motivation by using real life examples. The video/animation combination functions as a bridge between the abstract world of the textbook and real life. However, using video and animations can potentially lose much of its learning potential if students are not cognitively engaged when watching the videos [61]. By having every video/animation be combined with a quiz, students are likely to be more cognitively engaged when watching the videos since they know they will have a quiz about the subject. This can, in turn, increase the learning outcome of the animated solutions since students have time for deep reflection about the subject prior to watching the solution. This is one of the hypotheses that needs to be examined.



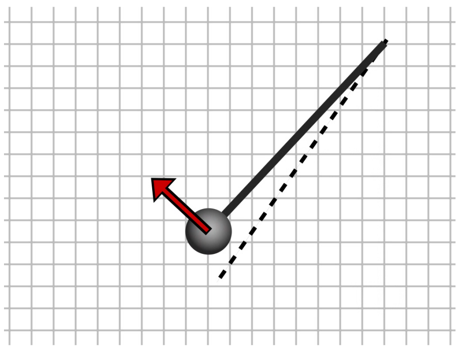

(a) Animation: Theory




(b) Video: Real life example




(c) Quiz



(d) Animation: Solution

Figure 1: Snapshots from an example of video/animation-aided quizzes. The session starts with an animation showing the theory of a pendulum in an accelerated car. Students are shown a video of a real life example, where the theoretical value of the car's acceleration (calculated from the angle of the pendulum) is compared to an accelerometer, followed by a quiz asking why the two graphs do not look the same. An animation aids the teacher explaining the solution.

# Bibliography


[1] Stymne, B., Nygård, M., Borre, K., Lundsgaard, A., and Zarrabi, S. (2008) Evaluering av ingeniørutdanning i Norge 2008. Tech. rep., Nasjonalt Organ for Kvalitet i Utdanning.

[2] Knights, R. (2002) *Five Easy Lessons. Strategies for Successful Physics Teaching*. Addison Wesley.

[3] Duit, R. and Treagust, D. (2003) Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, **25**, 671–688.

[4] Hein, T. (1999) Using writing to confront student misconceptions in physics. *Europ. J. Phys.*, **20**, 137–141.

[5] Masikunas, G., Panayiotidis, A., and Burke, L. (2007) The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, **15**, 3–20.

[6] Gleason, W. (1986) Better communication in large courses. *College Teaching*, **34**, 20–24.

[7] Geski, J. (1992) Overcoming the drawbacks of the large lecture class. *College Teaching*, **40**, 151–155.

[8] Gokhale, A. A. (1995) Collaborative learning enhances critical thinking. *Journal of Technology Education*, **7**, 22–30.

[9] Palincsar, A. S. (1998) Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, **49**, 345–375.

[10] Caldwell, J. (2007) Clickers in the large classroom: current research and best-practice tips. *CBE Life Sci. Educ.*, **7**, 9–20.

[11] Duncan, D. (2006) Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review*, **5**, 70–88.

[12] Beatty, I. (2004) Transforming student learning with classroom communication system. *Research Bulletin*, **3**, 2–13.

[13] Mazur, E. (1997) *Peer Instruction: A user's manual*. Upper Saddle River, NJ, Prentice Hall.

[14] Hansen, G. (2008) *'Men pausen tilhørte liksom faget'. Studenters opplevelse av bruken av studentresponssystem på forelesningen*. Master's thesis, Norwegian University of Science and Technology.

[15] Draper, S. and Brown, M. (2004) Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, **20**, 81–94.

[16] Hake, R. (1998) Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, **66**, 64–74.

[17] Rao, S. and DiCarlo, S. (2000) Peer Instruction improves performance on quizzes. *Advances in Physiology Education*, **24**, 51–55.

[18] Crouch, A. and Mazur, E. (2001) Peer Instruction: Ten years of experience and results. *American Journal of Physics*, **69**, 970–977.

[19] Kay, R. and Lesage, A. (2009) Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers and Education*, **53**, 819–827.

[20] Judson, E. and Sawada, D. (2002) Learning from past and present: Electronic response systems in college lecture halls. *Journal of Computers in Mathematics and Science Teaching*, **21**, 167–181.

[21] Cuban, L. (1986) *Teachers and machines: The classroom use of technology since 1920*. New York, NY: Teachers College Press.

[22] Chandler, R. (2009) Dynamic visualisations and hypermedia: Beyond the "wow" factor. *Computers in Human Behavior*, **25**, 389–392.

[23] Dufrense, R., Gerace, W., Leonard, W., Mestre, J., and Wenk, L. (1996) Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, **7**, 3–47.

[24] Poulis, J., Massen, C., Robens, E., and Gilbert, M. (1997) Physics lecturing with audience paced feedback. *American Journal of Physics*, **66**, 439–441.

[25] Lasry, N. (2008) Clickers or flashcards: Is there really a difference? *The Physics Teacher*, **46**, 242–244.

[26] Graham, C., Tripp, T., Seawright, L., and Joeckel, G. (2007) Empowering or compelling reluctant participators using audience response systems. *Active Learning in Higher Education*, **8**, 233.

[27] Nicol, D. and Boyle, J. (2003) Discussion in large classes: a comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, **28**, 457–473.

[28] Barber, M. and Njus, D. (2007) Clicker evolution: seeking intelligent design. *CBE Life Sciences Education*, **6**, 1–8.

[29] Duschl, R. A. and Osborne., J. (2002) Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, **38**, 39–72.

[30] Horowitz, H. M. (1998) *Student response system: Interactivity in a classroom environment*. White Plains, NY: IBM Corporate Education Center.

[31] Burnstein and Ledermann (2001) Using wireless keypads in lecture classes. *The Physics Teacher*, **39**, 8–11.

[32] Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J., Guild, N., and Su, T. (2009) Why peer discussion improves student performance on in-class concept questions. *Science*, **323**, 122–124.

[33] Len, P. M. (2007) Different reward structures to motivate student interaction with electronic response systems in astronomy. *Astronomy Education Review*, **5**, 5–15.

[34] James, M. (2006) The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, **74**, 689–691.

[35] Trees, A. R. and Jackson, M. H. (2007) The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, **32**, 21–40.

[36] Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., and Cooper, S. (2011) Does displaying the class results affect student discussion during Peer Instruction? *CBE Life Sciences Education*, **9**, 133–140.

[37] Brooks, B. J. and Koretsky, M. D. (2011) The influence of group discussion on students' responses and confidence during Peer Instruction. *Journal of Chemical Education*, **88**, 1477–1484.

[38] Boyle, J. and Nicol, D. (2003) Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal*, **11**, 43–57.

[39] Rice, R. and Bunz, U. (2006) Evaluating a wireless course feedback system: The role of demographics, expertise, fluency, competence, and usage. *Studies in Media and Information Literacy Education*, **6**, 1–32.

[40] d'Inverno, R., Davis, H., and White, S. (2003) Using a personal response system for promoting student interaction. *Teaching Mathematics and its Applications*, **22**, 163–169.

[41] Stuart, S. A. J., Brown, M. I., and Draper, S. W. (2004) Using an electronic voting system in logic lectures: one practitioner's application. *Journal of Computer Assisted Learning*, **20**, 95–102.

[42] Hatch, J., Jensen, M., and Moore, R. (2005) Manna from heaven or "clickers" from hell. *Journal of College Science Teaching*, **34**, 36–39.

[43] King, S. and Robinson, C. (2009) Formative teaching: A conversational framework for evaluating the impact of response technology on student experience, engagement and achievement. *39th ASEE/IEEE Frontiers in Education Conference, October 18-21.*.

[44] Draper, S. (2002), Ensuring effective use of prs: results of the evaluation of the use of PRS in Glasgow University, October 2001 - June 2002. `www.psy.gla.ac.uk/~steve/ilig/papers/eval.pdf` (accessed 9 January 2012).

[45] University of Wisconsin. *Student Response Systems*. `http://www4.uwm.edu/ltc/srs/faculty/best_practices.cfm` (last checked 15 July 2012).

[46] Reay, N., Li, P., and Bao, L. (2008) Testing a new voting machine question methodology. *American Journal of Physics*, **76**, 171–178.

[47] Knight, J. and Wood, W. (2005) Teaching more by lecturing less. *Cell Biology Education*, **4**, 298–310.

[48] Johannessen, A., Tufte, P. A., and Kristoffersen, L. (2004) *Introduksjon til Samfunnsvitenskapelig Metode*. Abstrakt Forlag AS, Oslo, Norway.

[49] Wilkinson, K. (2003) *Rethinking methods in psychology*, chap. Focus Groups, pp. 184–204. London: Sage Publications.

[50] Berg, B. L. (2007) *Qualitative research methods for the social sciences*, chap. Focus group interviewing, pp. 144–170. Boston, MA: Pearson Education.

[51] Charmaz, K. (2003) *Qualitative Psychology; A practical guide to research methods*, chap. Grounded Theory, pp. 81–110. London: Sage Publications.

[52] Charmaz, K. (2001) *Rethinking methods in psychology*, chap. Grounded Theory, pp. 27–49. London: Sage Publications.

[53] Glasow, P. (2005), Fundamentals of survey research methodology. `http://www.mitre.org/work/tech_papers/tech_papers_05/05_0638/index.html` (accessed 23 April 2012).

[54] Pinsonneault, A. and Kraemer, K. (1993) Survey research methodology in management information systems: an assessment. *Journal of Management Information Systems*, **10**, 75–105.

[55] Derry, S., Pea, R. D., Barron, B., Engle, R., Erickson, F., Goldman, R., and Sherin, B. (2010) Conducting video research in the learning sciences: Guidance on selection, analysis, technology and ethics. *Journal of Learning Sciences*, **19**, 3–35.

[56] Heath, C., Hindmarsh, J., and Luff, P. (2010) *Video in Qualitative Research. Analysing Social Interaction in Everyday Life*. Sage Publications.

[57] Kaartinen, S. and Kumpulainen, K. (2002) Collaborative inquiry and the construction of explanations in the learning of science. *Learn. Instr.*, **12**, 189–212.

[58] Anderson, K. T. (2009) Applying positioning theory to the analysis of classroom interactions: Mediating micro-identities, macro-kinds, and ideologies of knowing. *Linguistics and Education*, **20**, 291–310.

[59] Stangor, C. (2004) *Social Groups in Action and Interaction*. Psychology Press. New York and Hove.

[60] Knight, J. K. and Smith, M. K. (2010) Different but equal? how nonmajors and majors approach and learn genetics. *CBE Life Sciences Education*, **9**, 34–44.

[61] Rapp, D. (2005) *Visualization in Science Education*, chap. Mental Models: Theoretical Issues for Visualizations in Science Education, pp. 43–60. Springer.

# Paper A

## Designing and Developing a Student Response System for Mobile Internet Devices

Nielsen, K.L., Stav, J.B., Hansen-Nygård G., and Thorseth, T.M.
This chapter appears in *Learning with Mobile Technologies, Handheld Devices and Smart Phones: Innovative Methods* edited/authored by Joan Lu. 

# Chapter 4
# Designing and Developing a Student Response System for Mobile Internet Devices

**Kjetil L. Nielsen**
*Sør-Trøndelag University College, Norway*

**John B. Stav**
*Sør-Trøndelag University College, Norway*

**Gabrielle Hansen-Nygård**
*Sør-Trøndelag University College, Norway*

**Trond M. Thorseth**
*Sør-Trøndelag University College, Norway*

## ABSTRACT

*The authors present a Student Response System for modern Internet-capable mobile devices, which was developed in a European R&D project, co-funded by the European Commission. The goal was to make a system that is designed for speed, ease of use, and flexibility for use in lectures. The authors have tried to make a time efficient and intuitive system that does not compromise flexibility and that enables the teacher to use any lecture format he/she sees fit. The only requirement is a computer with an Internet connection; the teacher is not bound to specific presentation software. The system is Web-based, enabling students to use their own mobile device or computer. The cost for both educational institutions and students is kept at a minimum, lowering the threshold for using the system in education. As of today, the program is free of charge and can be found at histproject.no.*

## INTRODUCTION

A common challenge with traditional class lectures is the communication between teacher and students (Masikunas, Panayiotidis, & Bruke, 2007) and student interactivity (Draper & Brown, 2004). This is especially prominent at college and university levels where lectures often consist of large classes. With classes of up to hundreds of students, direct communication with the students becomes difficult, not only because of the size of the classes, but also because of time constraints.

DOI: 10.4018/978-1-4666-0936-5.ch004

Many students also find it difficult to give feedback or ask questions during class because of fear of being embarrassed in front of their peers (Geski, 1992; Gleason, 1986). The lack of student feedback during class can make it difficult for the teacher to assess if the students understand the subject being taught (Trees & Jackson, 2007). Students are rarely given time to reflect upon what is being taught in traditional lectures and their understanding of the material is rarely tested during class (Masikunas et al., 2007). This can result in a spiral where the students do not understand part of the lecture and risk losing a significant portion of it because of the inability to catch up. In addition to this, students' ability to stay focused falls dramatically after about 20 minutes, a factor that can amplify the spiral (Caldwell, 2007; Duncan, 2006).

One way of dealing with these challenges is using a Student Response System (SRS) (also referred to as Classroom Communications Systems, Electronic Voting System, Class Response Systems, and Audience Response Systems), a technology designed to provide communication and interactivity in large classrooms (Beatty, 2004). In a nutshell, an SRS is a technology that enables the teacher to ask questions to the students, often in the form of multiple-choice problems, and the students respond with a small handheld device, often referred to as ''clickers''. Responses can be given anonymously, lowering the threshold for student participation in the classroom. Studies have shown that anonymity is a key factor for students having a positive evaluation of lectures in large classrooms (Wulff, Nyquist, & Abbott, 1987) and many students appreciate the anonymity of the responses given from SRS (Trees & Jackson, 2007; Stuart, Brown, & Draper, 2004). Teachers can get feedback showing if the students follow the lecture as well as give students time to reflect upon the subject and see if they understand what is being taught (Dufrense, Gerace, Leonard, Mestre, & Wenk, 1996; Stuart et al., 2004).

There are several methodical approaches for using a Student Response System, for instance, Peer Instruction (Mazur, 1997). Students are given conceptual quizzes during class where they first give individual responses (without talking to their peers), followed by a discussion in small groups and respond once more. The teacher then goes through all responses and explains the correctness and incorrectness of the alternatives. Another approach is to omit the first vote and go straight to group discussion followed by a class-wide discussion among both students and teacher (Dufrense et al., 1996). For more methodical discussions of use of SRS, see Horowitz (1988) and Crouch (2001). For a comparison of the methodical approaches mentioned above, see Boyle and Nicole (2003) and Nicole and Boyle (2003). Several studies of SRS show that the students are satisfied with the opportunity to discuss in class and that they are given a chance to reflect and think about the subject being taught (Hansen, 2008; Masikunas et al., 2007). They see the benefits of SRS and usually do not mind that time is taken away from the ordinary lecture in order to be used for discussing and teacher explanation of quiz alternatives (Hansen, 2008). However, there are negative to dead-time, usage of lecture time that does not benefit to learning (waiting for the system to start, handing out clickers, technical problems and so on) (Caldwell, 2007; Hansen, 2008).

Most common commercial SRSs have consisted of systems that use either infrared- or radio senders and receivers, and some form of dedicated software. This is either stand-alone software or a plugin in common presentation software like PowerPoint, the latter being the most common solution. Infrared- and radio-based SRSs have several cost- and practical disadvantages. Receivers have to be installed in each classroom using the system, or at least a mobile receiver has to be brought each time the system is going to be used. These systems are often expensive when both receivers and clickers have to be bought, and some also have expenses for technical support. Some universities have made it obligatory for the students to buy their own clicker; however, the lack of standardization has resulted in some universities using more than one commercial SRS. Often the designated clicker only works for that particular system and consequently, students have to buy one clicker for each SRS used at the university (Duncan, 2006).

The introduction of Apple iPhone and iPod Touch in summer 2007 spawned a new trend of smart-phones and mobile devices. There have been many mobile phones with Wi-Fi before the iPhone, but they were often difficult to use because of small screens with low resolution and cumbersome navigation. They were not built for effective web browsing. With modern smart-phones with large touch-sensitive screens, Internet use is more practical, leading to a new generation of Student Response Systems (sometimes referred to 3rd generation Student Response Systems). In the last few years several Internet-based systems have emerged. Turning Point Technologies in the US has developed a system for iPod Touch and Blackberry (http://www.turningtechnologies.com/), while the university of Austin developed a system for Blackberries, netbooks and iPod Touch (Moca, 2009). Also iClicker have a system compatible with several devices, including mobile devices with Android 1.5 or later (http://www.iclicker.com/dnn/).

As a part of the Edumecca project (http://prosjekt.hist.no/edumecca), which was co-funded by the European Commission during the period of 2009-2010, we have designed and tested an easy, fast and flexible web-based SRS for modern mobile devices and PC/MAC. In this article, we will first present our design and development philosophy and what we wanted to achieve with our software. We then present our Student Response System where the basic functionality is explained, followed by a discussion and conclusion.

## DESIGN AND DEVELOPMENT PHILOSOPHY

There are numerous examples throughout history of new technologies that were supposed to revolutionize education, but that failed to do so (Cuban, 1986). For instance, the radio was predicted to be as common in classrooms as blackboards and motion pictures were going to replace most of the textbooks (Mayer, 2005). The main reason why these inventions did not live up to the predictions is that the focus was on the technology itself and not on how students think and learn (Mayer, 2005). Technology by itself does not facilitate learning, but can act as a helpful tool for teachers to make their lectures more effective. Chandler (2009) summarized the consequences of a technology-based approach rather fittingly:

*If we have learned anything from the history of the use of technology in education in the past 20 years, it is that as soon as learning processes are not a core consideration and pure technological capabilities, functionality, and the "wow" factor are made central, then chaos inevitably ensues (p. 392).*

We wanted to move away from a technology-based approach when developing our Student Response System. It should not be about how much the system can do, but rather how effective it can be as a tool for the teacher and students in the classroom. We have designed the system for speed, ease of use and flexibility. It should be very fast and time efficient, intuitive and easy to adapt to different lecture- and teaching methods. Another goal was to have it as software- and hardware independent as possible, both for the teacher as well as the students, making the system almost "invisible" and fitting together seamlessly with whatever lecture format the teacher prefers. Although we chose the iPod Touch for testing of the SRS addressed in this article, our goal is to have a system independent of whichever Internet-capable device used by the students. During our testing periods, we used SRS mainly in preparatory physics courses for engineering students at HiST, mathematics and society & technology as well as physics for building engineers, in classes ranging from 40-60 students. We relied heavily on student feedback through surveys and group interviews, dialogue with teachers and observations of the SRS being used in class to find the optimal system for effective classroom use. An overview of experiences and students' feedbacks during these testing periods will be presented in a later paper. For methodological guidelines for use of SRS in teaching, see Nygård-Hansen, Nielsen, Thorseth, and Stav (2011).

Time constraint can be an important limitation when trying to incorporate new elements into an already tight lecture-schedule. This makes it inevitable for compromises to be made when incorporating an SRS as part of a lecture, either it being using less time on each part of the curriculum or removing parts of it (Mazur, 1997). Speed is therefore essential for an SRS in order to minimize compromises to the curriculum. In order to make the system as fast and effective as possible, we have to ask: what is the main purpose of a Student Response System? There are several commercial systems that have a lot of different functionality, but in our view, the main purpose of an SRS is to enable the teacher to present problems/tasks with alternatives to the students, and enable the students to present the teacher with their answer, and doing so as fast and in as few steps as possible. The teacher should be able to start a vote within seconds without having to open up a lot of menus and submenus, and even be able to make up a question spontaneously "on-the-fly" without having to use a lot of time within the system. As well as being time efficient during class, the system should also be time efficient before class. Many commercial systems often need a lot of preparation before class, including having to make the questions in a specific format required by the software, registering students' devices and so on. We wanted an SRS where as little as possible preparation connected to the system was required; the teacher should be able to enter the class, start the program and be ready for class without additional system preparation.

In order to have a flexible system capable of being adapted into numerous lecture formats and methods, we wanted our SRS to be independent of the presentation software used by the teacher. Having it confined to PowerPoint (or similar software) could alienate teachers used to other lecture formats like smart-book notebook or just ordinary blackboard. We wanted the teacher to be able to use whatever format he/she is comfortable with and to enable the teacher to use the SRS without leaving this format. For instance, the teacher should be able to have a PowerPoint presentation in full screen and be able to start a vote without having to leave the full screen presentation. As well as being independent of presentation software, we also want it to be independent of which operation system used; it should work on Microsoft Windows and Mac OS X as well as on GNU/Linux.

The use of digital blackboards, or smart-boards, is becoming more and more common, and we therefore wanted the SRS to be designed for effective use in digital blackboards (although a smart-board is not required). In order to minimize steps and maximize efficiency, the teacher should be able to operate the SRS without having to walk away from the digital blackboard. Any unnecessary steps away from the digital blackboard, to a computer mouse or keyboard, are disturbing and not optimal when teaching is in focus. The SRS must be fast and intuitive to call up when needed without having to interrupt the flow of the presentation. When called up, the SRS should not be in the way of the lecture presentation and the teacher should be able to show the results together with the question.

In order to achieve this, we chose a minimalistic approach. As well as being minimalistic in visual appearance, having the system use as little of the screen as possible, the teacher should also not be overcome with options and choices. This can quickly become more of a disturbance than a benefit, and one also risks alienating teachers not used to advanced computer programs. As stated before, the main functionality of the system should be to send out buttons to the students' devices and collect responses. More advanced options, which the teacher is less inclined to use, should therefore not be intrusive and be hidden in underlying menus.

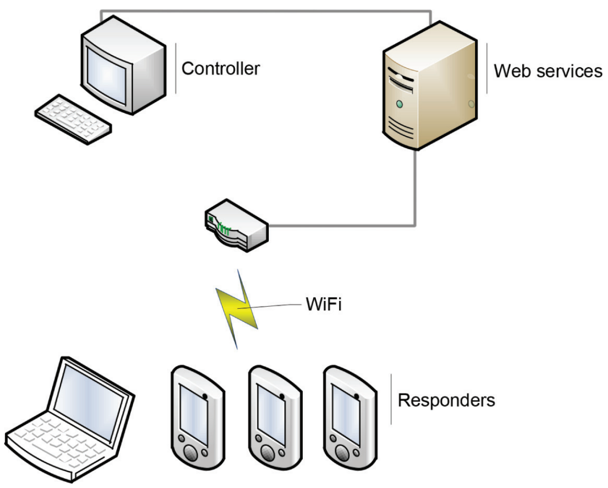
A summary of our design philosophy regarding the SRS can be listed in the following points:

- The SRS should be designed for speed, being easy, intuitive and time efficient to use in teaching situations.
- The SRS should be flexible towards different lecture methods and formats.
- The teacher should be able to present the questions in the format he/she sees fit.
- The teacher should be able to make questions "on-the-fly" during class and start votes within seconds.
- The SRS should be easy to use from a digital blackboard without having to leave the blackboard.
- The SRS should require minimum system preparation before class.
- Students should be able to use their own mobile device, PC/MAC or any combination of those.

## THE STUDENT RESPONSE SYSTEM

The SRS has been developed for web and uses the HTTP protocol to handle all Internet communication both from the student perspective and teacher perspective. The technical details are presented elsewhere (Pein, Lu, Stav, & Thorseth, 2011). The initial SRS was designed purely as a web interface that had two sides, a teacher control interface and the student responder interface. The teacher interface was later redesigned as an Adobe AIR (Flash) application (while the student interface remained a web page) for reasons that will be discussed later in this article. Figure 1 shows a simplified overview of the information

*Figure 1. Simplified overview of information flow in the SRS (© 2011, Sør-Trøndelag University College)*



flow in the system. The teacher (controller) and the students (responders) are connected to a server that handles requests from the teacher and sets up the web page that the students use. Modern mobile devices have the ability for easy access to bookmarked web pages, like on the iPhone where bookmarks can be placed on the home screen, similarly to apps. The student interface can be seen in Figure 2.

Each time the teacher starts the SRS, a three-letter session code is generated. The students have to enter this session code in the web page and an optional name. If a name is not entered, the vote is anonymous. By entering this code, the students are linked to the given session, prohibiting interference from other classes using SRS at the same time. Beyond this no other registration of devices are needed.

In a nutshell, the teacher is able to send out alternatives to the students' devices in the form of generic buttons. The teacher is presenting the questions in the format he/she sees fit. The buttons could be ''A'', ''B'', ''C''… (at the default the teacher chooses from 3 to 6 alternatives, but there is also a possibility to expand up to 20 alternatives), ''yes/no/don't know'' or ''true/false''. Figure 3 shows the student interface during a vote with four alternatives. Students get visual feedback when their vote is confirmed by change of background color. Students can change their mind by pressing another alternative as long as the vote is open. The SRS register the students' answers and results are shown in a histogram on the teacher's computer screen or digital blackboard when the vote is closed. The results are not shown on the students' devices. Students only have to consider generic buttons on the voting device; all other activity is centered on the teacher's presentation.

The teacher interface is a stand-alone program designed as a transparent film that is always on top of other programs. Since the SRS is always on top, it can be accessed at any time without having to leave the presentation format of the lecture, be it PowerPoint, PDF, Smart Board Notebook or any other presentation software. At the default state the SRS is mostly hidden from the user with only a hidden toolbar on the right side of the screen visible. The toolbar can be made visible by hitting the visible top part as seen in Figure 4. Making the toolbar visible also makes the session code visible in the top left corner as seen in Figure 5 (note that the session code is also visible on the student interface at all times after it has been entered). The SRS still takes up only a small part of the screen in order to have most of the lecture presentation still visible.

The teacher can now start a new vote by pressing "Run Vote". This will make a menu appear in a separate window as seen in Figure 6. This window can be moved and scaled as the teacher sees fit. There are several options from which the teacher can customize the vote, but the teacher can choose to ignore these options and just go with the default state. By default the results are shown automatically when the vote is closed. This can turned off, so that the results are not shown automatically. A ticking sound is used to tell students that the voting period is open. This sound can be disabled. The system has a time limitation in voting period, this time limit can be changed or disabled. By default the system gives the students a chance to choose only one alternative, but this can be changed to accept multiple choices.

When the teacher chooses the amount of buttons he/she wants to send out to the students' devices, a small control interface appears as seen in Figure 7. The teacher can start, pause or stop the vote (either manually or by waiting for the timer to run out), making the results appear as seen in Figure 8. The results are displayed as a histogram where bars can be highlighted by clicking on them. The teacher can also perform two votes where the results of both votes are only shown after the second vote. This is done by unchecking the "show results" option, starting a vote, checking the "show results" option and starting an identical vote. The results are then shown side by side for comparison.

Results from all votes are stored and can be retrieved from the SRS by going into a separate menu with the "History" button. Every student vote is registered with a timestamp and the alternative chosen. The system also registers if the student changed his/her mind during the voting process. However, it should be stated that this functionality of our SRS is under development. With the "Redirect" button, the teacher can forward the students to an external web page. The students will then receive a link instead of buttons. This can, for instance, be used for evaluation by sending the students to surveys made with services like Google Forms or to any web page relevant. Last, the teacher has the opportunity to configure the appearance of the control interface, changing color, opacity and so on.

*Figure 2. The student interface. In the right figure, the session code, here as "ncf", can be seen in the lower left corner (© 2011, Sør-Trøndelag University College)*

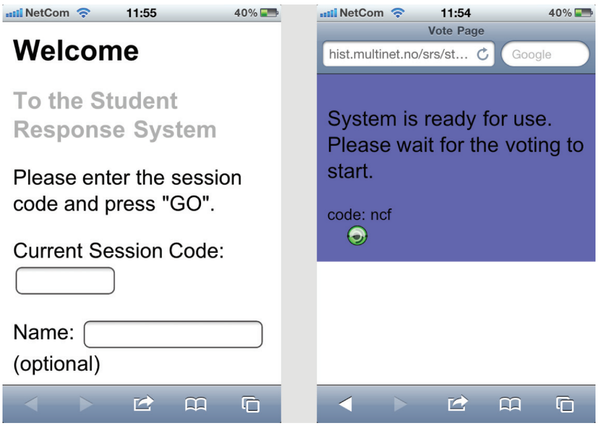



*Figure 3. The student interface while voting (© 2011, Sør-Trøndelag University College)*

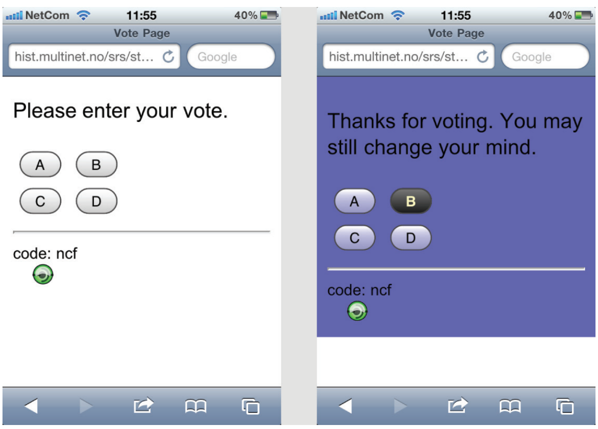
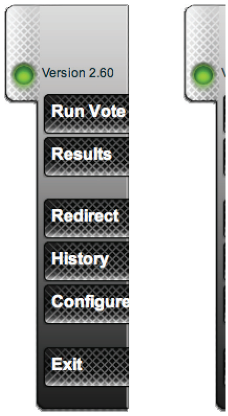


*Figure 4. The SRS toolbar when called out (left) and when hidden (right) (© 2011, Sør-Trøndelag University College)*




## DISCUSSION

One of the earliest questions that arose when designing our SRS was if questions and results should be presented on the students' devices. There was a version of our SRS where alternatives with picture and text could be presented, and the possibility of showing video on the devices was also discussed. Most 2nd generation clickers (with infrared/radio-senders) do not have a large, high-resolution screen (many do not even have screens) and it is therefore not practical or possible to show the question on the device. When starting to use modern mobile devices, like smart-phones, as "clickers," it may seem like a logical step to show the question and results on the students' devices since we now have the technology to do so. However, this assumption is problematic because it puts the technology in focus. There is a methodical as well as a practical argument that were the basis for the decision not to show the results and question on the devices.

We want the teacher to keep full control of the students' attention. By sending out the questions and results on the devices, the students' attention is taken away from the teacher and lecture and brought down towards the voting device. The teacher will risk losing the attention of the students, and have problems getting the students to switch their focus forward again after a discussion. We observed the importance of having full control over the students' attention when we first started using the system. The first version of the SRS did not have a timer and a ticking sound during voting. The students would start discussing during a quiz, but because of the volume of the discussion, the teacher had a difficult time getting the students' attention towards the blackboard. Many students even missed the vote because they did not realize that a vote had taken place. By keeping the focus directed forward, the teacher can maintain contact with the students and more easily assess if the question was understood. We see it as even more important not to show the results on the devices. The focus should be directed towards the teacher and his/her explanation after voting and not at the device.

There is, however, one major disadvantage by only sending out generic buttons on the devices and not having the questions and the responses linked. The system does not know what the question was when storing the results. If the teacher wants to examine the results for an earlier question, he/she has to remember what question was asked when and make that connection to the results manually (i.e., by hand). This is one advantage of SRSs connected to specific software like PowerPoint where the questions have to be made in a specific format and registered. There are similar SRSs where the questions and responses are not linked that, but where this is compensated for by having the program take a screenshot of the question with the responses (Barber & Njus, 2007). A logical next step for our SRS is to include a database for storing questions. Our system is a great and simple tool for the teacher who wants to use a SRS in class, but some teachers might find it difficult finding good questions and cumbersome to always have to search and prepare questions,

*Figure 5. The SRS as seen on the teachers' computer screen when the toolbar is made visible (© 2011, Sør-Trøndelag University College)*

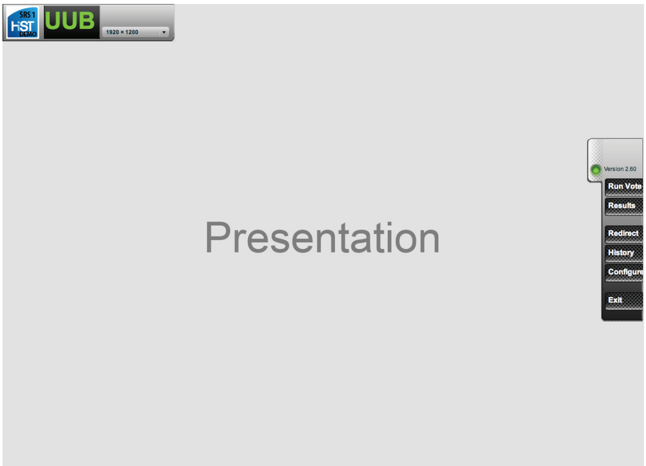



*Figure 6. The voting menu in the default state (left) and after "more" is clicked (right). The teacher chooses the type of buttons that will be sent out to the students' devices. The numbers (right figure) refer to the number of buttons that will be sent to the students voting device (© 2011, Sør-Trøndelag University College)*

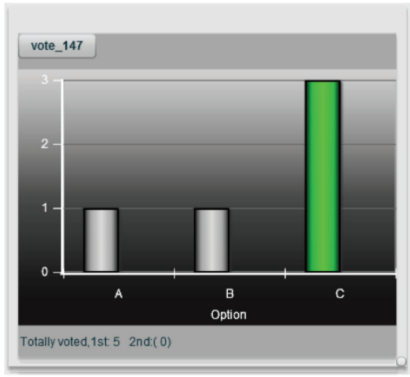

and would much rather like to search a database. This way we would also have a link between the question and the results stored in the program. However, implementing database functionality should be an addition and not replace the simple functionality of the SRS as it works today in order not to compromise the ease of use and flexibility of the system.

*Figure 7. Voting controller. The teacher can start, pause or stop the vote (© 2011, Sør-Trøndelag University College)*


When using Student Response Systems with Wi-Fi, the network is an Achilles heel. It is therefore a good idea to minimize the traffic on the network when a large amount of students are connected and receive information at the same time. Data traffic is an issue if one wants to send out large images or video to a large amount of devices, but we have experienced that even a small amount of data can also cause network problems if the number of students are high. Surveys and focus interviews of the students showed that classes with network problems were significantly less satisfied with SRS than other classes that did not have network problems. To keep data traffic at a minimum and to maximize the stability of the system (and the maximum amount of students the system can handle), it is beneficiary to send out a minimum amount of data to the units. Another benefit is if the students have to use their mobile networks in the absence of a Wi-Fi connection. Having data traffic at a minimum will then also minimize the cost for the students.

Although it seemed logical in the beginning to have both the teacher and students using the system through a web page, to be as independent of technology as possible, we experienced that using a web interface for the teacher creates several flow problems during a vote. For the teacher to start a vote and present the results, he/she must have a web browser on top of all other software used by him/her. This means leaving the format of the original lecture, which acts as a disturbance. The web page will cover the lecture material that the teacher is presenting and it will be difficult

*Figure 8. Results after a vote. Each bar can be highlighted by a single click (© 2011, Sør-Trøndelag University College)*




to show the results together with the lecture presentation. The teacher can of course scale down the web page to a small window, but this can be cumbersome when having to constantly readjust the window to either start a vote or show the results. Using a dedicated control application (as opposed to a web page) gives several advantages in the possibilities for the design of the software, and the applications can communicate easily with web services.

Technical problems are a very common source of dead-time when using a Student Response System, but time can also be lost by the SRS being cumbersome to use. If the teacher has to leave the original presentation format or if a large number of clicks and choices are needed to start a vote, precious seconds will be lost. The cumulative effect during a whole lecture can make students' satisfaction drop significantly. Not many seconds are needed before students become dissatisfied. A cumbersome program, needing several clicks and choices, also raises the threshold of teachers actually using a SRS in class. With our SRS, the teacher can start a vote with as few as 3-4 clicks (4 if the toolbar is hidden). The buttons are big and can easily be used with the touch of a finger from the digital blackboard as well as from a mouse. The layout of the menus and buttons makes it very fast to start a vote since the teacher does not need to go into submenus or drop-down menus. His/her options will be clearly shown and within reach, making it easy to start a vote within seconds. Making the session code visible on the top, and on every device at all times, makes it also more practical for students showing up late to class since it is unnecessary to stop the lecture to bring forth the code, saving additional seconds. By using only three letters and not mix numbers and letters, the session code is easy and fast to enter using the virtual keyboard on the voting devices.

Since we used iPod Touch for testing our SRS, it might have seemed logical to make the student interface as an app in app-store. However, this solution would limit the system to Apple products. One solution would have been to make a version for each OS (iOS, Android, S60 and so on), but this can cause logistical problems. Some students may still have a device for which the system was not designed and the teacher is dependent on the students having the latest version of the system. With a web-based interface, OS does not limit which device students can use. However, it raises the question, how many students have a modern mobile device where Internet use is practical? In our testing of the SRS, we lent out an iPod Touch to each student to make sure that everyone was able of participation. However, in a survey in two classes using the SRS, asking how many students had an own device they could have used for voting (including laptops), 68% out of 57 answered that they had a capable device like a smart phone or laptop. In a few years most students will most likely have a smart phone or similar that could be used for SRS. By using their own smart phone, students have one less device to think about. In a study by Draper & Brown (2004) the researchers experienced 25-35% of students forgetting their clicker. As students are very unlikely to forget their own mobile phone, this problem will be almost non-existent. There have also been reports where 5-10% of the students never bought or registered their clickers (Hatch & Jensen, 2005).

Ease of use and flexibility are two properties of a system that often is hard to make to coexist, as ease of use compromises the flexibility and vice versa. We have already shown that the system is flexible by having the SRS be software independent, but we also wanted a system flexible in different methodical SRS uses. For instance, the teacher should be able to use the Peer Instruction method (which includes two votes for each quiz, one before and after group discussion) without compromising the method. Our surveys and interviews showed that the students preferred having an individual vote before group discussion. They felt that showing the results of the first vote, however, could negatively influence the group discussion, that the focus of the discussion becomes centered around the alternative that had the majority of votes and not a discussion on each alternative to find out which was right or wrong (these findings will be discussed in a later article). The teacher should therefore be able to show both results of the votes at the end of the second vote and not show the results of the first vote without compromising the ease of use. Our SRS is also flexible to use outside of regular classroom lectures, for instance, distance learning and guest lectures (where the attendants are only present once). The teacher/speaker can easily use our SRS in these and similar situations with ease because of the lack of pre-class preparation needed for our SRS. The attendants do not need to download any software or register devices. They could just be presented with the link to the student page and given the session code in order to participate in votes.

## CONCLUSION

We have presented a Student Response System made for mobile Internet capable devices that is easy, intuitive and flexible to use in teaching situations. The system was designed for speed, making it ideal for large-scale lectures where time constraints are of the essence. The student interface is designed as a web page enabling students to use any device capable of showing web pages and not being required to buy commercial clickers. The teacher uses a dedicated control interface not connected to a specific program like PowerPoint, enabling him/her to use whatever presentation tool comfortable to the teacher. The SRS was designed for effective use with digital blackboards and the teacher can operate the system with ease from the blackboard without having to use a keyboard and mouse. The SRS enables the teacher to present questions in the format of his/her choice and start a vote within seconds without any need to prepare the question in the software, also making the SRS ideal for quick spontaneous "on-the-fly" questions during class. Although there is a limitation in that the question posed by the teacher is not linked to the responses collected by the SRS, the system allows for an extremely simple and effective means to receive feedback from the students and test their knowledge during class. Since only a computer with an Internet connection is required for the teacher and students may use their own web-capable device, this system provides a cost effective solution to educational institutions wanting to utilize a SRS. It is very likely that most students will have a modern web-capable device like smart phones within a few years, which will lower the threshold for using an SRS within a wide variety of educational institutes, some of which may not afford commercial systems.

## ACKNOWLEDGMENT

This project has been funded with support from the European Commission, contract 143545-2008-LLP-NO-KA3-KA3MP. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

## REFERENCES

Barber, M., & Njus, D. (2007). Clicker evolution: Seeking intelligent design. *CBE Life Sciences Education*, *6*, 1–20. doi:10.1187/cbe.06-12-0206

Beatty, I. (2004). Transforming student learning with classroom communication systems. *Research Bulletin (Sun Chiwawitthaya Thang Thale Phuket)*, *3*, 2–13.

Boyle, J. T., & Nicol, D. J. (2003). Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal*, *11*(3), 43–57. doi:10.1080/0968776030110305

Caldwell, J. E. (2007). Clickers in large classrooms: Current research and best practice tips. *CBE Life Sciences Education*, *6*, 9–20. doi:10.1187/cbe.06-12-0205

Chandler, R. (2009). Dynamic visualisations and hypermedia: Beyond the "wow" factor. *Computers in Human Behavior*, *25*(2), 389–392. doi:10.1016/j.chb.2008.12.018

Crouch, A. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970–977. doi:10.1119/1.1374249

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York, NY: Teachers College Press.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, *20*, 81–94. doi:10.1111/j.1365-2729.2004.00074.x

Dufrense, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, *7*, 3–47. doi:10.1007/BF02948592

Duncan, D. (2006). Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review*, *5*, 70–88. doi:10.3847/AER2006005

Geski, J. (1992). Overcoming the drawbacks of the large lecture class. *College Teaching*, *40*, 151–155.

Gleason, W. (1986). Better communication in large courses. *College Teaching*, *34*, 20–24.

Hansen, G. (2008). *...og pausen tilhørte liksom faget" about use of student response systems in large classes* (Unpublished master's thesis). Norwegian University of Science and Technology, Trondheim, Norway.

Hatch, J., & Jensen, M. (2005). Manna from heaven or "clickers" from hell. *Journal of College Science Teaching*, *34*(7), 36–39.

Horowitz, H. M. (1988). *Student response system: Interactivity in a classroom environment*. White Plains, NY: IBM Corporate Education Center.

Masikunas, G., Panayiotidis, A., & Bruke, L. (2007). The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, *15*(1), 3–20. doi:10.3402/rlt.v15i1.10903

Mayer, R. (2005). *The Cambridge handbook of multimedia learning*. Cambridge, UK: Cambridge University Press.

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.

Moca. (2009). *Gathering instant student feedback on mobile devices* (Case study series). Retrieved from http://www.utexas.edu/academic/diia/about/postcards/casestudy4 moca.pdf

Nicole, D., & Boyle, J. (2003). Discussion in large classes: a comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, *28*(4).

Nygård-Hansen, H. M., Nielsen, K. L., Thorseth, T. M., & Stav, J. B. (2011, June 1-2). *Developing and evaluating practical methodological guidelines for use of student response system in teaching.* Paper presented at the International Conference on Education & Learning in Mobile Age, Lake District, UK.

Pein, P., Lu, J., Stav, J. B., & Thorseth, T. M. (2011). *A robust and scalable restful web service for collecting student responses - student response system (srs)*. Paper presented at the International Conference on Education & Learning in Mobile Age, Lake District, UK.

Stuart, S. A. J., Brown, M. I., & Draper, S. W. (2004). Using an electronic voting system in logic lectures: One practitioner's application. *Journal of Computer Assisted Learning*, *20*, 95–102. doi:10.1111/j.1365-2729.2004.00075.x

Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems learning. *Media and Technology*, *32*(1), 21–40. doi:10.1080/17439880601141179

Wulff, D. H., Nyquist, J. D., & Abbott, R. D. (1987). Students' perceptions of large classes. *New Directions for Teaching and Learning*, *32*, 17–30. doi:10.1002/tl.37219873204

# Paper B

## Developing and Evaluating Practical Methodological Guidelines for Use of Student Response System in Teaching

Hansen-Nygård, G., Nielsen, K.L., Thorseth, T.M., and Stav, J.B.

# Chapter 6
# Developing and Evaluating Practical Methodological Guidelines for use of Student Response System in Teaching

**Gabrielle Hansen-Nygård**
*Sør-Trøndelag University College, Norway*

**Kjetil L. Nielsen**
*Sør-Trøndelag University College, Norway*

**Trond Morten Thorseth**
*Sør-Trøndelag University College, Norway*

**John Birger Stav**
*Sør-Trøndelag University College, Norway*

## ABSTRACT

*This article presents methodological experiences and evaluation results obtained during introduction and testing of a new online student response system (SRS) for modern mobile devices at Sør-Trøndelag University College, in Norway. The aim of the test period was methodological development, based on student evaluation. Using in-depth interviews with students, awareness of how SRS was comprehended by the students in their learning process increased. Several methodological choices and practical challenges were faced when introducing SRS. The procedures and methodological choices were based on published experience and the authors' assumptions. However, what was believed to be important pedagogical, were among the students perceived as positive but not in the way expected. The students have a clear perspective on their own learning process and gave insight into how SRS fit into their own learning process. Students' perceptions regarding methodology, in combination with their own experience of learning, appear as a necessary ingredient for an appropriate implementation and use of SRS in teaching.*

DOI: 10.4018/978-1-4666-0936-5.ch006

## INTRODUCTION

Key challenges associated with traditional teaching in higher education include low level of involvement among students, difficulties with implementation of various feedback activities and encouragement of active learning strategies such as in group discussions and larger class discussions. In an attempt to find possible solutions to such challenges several researchers and teachers have focused their attention to different technological tools. One of the tools that have received increasing attention over the past few years is student response systems (SRS).

SRS can shortly be described as a type of wireless technology aimed at promoting better communication, response and interactivity in large classrooms (Beaty, 2004). It is a technology that allows teachers to present a question or problem to a class and let students respond by using response devices. Responses are quickly summarized and aggregated for the teacher and students to see (Beaty, 2004). Based on the response data, both students and the teacher can get an idea whether key concepts are understood or misunderstood.

Research has identified several important advantages from use of SRS in teaching (Beaty, 2004). With the use of SRS students receive a clear confirmation on whether or not they actually have learned something and a clarification in relation to what they may have misunderstood (Rice & Bunz, 2006). Several studies show that student engagement increases significantly when SRS is implemented as part of their teaching (Horowitz, 1988; Dufrense et al., 1996; Gilbert et al., 1998; Everett & Ranker, 2002; Draper & Brown, 2004; Roschelle et al., 2004; Stuart et al., 2004; Masikunas et al., 2007). SRS also stands as a successful approach for managing discussions in large classes (Dufrense et al., 1996; Mazur, 1997; Draper & Brown, 2004; Masikunas et al,. 2007). Research shows that SRS can help create active discussion among students, which further can promote more active learning in the classroom (Boyle & Nicol, 2003).

Use of technology in education is by no means a new phenomenon (Rice & Bunz, 2006). History is full of attempts *and* subsequent failures in relation to introduction and implementation of various technological innovations designed to improve ordinary teaching and students' learning (Cuban, 1986). In light of this rather gloomy trend, it is perhaps appropriate to ask: what about the use of SRS? Or as Duncan (2006) asks: "are SRS just another educational fad?"(Duncan, 2006, p. 16). In other words, can use of SRS be classified as an instructional trend that will fade away as soon as the excitement has subsided? Considering that the majority of research shows several positive effects from use of SRS in classrooms, one cannot help but wonder what determines this technology "survival" from the fate of its failed predecessors, which leads to the following question: what separates a successful implementation of a response system from a less successful implementation?

Research shows that implementation and use of SRS, as a part of teaching, can act in many different ways. The system can be a supplement to teaching, or be a main pedagogical tool (Trees & Jackson, 2007). According to Beaty (2004), it is all about using the suitable pedagogical method. One example of a pedagogical choice regarding use of SRS is the questions (Duncan, 2006). The choice of questions in relation to the subject has to be considered, since the effect may vary with the subject (Stuart et al., 2004). Beaty (2004) recommends teachers to avoid simple questions that are based on memorizing, and rather use questions that call for careful appreciation and evaluation. Another pedagogical choice is the number of questions asked. It is recommended to create questions of good quality, rather than having a large quantity (Beaty, 2004; Duncan, 2006). According to Beaty (2004), SRS should not be used to fill the lecture with questions, but rather use it carefully with less questions of good quality.

One way to create student engagement is for the teacher to avoid reading the questions aloud, and hence avoid a teacher centric instruction (Beaty, 2004). Teachers often include group discussions as a part of the pedagogical method (Horowitz, 1988; Dufrense et al., 1996; Mazur, 1997; Gilbert et al., 1998; Crouch & Mazur, 2001; Boyle & Nicol, 2003; Nicol & Boyle, 2003; Beaty, 2004; Draper & Brown, 2004; Stuart et al., 2004; Kam & Sommer, 2005; Masikunas et al., 2007; Trees & Jackson, 2007). The discussion can be made in small groups after a question is given, but before the results are shown, either individually or in groups (Gilbert et al., 1998; Draper & Brown, 2004; Trees & Jackson, 2007). Alternatively, the responses are given individually first and then the collective response is discussed in groups. Another approach is the "peer instruction" (Mazur, 1997; Crouch & Mazur, 2001) that allow students to first respond individually to a question, then discuss with a peer student, before the response is given again.

A successful implementation of SRS is not only a matter of the teacher and pedagogical choice, but also very much the students (Dufrense et al., 1996; Duncan, 2006; Trees & Jackson, 2007). For some students, the transformation from a passive to an active participating student might cause a challenge (Dufrense et al., 1996). Given the potential that SRS has in the classroom, the possibility to move from passive towards active students, the way to success is dependent upon the students susceptibility, whether the students "by it" (Trees & Jackson, 2007).

A recently completed project, EduMECCA (2008-2010), an EU-funded KA3-ICT Lifelong Learning Project that consisted of 7 partners from different European countries, attempted to address the issues above. In this project an open, web-based response system for PC, MAC, iPod Touch, iPhone or any mobile devices that can read HTML, has been developed and tested. A description of the system and its functionality is given by (Nielsen et al., 2011). A significant part of the project has been to focus on how to implement and use the web based system in vocational training, as an integrated part of teaching and to develop a methodological guideline.

This article presents the first methodological experiences and evaluation results obtained during the fall of 2009 in physics and engineering education, when testing out a new type of SRS for next generation mobile handheld devices at Sør-Trøndelag University College, in Norway. During a testing period of five weeks, in a preparatory class for engineering with about 50 students, we gained initial experiences in relation to usage and method development. The overall goal for this trial was method development based on student evaluation. The student evaluation was carried out at the end of the testing period, and consisted of two focus group interviews.

## METHODOLOGICAL IMPLEMENTATION

In October of 2009 the system was, from a technical point of view, ready for the first test in class. The practical trial of the SRS began in Trondheim early that month. The trial was conducted in a preparatory class for engineering with 50 students, in the subject of physics, and lasted over a period of 4 weeks with7-hour teaching peer week. During this period, we developed initial experiences regarding the usage and methodological guideline and evaluated through interviews with students.

Following is a description of six methodological procedures we used during our implementation of SRS.

### 1. Introducing of SRS to the Students

The first time students were introduced to SRS, we arranged an approximately 20 minute introduction to the system, in which the students learned what the SRS was, why it was introduced and how we expected that the lessons with SRS would run during the course. The goal was to make students familiar with the system and explain why it was adopted. There was also a practical review demonstrating how an iPod works, how it is turned on, how enter SRS, how to vote, and how do they know that they have actually voted.

After the introduction, the students were placed in groups of three. Each group was given an iPod and a couple of test votes were given, so that students would get a picture of how the system worked.

## 2. Start Up

Normally, students picked up an iPod when they arrived in the classroom. The iPods were placed in front of the classroom, so that the teacher had an overview over them. When the teacher started the lecture, he/she presented the session code at the beginning. The students logged on with the given code and received confirmation that the iPod was linked to the session when the message "please wait" appeared. We decided to present the code at the start of the class period in order to reduce time-loss later when the focus zeroes in on the quiz questions. The procedure took approximately one minute and gave both teacher and students a confirmation that everything was functioning as it should before teaching started. After this procedure the teacher started teaching as usual.

## 3. The Quiz Questions: When and How to Present the Questions?

After approximately 20-30 minutes the first quiz was presented to the class. It can be difficult for students to pay attention throughout a class period of 45 minutes (Horowitz 1998), and that is why the first quiz question was posed at a time when the students can begin to lose their concentration. The teacher presented the quiz question by reading both the question with alternatives *aloud* to the students. If the question seemed difficult, time was taken to explain the question. The reason why both the question and its alternatives were read aloud, was to ensure that as many students as possible understood the question. Some students may have reading and writing difficulties and will have difficulty in grasping the question if they were to read it silently.

## 4. Small-Group Discussions

After the question was presented, the teacher encouraged the students to discuss the quiz question and its alternatives in small groups for a couple of minutes. Time was not spent on placing the students in groups; they discussed with the person/s beside them. The aim of the group discussion was to give the students a chance to be more involved in the actual teaching, as well as enable them to learn from one another by hearing other students' opinions and arguments.

## 5. Polling with a Timer and a Clock

Before the voting proper began the teacher must decide whether the voting was to be with a timer and a clock sound. We ran all polling's with a timer and clock sound. The timer was set to 30 seconds, but could be adjusted if necessary. The iPods can use up to five seconds to get the alternatives, which means that the students had 25 seconds to cast their vote, which we considered sufficient.

Our experience has been that both the timer and the sound of the clock are essential in order to create order and attention in the classroom during the polling. The sound reminds students that now is the time to vote and they have a limited time to respond. After the group discussions the students may be very involved and exited. The sound and time-pressure exerted by the timer, help the students to become quiet and get them to focus on the board and the coming results.

### 6. Teachers' Explanations Afterwards

When the polling closes the results appeared automatically. The teacher then walked through the results, highlighted the correct alternative and explained thoroughly why the various alternatives were correct or not.

After this, teaching proceeded as usual. The whole procedure – from the time when the question was posed, the students discussion, the voting itself and the teacher's explanation afterwards – usually took no more than five minutes.

It is important to emphasize the fact that our starting point regarding the use of SRS has been ordinary teaching. During a teaching period that lasts 45 minutes we have never presented more than a maximum of two quiz questions. The use of SRS has not, for us, been about filling a class period with quiz questions. Rather it has been about using the technology to give students a much desired break in which they get the opportunity to think for themselves, use their knowledge, and discuss and take part in a vote which again gives them concrete feedback on their own learning. Our goal has never been to turn the way teachers teach upside-down, but to try to make ordinary teaching more interactive.

## RESEARCH METHOD

After the trial, two student interviews were conducted. The interviews were performed as *focus group interviews*. The main reason for the choice of interview method is that focus interviews are recognized as an independent method of revealing the informants 'own perspectives about various topics (Johannesen, Tufte, & Kristoffersen, 2004).

A focus group interview constitutes a form a group interview where the conservation and discussion process is essential. One of the main advantages of focus group interviews is that, if properly managed, it can be extremely dynamic (Bergh, 2007). Interaction between informants can stimulate discussion in which informants respond to each other`s comments. Such a group dynamic if often described as a synergistic group effect, an effect that allows informants to build on what others have said or to enter into a collective "brainstorming" process. The idea is that the whole contribution to the process is more than the sum of its parts.

The interviews were analyzed using a type of analysis called *thematic analysis*, a wildly used approach to qualitative analysis (Boyatzis, 1998). Thematic analysis is a method for identifying, analyzing and reporting patterns (themes) within data (Braun & Clark, 2006). It organizes and describes your data set in detail, and interprets various aspects of the research topic (Boyatzis, 1998). The aim of the current evaluation was to bring out the experiences, views and opinions expressed by our students in relation to the use of SRS in physics' classes. Thematic analysis was considered an appropriate tool for achieving that goal.

## RESULTS

Overall students experienced the SRS as an integrated part of their classes – it was not seen as some additional element which disrupted normal classes. They saw great benefits of the system in relation to receiving feedback on their learning progress; experiencing increased involvement and engaging in academic discussions. The system became a natural part of physics teaching. According to students, this can largely be explained by the fact that the system did *not take time away* from regular classes. Regular routines of how the system would be used were implemented, and the students quickly got into these routines, which streamlined SRS usage. Student quotes;

*"The student response system doesn't interfere with normal classes, but rather integrates nicely."*

*"It was very well integrated into the classes. Not something that took a long time and dragged out time spent in classes. It simply worked – a subject was taught; then we did a quiz, and it just went smoothly. It was very good."*

Regarding the six methodological procedures, which were followed during the test period, students had following opinions and viewpoints:

### 1. "I Don't Think It Would Have Worked Without the Introduction, Really"

The first time students were introduced to the SRS, there was an approx. 20-minute introduction where students got an introduction to *what* the SRS is, *why* it was introduced in physics teaching, and *how the* lessons with SRS were to run in their course. Then there followed a hands-on demonstration of the system. Students were asked a few non-academic quiz questions and responded with the use of the iPods. This was done to give them training in how use of the SRS, in particular how the iPods are operated. From the student side, this was an effective way to be introduced to the system, and they assessed the knowledge gained in the introduction as valuable. By getting to know what would happen and how SRS would be used in classes, they built some expectations that they felt gave them a more considered opinion about the pros and cons of the SRS. Their first impression was that it sounded interesting, and they looked forward to using it. No one had changed their minds after the system was used in classes. Student quotes:

*"I think it was very nice to get to know what's going to happen. The presentation was nice. You understood a little bit more of what was going on."*

*"I don't think it would have worked without the introduction, really."*

*"I was positive from the start, and I'm positive now. I see that this is a beginning of a system that can only get better. So I think it's good. For me it has something to say for motivation in teaching. I think it's fun. It is bit more exciting than normal teaching."*

### 2. Handing Out and Returning iPods: Where to Place the iPods?

During the introduction, the teacher handed out iPods to students. For the remaining classes in which SRS was used, the iPods were placed in two suitcases by the blackboard. Students were told to pick up a unit from the suitcase on their arrival into the room. According to students, this scheme worked fine, their only comments being about the placement of the suitcases. For them it would have been better if the suitcases were placed by the entrance to the classroom, rather than the blackboard. If they were late for class, they found it embarrassing to go up to the board to collect an iPod while the teacher was teaching. Additionally, they were nervous about disrupting classes. The upshot was students arriving late would only pick up an iPod when the vote was about to start. This causes some stress since they had to quickly enter the session code, and if they typed the wrong code or took too long, they couldn't participate in the vote. Students` recommendation was to place the iPods by the door where they enter the room; it would be much simpler.

The students reported few problems with the *iPods* themselves. The iPods were found to be simple to use, it was intuitive and easy to understand the different aspects of the SRS elements which appeared on their iPods; session code, "please wait" and "thanks for your vote". The page that required the most work from the students was the page with the session code. For the students, this quickly became a routine task at the beginning of classes. Student quote:

*"It became a routine, because once the SMART Board was up and running, the code was shown on the screen, so yes, I think it worked very well."*

## 3. Efficiency through Reading Out Loud

Another routine students highlighted was that the teacher read the quiz questions and its alternatives *out loud*. That the teacher took the time to do this was valued as extremely important by the students. Through this procedure the teacher clarified and explained the question to the students, and thus prevented or cleared up any misunderstandings. According to students, it also helped to streamline the usage of SRS. If the students were to read through the questions themselves, they believe it would take longer than if the teacher read aloud – especially considering that students have different reading speeds. By having the teacher reading the question and alternatives, the students felt that they gained extra time to consider the question, and they could start making up their minds about the subject even as the teacher was reading. Some students also suffer from reading and writing difficulties and could not participate in classes with SRS unless the teacher read the question out loud. Student quotes;

*"I have a poor understanding of sentences and therefore need long time to read. I often have to read the same thing several times. So for me it was very nice that the teacher read it out loud."*

*"I think it was a great way to have it spoon-fed. We need that"*

*"It's very good and important that the teacher does it. It captures our attention, in a way. I don't think I would have bothered to read it if the teacher had simply said, "Read this yourselves." We listen to the teacher when he reads it, we want to hear what he says. I think it would have taken longer if we had to read it ourselves."*

## 4. Student Discussion: A Good Way to Learn

Before each vote, the students were encouraged to discuss the quiz questions and its alternatives among themselves for a couple of minutes. An opportunity the students valued. Working with other students are for them an effective way to learn. Hearing the perspectives, opinions and viewpoints of other students, are highlighted as important to achieve a better understanding of the subject matter. The students described the group discussions in following way:

### Group 1

**Per:** "I think I learned something from them, absolutely ..."

**Ole:** "It's always nice to get the opinion of the person sitting next to you"

**Jens:** "Yes, when you see the questions, you form an opinion that goes one way, and then along comes the person next to you with a different opinion. Thus, you get input from somebody who may think in a completely different way, and you just realize, "I never thought about that". Yes, you get a chance to discuss what the correct option is."

### Group 2

**Lise:** "It's very nice to be given the opportunity to speak with someone, especially since we're covering subject areas that are new to us. It's good to hear what others think, and together try to achieve a common understanding."

**Emma:** "Yeah I think it worked really well. We tried to reach an agreement on the correct answer. So, if we disagreed there would be a very good discussion. You knew that both sides couldn't possibly be right, so you'd turn the material a bit upside down and discuss it. Very good."

The discussion among peers was thus perceived as a valuable procedure in relation to the use of SRS. One reason for this is that the discussion had a clear *goal,* as it would end in a vote which gave them an immediate feedback on their learning. They didn't discuss for the sake of the discussion per se; rather they discussed to be better prepared to answer the quiz question. The goal was to find the correct answer that would further give them a positive feedback, and the discussion could increase their chances of achieving that goal. The feedback students would receive when using the SRS was thus a "bonus" that stimulated them to participate actively in the discussion. According to the students, this made the discussions focused and efficient. They only had a couple of minutes to discuss, and therefore had to work efficiently. One student group had the following comments about this;

**Emma:** "The point of the discussions that we had, was to figure out an exact answer. Otherwise, when we are discussing, I think the discussion very quickly loses focus, or at least becomes a rather "free-roaming" discussion."
**Lise:** "Yeah true, I think the voting is very important! I don't think we would have bothered to discuss with the person sitting next to us if it would have been for nothing; if I didn't cast a vote afterwards"
**Emma:** "You motivation increases."
**Ingrid:** "You put more into the discussion, to find the right answer."
**Lise:** "Yeah, you can really see the benefit of it!"

## 5. The Ticking Clock

The next routine that was discussed by the students was the *ticking clock sound* which was played back during voting. When the vote started, a timer was shown on the digital whiteboard which counted down from 30 seconds. Additionally, a ticking clock sound would be played back though the speakers. The students saw this as a positive aspect of the system, especially the ticking clock sound. For them, it served as a reminder that they had a limited amount of time to cast their vote. After a few minutes of discussion in preparation for a vote, the noise level in the class could be quite excessive, and their focus was not directed towards the blackboard, but against each other. A countdown timer without a sound would therefore have been of little use, and the students themselves reported that they could easily forget to vote unless the ticking sound was there. For them, it served as a signal that the discussions had now ended, and their attention should be directed towards the board. In addition, the combined countdown and ticking sound added time pressure, created an element of excitement, in terms of anticipation for the results of the vote. Student quotes;

*"When the sound comes on, you know that you have 30 seconds to cast a vote. Do not remove the ticking sound!"*

*"I think it worked really well."*

*"It became a bit of a quiz show atmosphere; we were all waiting with anticipation, hehe"*

## 6. Learning through Feedback

According to students, the use of SRS gave them valuable feedback on their learning and progression. To answer ta quiz questions and receive an immediate feedback, was for them a way to test themselves. By getting feedback on whether they had understood what the teacher had tried to convey, they got to test their knowledge in practice. Two of the students had the following to say about the feedback that the SRS gave them;

**Per:** "You get feedback on how well you have understood the topic. If you got the quiz question right, you received feedback that you'd actually understood the subject. You get feedback that you're able to use the right formulas and laws, and, yes, the material that the teacher has presented."

**Ole:** "Yes, you get feedback on whether you have understood it. It's about your own learning process, really. You get to see if you've learned something."

For the students, feedback is an important part of their learning. Feedback gives them an indication of their own learning progress. Normal feedback activities for these students include written tests and assignments. Feedback activities are normally not included as an in-class activity. The only opportunity students have to receive feedback during lectures is by raising their hand and either ask or answer questions from the teacher, a procedure they rarely do, since most of they find it very uncomfortable to raise their hand and talk aloud in front of a dozen other students. When asked whether feedback activities are included in normal classes, one of the students responded;

*"No, the teachers may ask, "do you understand?" and then they just look sheepishly at us and move on. None of us that dare to raise our hand and respond. In that sense, it's our own responsibility, but no, I certainly don't. Feedback activities are normally not included, which is a bit of a shame."*

Students want something they call *constructive* feedback in their academic life. This is feedback which, in addition to indicating if they're on the right track or not, explain *why* something is right or possibly wrong. From the student side, this feedback point out what they need to work on, as it highlights areas where they're struggling and need to focus on. Without such feedback, the students feel very much left to themselves, which makes it difficult for them to get an accurate assessment of their own learning and progression. One of the students said it quite clearly;

*"Without constructive feedback, how can we hope to improve?"*

In light of the students' desire for more constructive feedback, the SRS came as a long-awaited and most welcome addition to classes. Firstly, the system gave them an immediate feedback on their vote, in that they got to see if they had voted right or wrong. Furthermore, the teacher would go through each alternative after the vote and thoroughly explain *why* it was correct or incorrect. For the students, the teacher's explanation was perceived as a constructive feedback, and was highlighted as critically important for their own experience of learning. Through the teacher's explanation, the students got an understanding of why the various options were correct or incorrect. One thing is to cast a vote that turns out to be right or wrong; another matter entirely is to be able to understand *why* it is right or wrong. If they achieve such an understanding, they feel that they really learn something through the quiz questions. The students are keen to point out that the teacher should give adequate explanations for why the wrong options are incorrect. For the students, this is a way for those who answered incorrectly to understand *why* they got it wrong. One of the student groups explained it this way;

**Emma:** "Those of us who got the answer wrong have to be given to chance to understand that was wrong. Some part of the class usually got it wrong, and then it must be explained in such a way that we can understand where we went wrong. Because we obviously don't know if an option is wrong – otherwise we wouldn't have voted for it!"

**Lise:** "Yes, I feel it gives me a chance to understand what the subject is really about."

The second group had the following to say;

**Per:** "There's a reason why people have answered incorrectly, it's because they have misunderstood something, and then they have to be explained why the answer was wrong."

**Ole:** "Yeah I think the explanation from the teacher is very important. I think it is necessary that he explains why he uses certain laws, or other parts of the curriculum, and that he shows us why it is right or wrong."

**Jens:** "Spending some time to explain or discuss the different options in this way is; well, I feel that the quiz becomes a bit useless if you don't do that - if you don't spend enough time on it and do it thoroughly. The quiz then becomes – maybe not useless, but the quiz has a much greater effect on learning if you get an explanation why the answers are right or wrong."

## Room for Improvement; How to Get Feedback on "Actual Understanding"

From the student side, there is little doubt that the use of SRS can provide them with valuable feedback on their learning, particularly if the teacher gives them a thorough explanation after the vote. At the same time, however, they leave no doubt that the SRS may have a much *greater* potential than was used in their teaching.

Physics classes for these students are made up of two or three successive lessons. The teacher would begin by introducing new topics from the subject curriculum, and then give the students a quiz questions based on what had recently been presented, after approximately 20 minutes. According to students, this was a straightforward way to implement SRS in teaching, as they felt that the teacher's explanation after the vote contributed to their learning. As far as measuring their *understanding*, however, it was not optimal. Quiz questions simply came too early. Whether you test understanding or not depends, according to the students, on the time when the quiz question is asked – in particular, whether they've actually had time to learn something before the question is asked. In other words, if the teacher wants to use the SRS to measure students' understanding, the students must first be given time to work with the subject matter and acquire the academic skills needed to answer the quiz question. If the quiz question is presented too early in the session, the students may not have had time to acquire these prerequisites. In other words, instead of *measuring* their understanding or ability to apply knowledge to solve a problem, by using the SRS, the teacher *gives* them an understanding by giving a thorough explanation after the vote. Student quotes:

### Group 1

**Per:** "I would like to get a quiz at the end of the day too, in order to check if we've really understood it. After we've worked with the exercises for a period of time, and had time to process the material."

**Jens:** "Yes, I agree."

**Per:** "That would give a very good indication as to whether you've understood something or not. That would be a proper test!"

**Ole:** "Then we would have worked with it for a bit, and then we'll get to see if we've understood it."

### Group 2

**Emma:** "I somehow ... need time to understand it, in a way. Sometimes I think that the quiz questions seemed to come too early for me, in a way ... There were times when I just made a guess. I had somehow not received the scientific basis for properly discussing it. I felt that it was a bit unnecessary."

**Ingrid:** "Yes, we'd almost have to lie ahead, if we are to do it that way. The questions tend to be from the new subject area that we've just been through. So really, it might be best if he took us through the curriculum first, and included questions at the end of the class."

**Lise:** "Yes, to see that people had actually understood it."

From the students' side, this is really about what the teacher wants to use the SRS for, i.e., does the teacher want to *give* the student an understanding, by giving them a quiz question followed by a thorough explanation, or does the teacher want to *measure* their understanding. The students did not say that we'd selected a wrong way to use the SRS, they merely point out that to really measure their understanding, they must first be given a chance to an understand the subject matter, which is rarely the case after only 20 minutes. Proper understanding usually comes after the material has had time to mature – by doing exercises or assignments related to the subject matter, for instance.

## DISCUSSION

The interviews tell us that the students have a clear opinion on how to use the SRS system. The students explain profoundly the importance and possible outcomes from different pedagogical choices, and how this influences their learning. For us, the students' experiences and points of view increased our own understanding of consciousness in choice of pedagogical approach. Some choices were of more importance than we assumed, while others were not. We could not predict that the location in the classroom where we gave the iPods to the students could be an issue for the students. Some of our pedagogical choices were made according to published recommendations and other was not. Through the interview, we got an explanation from the users of SRS, the students, why different pedagogical approaches were experienced as important or not.

One of the methodical choices that received much attention from the students in relation to use of SRS, were the questions. Here there are many choices to be made; what type of question to ask and to what subject. For some subjects a simple "Yes" "No" type is suitable, while for others the question is suitable as a "brain teaser", a question meant to prompt discussion among students (Stuart et al., 2004). Another important aspect is how many questions to ask? In literature the recommendations are clear, don't ask too many questions (Beaty 2004), but focus on the quality of the question (Duncan, 2006). It is also recommended that teachers don't read the questions out loud to the students (Beaty, 2004). Letting students read it themselves, or discussing its meaning with their peers, is thought to prevent the classroom from being centered on teacher (Beaty, 2004). During our trial of the SRS, we chose to do the opposite; the teacher read the questions and the alternatives aloud to the students. Our initial reason for doing this was to clarify the questions and ensure that everybody got it, including those that might have reading or/and writing difficulties. The students on the other hand, perceived this as a routine that streamlined the use of SRS. When the teacher read the question, they read along with him/her and were thus ready to enter discussion immediately after the question was presented. Whether this procedure made the teaching more centered on the teacher, as suggested by Beaty (2004), we are not sure of. We received no signals from the students that they felt less involved due to this procedure, on the contrary, they pointed out that it provided all students an opportunity to participate. Some of the students had reading and writing difficulties, and felt that they would not have been able to participate without this teacher led procedure. Of course, in order to maintain student-centered activity, you can leave it to the students as a group to organize a way to make sure that everyone understands the question, but this requires that everyone feels comfortable exposing their own difficulties.

Teachers also have to consider when in the lecture the question is to be asked. Horowitz (1988) recommends asking questions approximately every 15 to 20 minutes, since the students attention normally drops significantly after this timespan. Our approach was to follow this recommendation and hence, we gave a question normally 20 minutes into the lecture. For every hour, lasting for 45 minutes, we normally prepared one or maybe two questions. Our focus was students' attention. What we did not consider, was the consequences this had for the students when they considered their own learning process.

When the quiz question should be introduced, was for the students a question about what the teacher wants to obtain with the question. The teacher may want to target measuring knowledge, and give the students a chance test themselves on the subject, or use the question as a tool to create an understanding of the subject. The students have a clear picture of this. If the question intend to measure knowledge, they first have to be given time to acquire an understanding of the subject. If the question is presented to early in the introduction of a new subject, the students' preconditions to answer is not right. They don't test their own understanding, because they haven't had time to build their own understanding of the subject. They guess the answer. The question did however help build an understanding, mainly due to the teachers procedure of going through the alternatives; which were correct, which were wrong, and why. If the students are given time to digest the subject, the scenario change. Under such conditions, SRS can be used to measure acquired knowledge. The pedagogical approach where we had this procedure of carefully summarizing the question and the alternatives, did not work as intended. While we thought that we gave the students a possibility to get feedback on their own understanding and help clearing up misunderstandings, we helped students build an understanding of the subject. For the students this was crucial, this was the point where they experienced learning.

If the teacher just points out the correct alternative and then continues with the lecture, the experience of learning is minimal. It is not through the immediate feedback from the SRS system that the students learn. If the teacher however, carefully sum up all alternatives and gives a detailed explanation for each alternative, they significantly improve their learning experience. For them, the confirmation that an alternative is right or wrong it is not important, but when they understand why one alternative is correct or wrong, this gives them an experience of a deeper understanding of the subject. The students experience is according to Kulhavy and Stock (1989), who distinguish between verifying versus elaborating feedback. Verifying feedback according to Kulhavy and Stock is when the wrong and correct alternatives are given. Elaborating feedback is when the reason for why the alternative is correct or wrong is explained. Kulhavy and Stock says that both kinds of feedback have to be present, I order for the students to have a positive learning process.

One pedagogical choice that has had a lot of focus in the SRS literature is implementation of small group discussions (Dufrense et al., 1996; Gilbert et al., 1998; Boyle & Nicol, 2003; Nicol & Boyle, 2003; Beaty, 2004; Draper & Brown, 2004; Masikunas et al., 2007; Trees & Jackson, 2007). Creating peer discussions is a challenge that teachers normally face in classrooms holding a large group of students (Kam & Sommer, 2005). The teacher can use SRS to create peer discussions, especially in large groups of students where gossip and social events often dominate (Masikunas et al., 2007). Creating a discussion can be time consuming and steal time from the teaching. The introduction of SRS does not guarantee active participation in group discussions, but can act as a tool to help the process (Dufrense et al., 1996; Masikunas et al., 2007; Trees & Jackson, 2007).

So what do the students say about this? According to our students, “the discussion had a final goal.” The students were given a question, and used the peer discussion as a tool to help decide on the correct alternative. This gave the discussions an extended context. The students could clearly see why the discussion was there, and what they could gain from it. In addition, the discussion had a purpose and they felt rewarded by the elaborating feedback at the end. The peer comments and arguments helped them gain understanding, and explaining to others their own arguments was recognized as a valuable part of the training. At the same time, the peer groups consisted of students who were trying to learn something new, a fact which limited the confidence they put to in the outcome. In regards to students’ experience of learning, the final elaborating feedback from the teacher was the one found most trustworthy. But the peer discussions helped them prepare and become more receptive to them.

## CONCLUSION

This article has presented and discussed students’ experiences of methodological choices and procedures in relation to the use of an online SRS for modern mobile devices. The system was well received, even during the development phase of the system. On the whole, students’ experiences can be characterized as positive and constructive.

Through interviews with students, we developed a deeper pedagogical understanding of the importance of different methodological choices. An understanding we will take with us in our further development and use of SRS. It became very clear that seemingly minor methodological changes can have big impact on students’ impressions and learning experiences. All choices teachers make can affect their overall impression, and determine whether students “buy the concept” of SRS. An increased awareness of the importance of practical choices combined with a close contact and communication with students, may increase the chance for teachers to succeed with use of SRS. In relation to the students who were interviewed in this study, we feel that it is correct to say that the key to their positive experiences was that they “bought it”, in the sense that they felt that it gave them something in return; a valuable learning experience, based on different methodological choices.

## ACKNOWLEDGMENT

This project has been funded with support from the European Commission. New Educational Models that Encourage Creative transfer of Competence and Acquaintance in Lifelong Learning (1.1.2009 – 31.12.2010) Contract 143545-2008-LLP-NO-KA3-KA3MP. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

## REFERENCES

Beaty, I. (2004). Transforming student learning with classroom communication system. *Research Bulletin (Sun Chiwawitthaya Thang Thale Phuket)*, *3*, 2–13.

Berg, B. L. (2007). Focus group interviewing. In Berg, B. L. (Ed.), *Qualitative research methods for the social sciences* (pp. 144–170). Boston, MA: Pearson Education.

Boyatzis, R. E. (1998). *Transforming qualitative information: thematic analysis and code development*. Thousand Oaks, CA: Sage.

Boyle, J. T., & Nicol, D. J. (2003). Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal*, *11*(3), 43–57. doi:10.1080/0968776030110305

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*, 77–101. doi:10.1191/1478088706qp063oa

Clark, R. E. (Ed.). (2001). *Learning from media: Arguments, analysis, and evidence*. Greenwich, CT: Information Age.

Crouch, A. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970–977. doi:10.1119/1.1374249

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York, NY: Teachers College Press.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, *20*, 81–94. doi:10.1111/j.1365-2729.2004.00074.x

Dufrense, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, *7*, 3–47. doi:10.1007/BF02948592

Duncan, D. (2006). Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review*, *5*, 70–88. doi:10.3847/AER2006005

Everett, M. D., & Ranker, R. A. (2002). *Classroom response system: An evaluation at an easy-access regional university.* Retrieved June 4, 2008, from http://www.einstruction.com/News/index.cfm?fuseaction=News.display&Meny=newsroom&content=FormalPaper&id=125

Gilbert, M., Massen, C., Poulis, J., & Robens, E. (1998). Physics lecturing with audience paces system. *American Journal of Physics*, *66*(1), 439–441.

Horowitz, H. M. (1988). *Student response system: Interactivity in a classroom environment*. White Plains, NY: IBM Corporate Education Center.

Johannessen, A., Tufte, P. A., & Kristoffersen, L. (2004). *Introduction to social science methods*. Oslo, Norway: Abstrakt Forlag.

Kam, C. D., & Sommer, B. (2006). Real-time polling technology in a public opinion course. *Political Science and Politics*, *39*(1), 113–117.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279–308. doi:10.1007/BF01320096

Masikunas, G., Panayiotidis, A., & Bruke, L. (2007). The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, *15*(1), 3–20. doi:10.3402/rlt.v15i1.10903

Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mazur, E. (1997). *Peer instruction: A user`s manual*. Upper Saddle River, NJ: Prentice Hall.

Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, *28*(4), 458–473. doi:10.1080/0307507032000122297

Nielsen, K. L., Stav, J. B., Hansen-Nygård, G., & Thorseth, T. M. (2011). *Designing and developing a student response system for mobile internet devices*. Paper presented at the International Conference on Education & Learning in Mobile Age, Lake District, UK.

Rice, R. E., & Bunz, U. (2006). Evaluating a wireless course feedback system: The role of demographics, expertise, fluency, competence, and usage. *Studies in Media & Information Literacy Education*, *6*(3), 1–32. doi:10.3138/sim.6.3.002

Roschelle, J., Penuel, W. R., & Abrahamson, L. (2004). The networked classroom: Electronic classroom networks can enhance student participation and achievement in mathematics and science. *Improving Achievement in Math and Science*, *61*(5), 50–54.

Stuart, S. A. J., Brown, M. I., & Draper, S. W. (2004). Using an electronic voting system in logic lectures: One practitioner`s application. *Journal of Computer Assisted Learning*, *20*, 95–102. doi:10.1111/j.1365-2729.2004.00075.x

Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, *32*(1), 21–40. doi:10.1080/17439880601141179

# Paper C

## Investigating Peer Instruction: How the Initial Voting Session Affects Students' Experiences of Group Discussion

Nielsen, K.L., G. Hansen-Nygård and J.B. Stav.
*ISRN Education*. Article ID: 290157. 2012.

International Scholarly Research Network
ISRN Education
Volume 2012, Article ID 290157, 8 pages
doi:10.5402/2012/290157

*Research Article*

# Investigating Peer Instruction: How the Initial Voting Session Affects Students' Experiences of Group Discussion

**Kjetil L. Nielsen, Gabrielle Hansen-Nygård, and John B. Stav**

*Sør-Trøndelag University College, 7004 Trondheim, Norway*

Correspondence should be addressed to Kjetil L. Nielsen, kjetil.l.nielsen@hist.no

Received 16 March 2012; Accepted 3 April 2012

Academic Editors: F. Jimenez, A. Kara, K. Kiewra, and K. Y. Kuo



Peer Instruction is a popular method of implementation when using Student Response Systems (SRS) in classroom teaching. The students engage in peer discussion to solve conceptual multiple choice problems. Before discussion, students are given time to think and give individual responses with a voting device. In this paper, we investigate how this initial voting session affects students' experiences of the following discussion. The data is based on student interviews which were analyzed using analytical tools from grounded theory. The students emphasize the individual thinking period as crucial for constructing explanations, argumentation, and participation during discussions, and hence for facilitating learning. However, displaying the results from the initial vote can be devastating for the quality of the discussions, especially when there is a clear majority for a specific alternative. These findings are discussed in light of recent quantitative studies on Peer Instruction.

## 1. Introduction

The traditional one-way teacher style of lecturing can be effective when delivering factual content, but it is not as effective for facilitating cognitive skills [1, 2]. Engaging students in active-learning activities can be an important factor for mastering skills such as critical thinking and problem solving [1, 3], skills that are often lacking in novice science students [4]. Although such skills are vital to evaluate scientific evidence and theories, students also have to be able to generate and present explanations and arguments for their evaluations, that is, to be fluent in the scientific language [5]. Novice students also suffer in that they are often unable to put into words, or at least scientific words, how they approach the use of theories to solve problems. Engaging students in peer discussions can challenge them to generate explanations and convincing arguments for their solution and in this way also facilitate deeper understanding of the scientific phenomena [5].

One way of engaging students in active-learning activities is to use a Student Response System (SRS). Such systems are often used during the lecture to present students with multiple choice questions, which they will discuss with their peers in small groups before answering with a voting device [6–10]. The result from the voting session is displayed in the form of a histogram which can give the teacher an indication of the level of understanding among the students and if they are able to follow the lecture [10, 11]. SRS use in classroom teaching has been shown to increase learning [12–15] including increased conceptual understanding in physics courses [12, 16, 17].

Although SRS can be an outstanding tool for facilitating peer discussion, different choices of implementation can have a high impact on the quality of the discussions. For instance, giving credits for SRS participation has been shown to increase the overall participation in the class [18]. However, in a study by James [19], the researcher found that if each individual student was "punished" for voting incorrectly, that is, that he/she was not given as much credit as voting correctly, the group discussions tended to be dominated by students with greater knowledge while other students remained passive. If the students were only credited for participation, they would be more inclined to participate and explore different explanations and ideas.

One popular methodological implementation of SRS is with the Peer Instruction technique [17]. During the lecture students are presented with a problem, often where conceptual ideas are in focus, to challenge their knowledge about the subject as opposed to simply looking up the answer in the textbook. Students are given time to think on their own for about 2-3 minutes and answer individually before they are encouraged to engage in discussion with nearby students. The discussions are concluded with a revote and an explanation by the teacher. Dufresne et al. [11] describe a similar method, but where the initial voting session is omitted. Instead, students start discussing immediately after the question is presented and the group discussions are concluded with a class-wide discussion instead of teacher explanation.

In a comparative study between Peer Instruction and the method described by Dufresne et al. [11], students felt that without the initial voting session they would be more inclined to be passive in group discussions and that the discussions would be more likely to be dominated by confident and/or "stronger" students [20]. Students reported that they used the initial voting session to formulate their own answer, which they in turn could use in the following discussion, and that they therefore would be more likely to engage in dialogue and defend their views. The researchers argued that the lack of an individual thinking period before discussion would result in less cognitive conflict at the start of discussion and thus students would be more inclined to accept dominant explanations.

There have been several recent quantitative studies on different aspects of Peer Instruction [21–24]. Previous studies have shown that the amount of correct votes increases after the discussion [16, 25]. One interpretation of these results could be that students choose the same as stronger students and not that they change their answer as a result of learning. In a study by Smith et al. [23], the researchers found evidence that the increase of correct votes is indeed a result of increased learning and not primarily due to the influence of other students. Even in groups where no students initially had the right answer, they observed an increase in learning. Smith et al. [24] showed that both the peer discussion and teacher explanation are important for facilitating learning with Peer Instruction. Excluding either the discussion or teacher explanation showed less learning than the combination of both.

A common practice when using Peer Instruction is to show students the results from the initial voting session prior to the group discussion [26]. Perez et al. [22] found that the probability of students switching to the alternative with the majority of votes increased by 30% if the results were displayed to the students. The researchers provided several interpretations for these findings. A clear majority could function as a stimulus for focused discussions and students might discover flaws in their original reasoning by trying to identify why most students chose one particular alternative. Another interpretation was that students simply switched to this alternative based on the consensus of nearby students. Confidence in their own choice has been shown to be significantly higher if students can see that they initially voted for an alternative in the majority [21]. This was persistent whether the alternative was correct or not.

The study in this paper was a part of EU-cofounded projects (EduMecca, Do-it, Done-it and Global-SRS) at Sør-Trøndelag University College in Trondheim, Norway. One goal of these projects was to develop an online SRS designed for effective classroom teaching, where students can use their own mobile device, such as smart-phones, as a voting device as compared to the traditional "clickers". As well as evaluating the system from a technical point of view, we also investigated different methodological implementations. More information about these projects can be found at our web page (http://www.histproject.no/).

The inspiration for this study was the findings of Nicol and Boyle [20]. We wanted to go deeper into, and try to examine, the effects of the initial voting session (and thinking period) with Peer Instruction. To do so, we divided the study into two parts: a quantitative part based on survey data and video of students engaged in peer discussion during class and a pure qualitative part based on in-depth analysis of student interviews. In this paper we focus on the interviews and investigate students' experiences regarding use of Peer Instruction and the effect of the initial voting session. The analysis of the video material is ongoing and will be presented at a later date. The paper starts with a description of study design and analysis methods followed by a presentation of the results. We conclude with a discussion of our findings; in particular, we discuss the results in light of the recent quantitative studies on Peer Instruction.

## 2. Method

The study was conducted in an introductory physics course for preparatory engineering students. The lectures usually consisted of $2 \times 45$ min. sessions generally constructed so that the first 45 min. dealt with new theory while the second focused on practical work, mainly problem solving. Four parallel classes with of a total of seven teachers (three classes used two different teachers) used SRS for eight weeks. One of the authors (K. L. Nielsen) was the teacher of the class with a single teacher. The second author (G. Hansen-Nygård) was present in lectures where SRS was used to function as an observer. Each class consisted of 50–70 students, the majority being male students. The theoretical part of the lectures was traditional teacher-style lectures (using digital blackboards) except that the teacher would present the students with 1–4 quizzes consisting of conceptual multiple choice questions. Each student borrowed an Apple iPod Touch which was used as a voting device. When SRS was used, the iPods were handed out at the beginning of the lecture and collected at the end. SRS was used in two of a total of three lectures each week due to some lectures running parallel timetables. To minimize the amount of variables in our data, all classes used the same set of multiple choice questions. We prepared around 50 concept questions which the teachers used in the course of the study.

In order to investigate the effects of the initial thinking period with Peer Instruction, we also used another method which we called "Classic" as the reference. This method was similar to Peer Instruction apart from the initial voting session (and thinking period) being omitted; students started discussing immediately after the question was presented. In both methods, the discussions were concluded with a vote and a teacher explanation. During this project, two classes started with Peer Instruction while the other two started with Classic. After four weeks, we switched the methods in all classes.

In order to get as reliable feedback from the students as possible, we did not emphasize that we were going to compare different methods. The teacher would just give different instructions based on the method used; for instance, to think individually without speaking to their fellow students in the case of the initial voting session of Peer Instruction. In both methods, students were encouraged to discuss the presented question in small groups for 2–4 minutes. The initial thinking period with Peer Instruction usually lasted about 1-2 minutes.

The students were interviewed twice, once after Peer Instruction and once after Classic. One exception was one class which already had experience with SRS and the Classic method. That particular class had already been interviewed about SRS, and therefore was only interviewed once at the end of the 8-week testing period. The interviews were conducted by one of the authors (GHN). The first interview was about three hours and consisted of questions regarding students' experiences with SRS in general. The second interview was shorter (around 60 minutes) and focused directly on the differences between the two methods. Many of the topics and questions in the second interview were created based on student feedback in the first interviews. The student interviews were conducted as focused (semistructured) group interviews, which is recognized as a reliable method of revealing informants' perspectives [27]. There were four groups (one from each class) with four students per group consisting of both male and female students.

The interviews were analyzed using analytical tools from grounded theory. We want to emphasize that we have not conducted a true grounded theory analysis, that is, with theory sampling to achieve theoretical saturation, but rather we have borrowed the analytical tools. These include a three-step coding scheme (line-by-line coding, focused coding, and categorization) adapted from Charmaz [28, 29]. This method is an appropriate direction for analysis of topics such as personal experiences, opinions, feelings, and attitudes [28]. The first step is to examine each line of data material and code it to define events that appear or are represented [28, 29]. We relied heavily on "in vivo" codes, that is, using the interviewers own words in the early stages of coding. This was to avoid misinterpretations of students' utterances and assure that we maintained a close relationship between the codes and what expressed by the students. The next phase is to focus on several lines or paragraphs of the interviews (focused coding) where the most significant line-by-line codes are identified. Thus, we are left with a smaller number of codes that give a more accurate description of the data. In the last step, categorization, focused codes are treated more analytical and conceptual [28]. Each focused code is described in detail: its properties, its consequences, how they relate to other focused codes, and conditions under which they arise, is maintained, and changes [29]. This process often led to several focused codes being merged when very close relationships were discovered. In the end we have a small number of categories that describe students' most significant experiences, in this case their experiences of the two SRS-sequences, Peer Instruction, and Classic.

## 3. Results

The analysis of the interviews resulted in three categories: (1) *Argumentation and explanation*, (2) *Peer Instruction: Opportunity for individual thinking,* and (3) *Seeing the results: Authority of the majority*. The first category deals with the students' experiences of generating explanations and presenting arguments in group discussions, and how they perceive explanations and arguments from fellow students (and to some extent the teacher). It is not about the content of the arguments and explanations *per se*, but rather about the students' own experiences of the process of generating and presenting them and how this relates to their learning. The second category is about thinking without the influence of others and how Peer Instruction gives an opportunity to reflect more deeply upon the questions and forming one's own opinion, resulting in increased participation and confidence during discussion. The last category focuses on seeing the results from the initial vote in Peer Instruction and how a clear majority can influence students' decision-making and the group discussion.

*Category 1: Argumentation and Explanation.* Learning physics is more than just memorizing formulas. Although memorization is an important part of their learning process, students also emphasize the importance of being able to reflect upon and solve problems, preferably without the help or influence of others (Category 2). However, according to the students, the best confirmation that they have learned physics is when they are able to explain the solution to other students. Then they have to challenge themselves to explore other ways of thinking in order to generate explanations and arguments that will make the solution both convincing and understandable. In addition, explanations from their peers are often easier to understand because students have the same foundation and speak "the same language." Students feel that the teacher is on a "higher level" and often uses complex words and phrases which can make his/her explanation hard to follow.

> You have to sort of re-learn it when you are going to explain it to others.
>
> —
>
> It becomes other words then [when peers explain], because it is like-minded people who repeat what the teacher said in a different way.

According to the students, good group discussions with SRS start with an uncertainty or disagreement about the answer and all students collaborate towards a consensus.

Different students remember different things, making it important that everyone participates. Also, if a student does not object or ask questions, it can be interpreted as agreement and that he/she understands the arguments being presented. In other words, being passive can result in students not being given enough explanation to understand the solution, and selecting the correct answer does not benefit learning if they have not understood why it is correct.

> I think it is important that everyone participates so that they can follow and understand why it is not like this and why it is like this.
>
> —
>
> If I do not say what I think, then maybe they just expect me to agree with them. And then no-one cares to explain it. So even if I vote the same as them and get the right answer, it does not help if I do not understand it.

When students feel that they understand the question and have a good argument, they are more likely to be active during the discussion and try to convince their fellow students. Being convinced or proven wrong is an important part of their learning. As one student puts it:

> I think you remember it better for later if you are proven wrong.

On the other hand, if the students are uncertain about the question or the solution, they are more inclined to be passive and withdraw from the discussion. Another factor that can increase passivity is sitting with students they do not know well. They are afraid of making a fool of themselves when they do not know how their fellow students will react to their arguments. Passivity in group discussions is also prominent if a student in the group is regarded as being skillful or "strong." The arguments and explanations of stronger students are valued higher than those with equal or lesser skills. Students are then more likely to only listen to these explanations rather than try to find out the solution for themselves, often accepting the others' conclusions without fully understanding them.

> You are more passive when you are uncertain, because then you listen to what the others say. And it might be the case that you did not understand the question and the alternatives the way you should have, and when the guy next to me says that it is like this and this: "OK, then maybe it is like this then". And then you look at it and "OK, I think this also", so without being certain of the answer I will vote the same.
>
> —
>
> When we sit and discuss there are certain people in the class you know are very skillful. So if someone in this group talks very loud or a person you know is very skillful says "No, it is B because…", and then they start to talk, then everyone else in the group will just shut up and listen to what that person has to say. Then it becomes like "Yeah, OK. So maybe it is like this", if you are uncertain. Then it is easy that you just listen to the others in the group rather than try and find the answer for yourself.

The results from the voting session can thus give the teacher a "false" image of the level of understanding among the students. Even though the majority of students have voted for the correct answer, it does not mean that the majority have understood why it is correct. Students therefore emphasize the importance of the explanation given by the teacher after group discussions. The teacher should explain thoroughly both the correct and incorrect alternatives. Learning is not only gained by understanding why an alternative is correct. Understanding why an alternative is incorrect can be just as fruitful and often crucial, especially for students who have voted for the incorrect alternative.

> Even if the majority has voted correctly, I still think a short explanation is needed. Because the situation can be so that you just voted what the guy next to you thinks is correct and you just follow him. So it is good that we still get a short explanation.
>
> —
>
> There was one question we sat and discussed which we were absolutely sure was the correct answer. Then it turned out that it was actually wrong, and then I think that it is good that they [the teachers] explain. That they do not just say that "This is correct" and just explain this alternative, but that they go through all the alternatives and explain why they are wrong or show us with illustrations. Because then it is much clearer: "Yes, of course it is like this".

*Category 2: Peer Instruction: Opportunity for Individual Thinking.* The SRS gives students an opportunity to engage in solving problems in group discussions during lectures. If this is to really benefit their learning, however, students emphasize that it is important that they are able to reflect more deeply upon the problem at hand and struggle with it. Students feel that Peer Instruction, with its initial thinking period, gives them an opportunity to actually involve themselves in the question before discussions. This will not only enhance learning during discussion, it will also make it easier to remember the solution and explanation given by the teacher after the discussion.

> Yes, we have enough time to involve ourselves in the question [with Peer Instruction], and that is what's important. That you have thought deeply

> about it. Because it's first then you really have a benefit of the answer.
>
> —
>
> If you first have pondered over something, and maybe you did not find the solution, and then you get the explanation, then it is much easier to remember it [the solution]. When you first had the opportunity to struggle with something, and then you get the explanation, then you remember much better.

For students to get involved in the questions, they feel that it is very important that they are able to form their own opinion without the influence of others. This can be difficult to achieve when using the Classic method because they go straight into group discussion after the question is presented. They are not given the opportunity to think individually without being "colored" by other students.

> I think it was much better [with Peer Instruction], because you are allowed to think for yourself and not having objections from everybody else.
>
> —
>
> You get a chance to make up your own mind before you get colored by what everybody else thinks.
>
> —
>
> You do not have time to think before you are influenced by the others' opinions [with Classic].

Students use the initial thinking period with Peer Instruction to construct their own mental image of the problem and an explanation which they can use in the following discussion (Category 1). When using the Classic method, explanations have to be generated during the discussion period (or when the teacher is reading the question), but often students experience that they barely have time to think before someone takes control and starts talking. When these students start to "think out loud", other students will have a difficult time thinking and working out logical arguments (unless they step in and actively participate by uttering their own thoughts). They can quickly be drawn towards the arguments and conclusions of the students taking control and vote the same alternative without having been able to think for themselves and understand the answer.

> I think that it blocks your own thoughts if you first have to listen to other people's ideas without having thought about it yourself; that you in a way forget to think for yourself.

With an explanation ready, students feel they have much more to contribute to the discussions. Everyone in the group is more likely to be heard because they can present more convincing arguments for their opinions (Category 1). The group consists of stronger individuals who are more inclined to defend their views when they are given time to think for themselves. This way, both stronger and weaker students can benefit more from the Peer Instruction method since weaker students have greater opportunities to construct and present arguments, while stronger students have a higher probability of having their arguments challenged.

> After we have thought by ourselves, we have so much more to say, rather than when we went straight into discussion.
>
> —
>
> The initial thinking time you have, I think that is great! Then you are able to reflect over what you think so that you have a better basis for participating in the discussion in your group.
>
> —
>
> I feel that there will always be someone who dominates more, but now everyone had something to bring to the table, and everyone was heard because you had a better explanation for your opinions.

With Peer Instruction the students find using SRS more serious and orderly. When using Classic it can be hard to know, according to the students, when to stop thinking for themselves and start discussing. With Peer Instruction, however, they know that the discussion starts immediately after the initial voting session. It thus becomes easier to focus on what they should do, when they should think for themselves and when they should discuss. There is more time to come to a consensus and find a shared solution that everyone is comfortable with. Since they are likely to have formulated an explanation during the initial voting session, more of the discussion time is used for actual discussion.

> The latter was the absolute best, yes, really [Peer Instruction]. It was a little more serious from the start really, because you had to work alone and had thought by yourself first. You put a little more into it than if you went directly to the group discussion, at least I think so. Also you got started at once, you got a better focus.
>
> —
>
> We also noticed that it was difficult to reach a good explanation [with Classic] that we agreed and where we felt that "Yeah, it must be right." It was not always we were quite there yet, no, because there was such a short time. It was better when we had thought about it beforehand, and we had an explanation ready. So when we started to discuss there was more time to agree, and feel that we had the answer.

*Category 3: Seeing the Results: Authority of the Majority.* Students clearly prefer Peer Instruction and experience it as the best method with regard to their own learning. However, the method is not without its weaknesses. After the initial vote the results are shown on a histogram. This last category is about how seeing these results can affect the quality of the discussions and how this in turn affects the students' decision making. If there is a clear majority that has chosen a specific alternative, the discussion can often be guided towards this alternative. It is very likely the students just assume that this alternative is correct and try to work out *why* it is correct, rather than go through all alternatives to work out what is correct and what is not. With the Classic method they do not see any results before discussion and so every alternative is considered equal when they enter discussion.

> You become guided, or misguided, and lose focus of what you are supposed to discuss because so many in the class have voted "B".
>
> —
>
> If you just see that "B" has gotten most of the votes, then you might just end up with trying to explain why "B" is correct rather than trying to find out what is the right answer.
>
> —
>
> Because if 80% have voted for one of the alternatives, then there is a high probability that it is correct, right? So then the discussion is focused on finding out why it is correct. And if we hadn't chosen this [alternative], it would have been better if we did not know about it.

Several students point out that it is not necessarily the majority that has chosen the correct alternative. Despite this, students feel that they still would be very likely to choose the same as the majority no matter what arguments are presented in the discussion, and without necessarily understanding why it is "correct".

> If most have voted "B" and you answered "A", you're very inclined to answer "B" however the others argue for or against it.

When students feel very uncertain they simply go for the alternative they perceive as most likely, which in most cases will be equal to the alternative which has got a clear majority of the votes (if any). If they have chosen an alternative with the majority of votes in the initial voting session, students become more confident in their choice, lowering their threshold for presenting arguments (Category 1). Even though they may not have the best arguments for this alternative, the students feel they are likely to be more certain of its correctness and defend their choice. The picture is reversed if their initial vote is in the minority. According to the students, they have to be very confident in their arguments to defend such an alternative.

> *Interviewer*: So the 20% of students who have voted for alternative "C" might not fight as hard to defend alternative "C" because 80% have voted alternative "B"?
>
> *Everyone*: Yes.
>
> *Student*: It's not even certain that people will admit that they have answered "C". It is likely that you will just ask the group "OK, why is it B?" That you just assume that it is "B" because the majority has voted "B". So then you will just try to explain why this is correct, even though it might not be the correct answer. It might be the case that the majority voted incorrectly.

Seeing a clear majority displayed in the results from the initial vote might have most effect on the group discussion when everyone in the group has answered the same as the majority. The students agree among themselves in the group and the majority of the class agrees with them. Then they are likely to lose interest in the question and talk about something else.

> It just wasn't interesting to talk about it then [when they saw the results from the initial vote].

Therefore students feel that it would be much better to not see the results of the initial vote until after the re vote. They find it interesting to see if the class have changed their mind during discussion, so several students point out that the best method would be to use Peer Instruction, but to wait to show the results of the initial vote until after the discussion to minimize "damage" to the discussion. As one student puts it:

> If you remove it [the results from the initial vote] there would have been better discussion per person in the class.

## 4. Discussion and Conclusion

We have studied students' experiences of Peer Instruction with and without the initial voting session using focus group interviews and a grounded theory-based analysis. Students value the initial voting session as a means of delving more deeply into the question, generating a mental image of the problem and constructing an explanation with convincing arguments to use in the following discussion. This is consistent with findings of Nicol and Boyle [20]. In order to generate good explanations, students need time to clarify their thoughts and reflect more deeply upon the problem at hand [5]. Without the initial voting session such a process becomes very difficult because they will only be "allowed" to think and reflect until the first student starts to speak or "think aloud." The consequence is often less participation and higher probability of accepting explanations presented by "stronger" students.

Although the students emphasize the need to construct convincing arguments in order to defend their views, they do not necessarily feel this merely to "win" the discussions. A simplistic view of argumentation is to view it as a battle where one tries to defeat one's opponent. Duschl and Osborne [5] argue that argumentation is also an exploration to find and fill out holes in one's knowledge. This view is supported by the students. They want to be convinced of the arguments presented, and to be proven wrong can function as a high facilitator of learning as it can challenge their thinking and reveal flaws in their understanding. The initial voting session results in more arguments and ideas being presented during discussion and the students are more likely to come to a consensus.

Another consequence of a greater number of ideas presented at the start of discussion is an increase in the probability of disagreement. Difference is an important part of learning, as Duschl and Osborne [5] stated fittingly: "without difference, there can be no argument, and without argument, there can be no explanation" (p. 53). This is not to say that initial conflict necessarily translates to increased learning. It is also important that the students actively engage in either trying to convince their neighbor or to be convinced, that is, that they state their thoughts and (dis)agreement. Passive compliance and/or insufficient verbalization during group discussion can have a detrimental effect on the learning outcome of discussions [2].

Our research is consistent with the findings of Perez et al. [22] and Brooks and Koretsky [21] that displaying the results of the initial voting session can affect students' decision making and confidence during discussion. As predicted by Perez et al. [22], the students in our research experience that an alternative with a clear majority of the votes becomes a center of focus for the group discussion, although not in a positive sense. Students point out that they will not necessarily try to find out the reason why the majority has chosen one specific alternative; they will automatically *assume* that it is correct. The histogram becomes an argument in itself, an argument much stronger than those presented in the discussion or through individual reasoning. The focus becomes on finding out *why* the alternative is correct and not *if* it is correct.

The bias from seeing the results from the initial voting session was also shown to be stronger for more difficult questions [22]. Although our students do not specify this connection in particular, they do emphasize a stronger influence when they feel uncertain, and it is likely there is a high correlation between the difficulty of the question and the level of uncertainty among the students. By not showing the results prior to discussion, every alternative is initially considered equal. Rather than trying to "force" a solution upon an alternative with a clear majority, students are more likely to evaluate each alternative more thoroughly.

Without a total evaluation of all alternatives, students will be more dependent on the teacher in order to accept the correctness of the alternative, rather than having it come through their own argumentation [5]. This is a prominent feature in novice students [25], and it is therefore not surprising that the students in our research emphasize the importance of the teacher explanations following the discussions. Our students stress the importance of the teacher carefully explaining the correct alternative and why it is correct in order to be convinced of the solution. This is consistent with the findings of Nicol and Boyle [20]. In addition, our students also emphasize the importance of the teacher explaining why the *incorrect* alternatives are incorrect. If they do not choose the correct alternative, or are not sure of its correctness, they need an explanation to be convinced and fully understand the solution.

Smith et al. [24] argued that there is a synergy effect between peer discussion and the teacher explanation, making the combination facilitate more learning than either on its own. A majority of the students in their study also agreed that the peer discussion made them more prepared for the following explanation. The students' experiences in our study are consistent with these findings in that the students feel that reflecting and struggling with a question makes it easier to remember the following explanation from the teacher. The initial voting session is crucial for this to happen as deeper reflection is difficult to achieve during peer discussion if they do not have time to formulate their thoughts without the influence of others.

In the study by Smith et al. [24], many students even reported frustration when the teacher explanation was excluded when using Peer Instruction. An explanation for this can be that the teacher explanation functions as a closure for the SRS session and also removes any last doubts. Our students feel that they are very seldom 100% sure about the correctness of the answer, and therefore they need the feedback that they have not only chosen the correct answer, but have also understood the solution correctly. Not receiving this feedback is likely to cause frustration.

In summary, the students in our research experience Peer Instruction as more beneficial to learning when the initial voting session is included, but where the voting results are not shown until after the revote. Our study has shed more light on recent findings on Peer Instruction from the students' point of view. Nevertheless, more research is required to obtain a more complete picture of the effect of the initial voting session; for instance, to verify students' claims of more fruitful discussions, where more arguments are presented. Our ongoing video analysis of students engaged in peer discussion, both with and without the initial voting session, should be able to give more insight to their experiences.

## References

[1] A. A. Gokhale, "Collaborative learning enhances critical thinking," *Journal of Technology Education*, vol. 7, no. 1, pp. 22–30, 1995.

[2] A. S. Palincsar, "Social constructivist perspectives on teaching and learning," *Annual Review of Psychology*, vol. 49, pp. 345–375, 1998.

[3] M. Prince, "Does active learning work? A review of the research," *Journal of Engineering Education*, vol. 93, no. 3, pp. 223–231, 2004.

[4] J. K. Knight and M. K. Smith, "Different but equal? How nonmajors and majors approach and learn genetics," *CBE Life Sciences Education*, vol. 9, no. 1, pp. 34–44, 2010.

[5] R. A. Duschl and J. Osborne, "Supporting and promoting argumentation discourse in science education," *Studies in Science Education*, vol. 38, no. 1, pp. 39–72, 2002.

[6] I. Beaty, "Transforming student learning with classroom communication system," *Research Bulletin*, vol. 3, pp. 2–13, 2004.
[7] S. W. Draper and M. I. Brown, "Increasing interactivity in lectures using an electronic voting system," *Journal of Computer Assisted Learning*, vol. 20, no. 2, pp. 81–94, 2004.
[8] C. D. Kam and B. Sommer, "Real-time polling technology in a public opinion course," *Political Science and Politics*, vol. 39, no. 1, pp. 113–117, 2006.
[9] G. Masikunas, A. Panayiotidis, and L. Bruke, "The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning," *Research in Learning Technology*, vol. 15, no. 1, pp. 3–20, 2007.
[10] S. A. J. Stuart, M. I. Brown, and S. W. Draper, "Using an electronic voting system in logic lectures: one practitioner's application," *Journal of Computer Assisted Learning*, vol. 20, no. 2, pp. 95–102, 2004.
[11] R. J. Dufresne, W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk, "Classtalk: a classroom communication system for active learning," *Journal of Computing in Higher Education*, vol. 7, no. 2, pp. 3–47, 1996.
[12] R. R. Hake, "Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses," *American Journal of Physics*, vol. 66, no. 1, pp. 64–74, 1998.
[13] R. H. Hall, H. L. Collier, M. L. Thomas, and M. G. Hilgers, "A student response system for increasing engagement, motivation, and learning in high enrollment lectures," in *Proceedings of the Eleventh Americas Conference on Information Systems*, Omaha, Neb, USA, August, 2005.
[14] S. J. Pollock, "Transferring transformations: learning gains, student attitudes, and the impacts of multiple instructors in large lecture courses," in *Proceedings of the Physics Education Research Conference*, vol. 818 of *AIP Conference Proceedings*, pp. 141–144, Salt Lake City, Utah, USA, August 2005.
[15] S. P. Rao and S. E. DiCarlo, "Peer instruction improves performance on quizzes," *Advances in Physiology Education*, vol. 24, no. 1, pp. 51–55, 2000.
[16] C. H. Crouch and E. Mazur, "Peer Instruction: ten years of experience and results," *American Journal of Physics*, vol. 69, no. 9, pp. 970–977, 2001.
[17] E. Mazur, *Peer Instruction: A User's Manual*, Prentice Hall, Upper Saddle River, NJ, USA, 1997.
[18] P. M. Len, "Different reward structures to motivate student interaction with electronic response systems in astronomy," *Astronomy Education Review*, vol. 5, no. 2, pp. 5–15, 2007.
[19] M. C. James, "The effect of grading incentive on student discourse in peer instruction," *American Journal of Physics*, vol. 74, no. 8, pp. 689–691, 2006.
[20] D. J. Nicol and J. T. Boyle, "Peer Instruction versus class-wide discussion in large classes: a comparison of two interaction methods in the wired classroom," *Studies in Higher Education*, vol. 28, no. 4, pp. 457–473, 2003.
[21] B. J. Brooks and M. D. Koretsky, "The influence of group discussion on students' responses and confidence during peer instruction," *Journal of Chemical Education*, vol. 88, no. 11, pp. 1477–1484, 2011.
[22] K. E. Perez, E. A. Strauss, N. Downey, A. Galbraith, R. Jeanne, and S. Cooper, "Does displaying the class results affect student discussion during peer instruction?" *CBE Life Sciences Education*, vol. 9, no. 2, pp. 133–140, 2010.
[23] M. K. Smith, W. B. Wood, W. K. Adams et al., "Why peer discussion improves student performance on in-class concept questions," *Science*, vol. 323, no. 5910, pp. 122–124, 2009.
[24] M. K. Smith, W. B. Wood, K. Krauter, and J. K. Knight, "Combining peer discussion with instructor explanation increases student learning from in-class concept questions," *CBE Life Sciences Education*, vol. 10, no. 1, pp. 55–63, 2011.
[25] J. K. Knight and W. B. Wood, "Teaching more by lecturing less," *Cell Biology Education*, vol. 4, no. 4, pp. 298–310, 2005.
[26] J. E. Caldwell, "Clickers in the large classroom: current research and best-practice tips," *CBE Life Sciences Education*, vol. 6, no. 1, pp. 9–20, 2007.
[27] A. Johannessen, P. A. Tufte, and L. Kristoffersen, *Introduksjon til Samfunnsvitenskapelig Metode*, Abstrakt Forlag AS, Oslo, Norway, 2004.
[28] K. Charmaz, "Grounded theory," in *Rethinking Methods in Psychology*, J. A. Smith, R. Harre, L. Langenhove et al., Eds., pp. 27–49, Sage, London,UK, 2001.
[29] K. Charmaz, "Grounded theory," in *Qualitative Psychology; A Practical Guide to Research Methods*, J. A. Smith, Ed., pp. 81–110, Sage, London, UK, 2003.

# Paper D

## How the Initial Thinking Period Affects Student Argumentation during Peer Instruction: Students' Experiences versus Observations

Nielsen, K.L., G. Hansen-Nygård and J.B. Stav

Preprint 2012. Submitted to *Studies in Higher Education*.

# How the Initial Thinking Period Affects Student Argumentation during Peer Instruction: Students' Experiences versus Observations

Kjetil Liestøl Nielsen, Gabrielle Nygård Hansen and John Birger Stav
*Sør-Trøndelag University College, 7004 Trondheim, Norway*

We have compared students discussing multiple choice quizzes during Peer Instruction with and without the initial thinking period before discussion. Video clips of students engaged in peer discussion in groups of three of varying group combinations, a total of 140 different students in all, were compared to students' own experiences extracted from group interviews (16 students in groups of four and a total of seven interviews) and survey results (109 responses). The initial thinking period was found to increase argumentation time during discussion, consistent with students' own experiences. However, while students felt that the initial thinking period increased participation and contribution of ideas among all group members, we only found significantly improved discussion for two out of three group members, those already most active. We did not find any statistically significant difference for the least active students with or without the inclusion of the initial thinking period.

## 1. INTRODUCTION

For several decades researchers have explored the benefits of using electronic voting systems, also referred to as a Student Response System (SRS) in lectures (e.g. [1, 2]). A common way of using such a system is to present students with a multiple choice quiz, which they will discuss in small groups before voting with a small handheld device. Using an SRS in class has been shown to have many positive effects on lectures, for example, increasing student engagement (e.g. [3]) and improving student performance (e.g. [4]). In more recent years, more researchers have studied various aspects of using SRS, comparing different ways of using such systems, rather than comparing SRS to its not being used. For instance, the teacher explanation of the solution after the discussion and voting session has been shown to be an important part of increasing the learning gains of SRS-use [5]. Another example is the study by James [6] who found that grading SRS-participation, where more points are given to the correct answer, can negatively influence group dynamics during discussions.

A popular approach when using SRS is to include an initial thinking period and voting session before entering discussion, often referred to as the 'Peer Instruction' sequence [7]. Students are given about a minute to think individually on the question without talking with fellow students. After giving their vote, students enter group discussion for a few minutes where they argue for their choice. The session ends with a revote and the teacher explaining the solution. Students value this initial thinking period as they use it to generate an explanation which they can use in the following discussion, and are as a result less likely to be dominated by more skilled students [8].

In this paper, we explore this initial thinking period in more depth and illustrate how it affects the following discussion. In particular, we compare students' own experience of these effects to observations of students engaged in peer discussion during SRS, with and without the initial thinking period and voting session. We start this paper with a description of study design and research method. The results are divided into two sections: First we give a short summary of the results from the interview analysis, in particular those results directly covering the effects on students being given time to think and reflect before discussion. A more detailed presentation of the results from the interviews can be found in [9]. The second part presents the quantitative results from the video analysis. The consistencies and inconsistencies between part one and part two are discussed and we explore possible explanations for the results. The paper ends with a summary and conclusions.

## 2. METHOD

### 2.1. Description of the classes

In Norway, students are required to have a certain level of physics and mathematics (among other subjects) from senior high school in order to start on an engineering degree. At Sør-Trøndelag University College (HiST), in line with other University Colleges in Norway, students without these requirements can attend a preparatory course, lasting a full year and consisting of the required curriculum from senior high school. At the time of this study there were four such classes at HiST with approximately 50 students per class, one of which was taught by the first author. Apart from laboratory exercises and tests, there is no mandatory attendance for the lectures, which usually consist of 2x45 min. sessions divided between traditional teacher-style lecturing (using digital blackboards) and textbook problem solving.

### 2.2. Study design

During the spring semester of 2010, an in-house SRS was used for a period of eight weeks in all four classes in physics. There were no changes made to the lecture format during this period, apart from the teachers presenting students with 1-4 quizzes during the theoretical part of the lectures. In order to investigate the effects of the initial thinking period with Peer Instruction, we introduced a similar method, 'Classic', as a reference. The Classic sequence is similar to Peer Instruction, with the only difference being omission of the initial thinking period and voting session. Instead, students engage in discussion immediately after being presented with the quiz. Students were usually given 2-4 minutes to discuss. We encouraged teachers to aim for a discussion time of about three minutes, but teachers could adjust the time if needed. Two classes started with the Classic sequence while two started with Peer Instruction. After four weeks we switched the methods in all classes. We prepared around 50 concept questions which the teachers used during the study. All teachers used the same questions in order to reduce variables.

We interviewed four groups of students twice (one group from each class), the first time right before we switched SRS-sequences and the second at the end of the testing period. Each group consisted of four students, both male and female. During the first interview, we did not emphasise that our main focus was to compare different SRS-sequences in order to avoid biasing students' opinions for the next testing period. Consequently, the first interviews consisted of questions regarding students' general experience with SRS, ranging from positive and negative aspects of SRS, the role of the teacher, the quizzes, how they experienced group discussion and so on. One class had already had experience with SRS from the autumn semester (where the same students were interviewed about their SRS-experience) and the group from this class was only interviewed at the end of the testing period. The last interviews were more specific, directly targeting the differences between the two methods. The interviews were analysed using analytical tools from Grounded Theory (see [9] for more details). In addition to interviews, we conducted a survey at the end of the semester. The questions on the survey were mostly based on themes that arose during the student interviews.

In order to gain a deeper understanding of the effects of the initial thinking period, we filmed students discussing in groups during the lectures for four weeks, two weeks before and two weeks after switching SRS-sequences. Prior to filming, students signed permission waivers. Two video cameras were positioned in the front corners of the classrooms. In addition, we placed small audio recorders with groups which were chosen at random during each lecture. We opted for having students in groups of three and encouraged them to sit together to form such groups, but otherwise we did not interfere with how students formed groups in order to have the data material represent real life situations as closely as possible. We filmed students discussing a total of 15 different quizzes. However, due to time constraints, only six of these were analysed. The six quizzes were chosen based on having at least one quiz from each week of filming and the same amount before and after switching methods.

### 2.3. Video analysis

*Preparing the video material*

The video material was organised in small video clips, one for each group discussion. The video clips were synchronised between the two camera angles, composited to only show the relevant group, and audio from the recorders. By using two different camera angles, students' faces in the group were visible during the discussion, which simplified transcription. The video clips were initially transcribed by four assistants, who were trained to identify and transcribe only the parts of the discussions relevant to the quiz (i.e. excluding small talk such as plans after schools). They were also instructed to include a time code at the start and end of each utterance by the students. After the transcriptions were complete, a new assistant examined all videos marked for analysis to make sure all times codes were accurate and that all relevant discussion was included.

*Coding*

All utterances were coded based on a set of predetermined categories adapted from Kaartinen and Kumpulainen [10] and James [6]. The codes were separated into two main categories: argumentation and comments. Argumentation was defined as utterances that included an idea or explanation, i.e. utterances that included conceptual information which was presented with the intention of being related to the solution of the question. The argumentation category included three subcategories which described the purpose of the argument, the extent of the argument consisting of an idea that was new or had already been presented, and a description of the language used. Comments also had a set of sub-categories describing the nature of the comment, for instance, asking for opinions or stating uncertainty. However, in this paper we only focus on the two main categories and not the different sub-categories.

What constituted as an idea or argument was often context dependent. For instance, in a quiz about light and interference, the student utterance 'red light has the longest wavelength', could be interpreted as presenting an idea. If this was the first utterance in the group, the student was presenting the idea of longer wavelength being relevant to the answer to the quiz. On the other hand, if this utterance was preceded by another student asking 'OK, so which has the longest wavelength?', the utterance 'red light has the longest wavelength' would not be considered an argument since the intention was not to present or justify longer wavelength as a concept connected to the solution, for instance, by trying to rephrase the original argument, nor does it contain any new conceptual information. On the other hand, if the student were to answer 'red light has the longest wavelength, which means that it has less energy', (s)he is expanding the idea of longer wavelength being relevant by introducing the concept of energy and how the wavelength relates to this. Consequently, the coding process included careful considerations of what had previously been said in the discussions in order to assess the correct context. The coding process included several iterations to make sure the codes represented the actual material.

*Calculating Bias*

An important part of our video analysis was to investigate the level of 'dominance' or 'bias' in the groups, i.e. how much one student argued compared to others. This notion of bias was also used by James [6], who investigated students discussing in pairs. If one student in a discussion pair presented 80% of the ideas, P1, while the second presented the remaining 20%, P2, bias was calculated as P1-P2 = 60. We wanted a similar measurement for our data, but having three students in each group, the simple scheme of P1-P2 would not suffice. Instead we used the equation:

$$bias = 100\frac{STD(P1, P2, P3)}{STD(100, 0, 0)} \quad (1)$$

,where STD stands for population standard deviation, to calculate the bias. A bias of 100 translates to discussions where all ideas were presented by a single student, while a bias of 0 translates to discussions where every student in the group presented the same amount of ideas. This equation gives the same results as the method described by James [6] when used for two students instead of three. However, in our study we did not count the number of ideas, which was done in the study by James [6]. Instead, we measured the argumentation time based on the time codes placed during transcription. Students would often start the discussion by presenting new ideas or by justifying their opinions, but further into the discussions, argumentation often consisted of either restating, rephrasing or expanding on old ideas. This made counting individual ideas difficult. It has been shown that a word count can be used instead of counting ideas with similar results [11]. In our study we found that some students would speak very softly, especially during the Classic method, making it sometimes hard to identify specific words due to ambience noise although the nature of the argument could be identified as well as the length (in time). Speech time was therefore considered as a more reliable option in our study.

*Statistical analysis*

The video material presented a complex data set with regard to statistical testing. The group combinations would also change from day to day and not all students were present in the video material during both Peer Instruction and Classic. As a result, the material was subjected to different levels of statistical dependence. We chose to average the speech time for each student in identical group combinations and treated the two data sets (student speech time with Peer Instruction and Classic) as independent. In other words, two data points from the same student were treated as independent if they came from different group combinations. This way we had two almost fully independent data sets, with 126 data points from Classic and 153 from Peer Instruction, which simplified the statistical testing.

There were two group combinations which were present during both Peer Instruction and Classic. However, removing the data points from these groups, in order to maintain our definition of independence, did not change any statistical significance (although removing these slightly decreased the p-value in the cases with statistical significance ($p < 0.05$) and increased the p-value in the case of no statistical significance ($p > 0.05$)), and therefore we did not remove them from the calculations. The data sets failed normality tests performed in SPSS (version 19), and the Mann Whitney U test for nonparametric data was therefore chosen as the main tool for comparing the two samples. Non-normal distributions often include a high amount of skewness. Thus, we mainly focus on the medians when presenting our results as it may give the reader a more representative impression of our findings. Nevertheless, we will also present the mean values and standard deviations for comparison.

## 3. RESULTS

### 3.1. Student Experiences

*Student experiences: argumentation*

The initial thinking period with Peer Instruction was seen as invaluable for good group discussions and clearly preferred by the students (Figure 1). They could gather

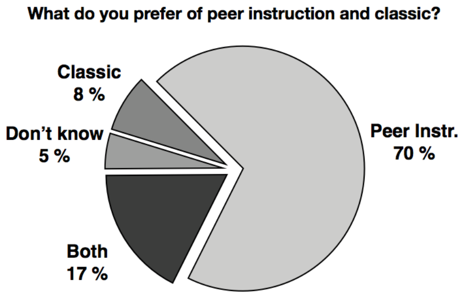



FIG. 1: Responses from the survey. N = 109.

their thoughts and reflect on the questions without the influence of others. In other words, the students could develop an opinion which had not been 'coloured' by other students and which they therefore could call their own. This was also reflected in the survey (Table 1, statement S1). Feeling an ownership to their opinions and explanations was one of the most emphasised aspects during the student interviews.

I think it was much better [with Peer Instruction], because you are allowed to think for yourself and not having objections from everybody else.

—

You get a chance to make up your own mind before you get coloured by what everybody else thinks.

—

You don't have time to think before you are influenced by the others' opinions [with Classic].

In the Classic method students were thrown straight into discussion and, consequently, reflection about the question had to be done during the discussion time. The problem was, according to the students, that when other students started to speak, it was very difficult to reflect on the question. As one student put it:

I think that it blocks your own thoughts if you first have to listen to other people's ideas without having thought about it yourself; that you in a way forget to think for yourself.

*Student experiences: bias*

Without the opportunity to reflect over the quiz before discussion, students felt that they were more susceptible to accepting arguments presented by 'stronger' students (S5). Opinions from students with a reputation for high skills are valued higher than those of lesser skills, which could result in explanations being accepted by the rest of the group even though they might not have fully understood them. This was especially prominent if students were uncertain about the solution. With the Peer Instruction method, however, students felt they would be more prepared for the discussions (S3) because they have used the thinking period to reflect on the problem. Being more prepared, students would also be more likely to participate in argumentation. According to the students, they had more to say and the group as a whole would contribute more (S4), and everyone was heard to a higher degree since they could present more convincing and thought out arguments. This also meant that the 'actual' discussions would begin sooner when students had an explanation ready (S2).

After we have thought by ourselves, we have so much more to say, rather than when we went straight into discussion.

—

The initial thinking time you have, I think that is great! Then you are able to reflect over what you think so that you have a better basis for participating in the discussion in your group.

—

I feel that there will always be someone who dominates more, but now [with Peer Instruction] everyone had something to bring to the table, and everyone was heard because you had a better explanation for your opinions.

### 3.2. Video analysis

*Video analysis: argumentation*

Speech time relevant to the quiz during group discussion had a median of 15.7s per student for Classic and 20.3s for Peer Instruction (PI). The difference was found to be statistically significant (p=0.035). By dividing speech time into argumentation and comments, we see that the difference in speech time was mostly due to argumentation, which had a 91% increase in medians from 3.9s on Classic to 7.4s on PI (p=0.007). There was no statistical difference between commenting time (p=0.300). The distribution of argumentation time for Classic and PI can be seen in Figure 2. The medians, mean values and standard deviations are summarised in Table 2.

*Video analysis: Bias*

We found a statistically significant decrease in bias from Classic to PI (p=0.031), with a median of 59.9 on Classic and 50.2 on PI. The distribution of the bias can be seen in Figure 3. Most noticeable is the high column between 90 and 100 on Classic, which represents discussions where (almost) all arguments were presented by a single

TABLE I: A selection of answers from the survey regarding the effects of the initial thinking period. The table shows the mean value and standard deviation (in parenthesis) for the responses. 5 = totally agree/very strong degree, 1 = totally disagree/very small degree. N = 109.

| Statement | mean (standard deviation) |
|---|---|
| 1. The individual vote with Peer Instruction gives me a better opportunity to think for myself and generate my own opinion surrounding the question. | 4.00 (0.87) |
| 2. I feel the actual discussions start faster with Peer Instruction versus Classic. | 3.72 (1.05) |
| 3. The individual voting session makes me more prepared for participating in the following discussion. | 3.88 (0.93) |
| 4. The group as a whole contributes more during group discussion when we had an individual vote before discussion. | 3.75 (0.88) |
| 5. When we go straight into group discussion without the individual vote, it is easier to be influenced by dominating parties in the group. | 3.72 (0.90) |

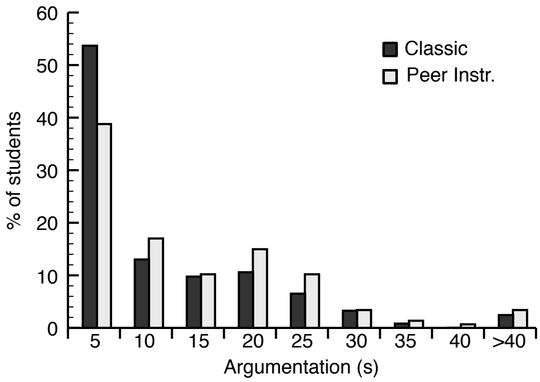



FIG. 2: Comparison of argumentation time (s) between Peer Instruction (N=153) and Classic (N=126).

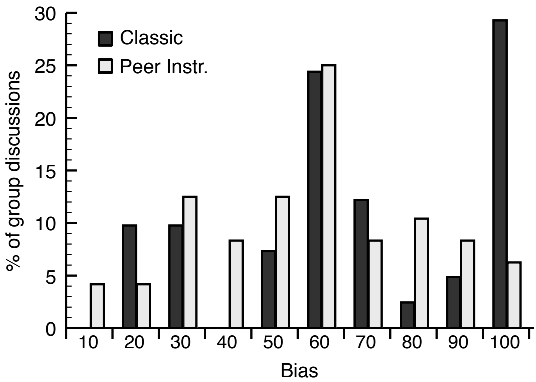



FIG. 3: Comparison of bias between the Classic (N=41) and Peer Instruction (N=50) method. One group during Peer Instruction and one during Classic did not have any argumentation and was therefore removed from the calculations of bias.

TABLE II: Total speech time, argumentation time and commenting time compared between Classic (N=126) and Peer Instruction (N=153)). The table shows the **Median** (Mean value, Standard deviation) in seconds, as well as the p-value from the Mann-Witney U test. Statistically significant differences have p-values marked in bold letters.

| Discourse | Classic | Peer Instruction | p-value |
|---|---|---|---|
| Speech time | **15.7**(19.7, 15.9) | **20.3**(23.0, 15.8) | **0.035** |
| Arguments | **3.9**(8.2, 10.1) | **7.4**(11.1, 11.1) | **0.007** |
| Comments | **9.6**(11.5, 8.9) | **10.6**(11.9, 7.8) | 0.300 |

student. To get more insight into the bias, we divided argumentation time between the most dominant student (i.e. longest argumentation time), D1, the second most dominant student, D2, and the least dominant student in the group, D3. Although we see an increase in median argumentation time for all group members from Classic to PI, only D2 is statistically significant (p=0.014) with a median of 4.2s for Classic and 6.9s for PI. These results are summarised in Table 3.

This suggests that the decrease in bias is mostly due to D2 students becoming more involved rather than D3. This claim is further supported by comparing the share of argumentation time between each member of the group for Classic and PI, and also, by comparing the number of students having no argumentation time. D1 students decrease their share of argumentation time with a median of 72.4% for Classic and 57.4% for PI, the difference being statistically significant (p=0.046). For D2 students, we find a statistically significant (p=0.048) increase from 19.9% with Classic to 30.1% with PI. Although we also find an increase for D3 students, the difference was not statistically significant (p=0.246).While both D2 and D3 saw a decrease in not having any argumentation time from Classic to PI, only D2 is statistically significant

TABLE III: Calculation of bias as well as argumentation time and share for the different dominance level. The table shows the **Median** (Mean value, Standard deviation) in seconds, as well as the p-value from the Mann-Witney U test. Statistically significant differences have p-values marked in bold letters. One group during Peer Instruction (N=51) and one during Classic (N=42) did not have any argumentation and was therefore removed from the calculations of bias and argumentation share.

| Domination | Classic | Peer Instruction | p-value |
|---|---|---|---|
| Bias | **60.0**(63.1, 28.3) | **50.2**(51.4, 24.3) | **0.031** |
| D1 arg. time (s) | **15.3**(16.3, 12.2) | **17.1**(19.7, 11.3) | 0.127 |
| D2 arg. time (s) | **4.2**(6.2, 6.3) | **6.9**(10.1, 9.0) | **0.014** |
| D3 arg time (s) | **0.0**(2.0, 3.9) | **1.0**(3.4, 5.1) | 0.116 |
| D1 arg. share (%) | **72.4**(72.5, 29.3) | **57.4**(63.8, 17.6) | **0.046** |
| D2 arg. share (%) | **19.9**(21.5, 16.1) | **30.1**(28.2, 13.3) | **0.048** |
| D3 arg. share (%) | **0.0**(6.0, 8.5) | **5.1**(8.1, 9.3) | 0.236 |
| D2 arg. = 0 (%) | 24 | 6 | **0.017** |
| D3 arg. = 0 (%) | 57 | 45 | 0.300 |

(p=0.017 and 0.300 for D2 and D3 respectively using Fisher's exact test) with 27% with Classic and only 6% with PI. A summary of the video analysis with respect to bias and the different dominance levels is shown in Table 3.

As we can see in Table 3, D3 had very little of the argumentation time for both Classic and PI. However, this does not mean that D3 students did not participate during discussions. Although D3 students took little part in the discussion with regard to presenting ideas and explanations, they were still to a large degree active with relevant comments. With Classic, D3 students had 22.6% (median) of the commenting time in the group and 21.2% with PI.

*Student positions in the groups*

Students were positioned in a linear manner in the groups, so that one student sat to the left, one in the middle and one to the right with all students facing the teacher. We wanted to ascertain if there was any correlation between student position and dominance level. To test whether there was any statistically significant difference in seating position we used the Chi Square test in SPSS. Since we do not differentiate between the left and right position, we calculated the average number of times students sat in either the left or right positions and compared it to the number of times students sat in the middle.

The results showed a significant difference for D1 and D3 students, but not for D2. D1 had a higher tendency to sit in the middle position of the groups, while D3 was more inclined to sit in one of the outer positions. D1 sat in the middle 47% of the time, with an average of 27% in the outer positions (p = 0.027), while D2 was 36% in the middle, with an average of 32% in the outer positions. The difference for D2 was not statistically significant (p = 0.701). D3 was in the middle 18% of the time with an average of 41% in the outer positions (p = 0.004). The results are summarized in Table 4.

TABLE IV: Student positioning for the different dominance level. Statistically significant differences have p-values marked in bold letters. *Note that this is the number of data points used in the statistical calculations.

| Dom. | Mid. pos. | Out. pos. | $\chi^2$ | p-value | N* |
|---|---|---|---|---|---|
| D1 | 47 % | 27 % | 4.909 | **0.027** | 66 |
| D2 | 36 % | 32 % | 0.148 | 0.701 | 61 |
| D3 | 47 % | 27 % | 8.321 | **0.004** | 53 |

## 4. DISCUSSION

### 4.1. Argumentation

With an increase of 91% in argumentation time medians from Classic to Peer Instruction, students' claim of more productive discussions, where more ideas and explanations were presented, is supported. With an opportunity to make up their own minds and generate an explanation before engaging in discussions, more of the discussion time can be used for presenting explanations instead of generating them. Some aspects of students' own descriptions of why there is less productivity with the Classic method seem to closely resemble the notion of 'production blocking', which is usually used to describe productivity loss in brainstorming groups ([12, 13]. Group discussions during SRS have several similarities to brainstorming groups, such as having a limited time to generate and present ideas and only one student being able to present ideas at a time (unless (s)he is interrupted). Production blocking refers to the aspect of students having to take turns in order to present ideas, which can lessen the effectiveness of the groups.

Students found it distracting hearing others' ideas as it could either 'block' or 'colour' their own thought process. This 'thought distraction' was presented as a possible explanation for why production blocking occurs in brainstorming groups, but some researchers have failed to find evidence of thought distraction and have argued against it [13, 14]. Still, it may be a viable explanation in our setting because of the differences between quiz discussions and brainstorming groups, the latter focusing on presenting a large amount of ideas, with no requirements on their validity, and which should not be subjected to peer evaluation. Generating a more complete explanation about the quiz solution requires a deeper cognitive process and organisation of thoughts. Since the explanations have to be convincing as well, the cognitive requirements increase further. It is likely easier to distract this deep cognitive process and thought organisation, and consequently, 'thought distraction' might be a part of the explanation why we find less argumentation during the Classic method.

Another explanation of production blocking, which is applicable in our setting, is simply due to students having to wait for their turn to speak, creating a delay between generating and presenting ideas. This has been found to interfere with the cognitive process of generating ideas [15]. The unpredictable nature of these delays can reduce the flexibility of idea generation while long delays can in addition disrupt thought organisation [15]. The initial thinking period with Peer Instruction gives the students a predictable time period where they know they can reflect on the solution without being interrupted, which could lessen the degree of the effects described by Njistad, Stroebe and Lodewijkx [15]. Although students still have to wait for their turn to speak (unless they interrupt the current speaker), the ideas and arguments are to a larger degree already generated during the initial thinking period and not during group discussion.

### 4.2. Bias

Researchers have argued that without the initial thinking period, there will be less 'cognitive conflict' at the start of discussions and students are thereby more inclined to accept dominant explanations [8]. In addition, the results from the interviews suggest that there is a high presence of 'evaluation comprehension' during discussion, i.e. students refraining from sharing ideas because of fear of negative peer evaluation [12]. This effect can be more prominent if some group members are regarded as experts [16]. Also, without having had the time to work on a convincing explanation before discussion, they are less confident in their arguments which could increase the degree of evaluation apprehension present. According to our students, opinions from stronger students had a strong impact during group discussion, but it should be noted that dominant students does not necessarily correspond to those actually being most skilled. Research have found that dominant group members can appear more competent to the rest of the group, even though they might lack competence, as a result of their behaviour making them seem as experts [17].

We did observe less of a tendency of domination or bias when using the Peer Instruction method compared to Classic. However, our results suggest that this is due to better communication between two group members, those already active to a large degree, D1 and D2, rather than all three. A possible explanation for this result could be due to students sitting in a straight line in their groups, all facing towards the teacher and digital blackboard. Studies have found that in groups with similar seating positions there is significantly less communication than when group members are sitting face to face [18, 19]. If the middle student speaks, for instance, to the student to the left, it creates a communication barrier towards the right. The middle student is facing away from, and sometimes even turns his/her back on the student to the right which makes it difficult for this student to effectively communicate with the rest of the group. The student in the middle position can reduce this barrier, for instance, by leaning back so that there will be a face to face connection between the left and right student, but (s)he still has to shift his/her attention between the right and left in order to effectively communicate with both students. The only way for a common centre of attention to occur is by focusing on the notebook in front of the student in the middle position, or on the quiz at the blackboard. The former is a possible scenario with explanations that involve drawing illustrations and manipulating equations.

The effects of the communication barrier are probably increased further by the fact that the most dominant students (with regard to argumentation) are often positioned in the middle, where communication is easier. Although one could argue that the most dominant students, D1, become D1 students simply by sitting in the middle position, research has shown that students with a dominant personality are more likely to choose such 'high-talk' positions [20]. Students that are shy, regard themselves as 'weaker' or who do not wish to engage in discussions, are then more likely to be seated in one of the outer positions where communication can be more difficult. The most obvious solution is to simply avoid such linear positions, so that all group members can face each other, but this might not be practical or even possible in certain situations such as some university lecture halls. In such situations it might be more beneficial to have students discuss in pairs to avoid 'outliers'.

The most interesting question still remains: if the students feel that the whole group participate more and everyone is more likely to present explanations with Peer Instruction, why do we not find a significant difference for both D2 and D3 students? Both the interviews and surveys show that students felt that they were more likely to participate in presenting arguments and explanations when they had the time to think before discussion. Although there was no question in the survey that specifically stated 'I contribute more during discussion with Peer Instruction', this seems likely to represent the experiences of the majority of students when we look at the survey results as a whole as well as comparing them to the results from the interview analysis. At least half of the interviewed students were present in the video analysis and we identified all dominance levels among the interviewed students. This indicates that the results from the interview analysis are not confined to one particular dominance level.

One possible explanation of the contradictions between students' experiences and observation is that there might have been an element of 'illusion of group productivity', an effect where group members value the productivity of the group higher than the measured outcome, as a result of the group producing more than a single group member would alone [12]. The increased collaboration and argumentation between D1 and D2 students might give the group members an illusion that the group as a whole, i.e. all group members, see an increase in participation and argumentation.

Furthermore, research has shown that group members often fail to differentiate between ideas presented by themselves and other group members, resulting in an overestimation of their own performance (Stroebe and Dielh 1994). The students in our study emphasised strongly the notion of being able to have an opinion of their own with Peer Instruction, i.e. an opinion that had not been 'coloured' by other students. By having thought more deeply about the questions, students are to a greater extent able to judge the quality of presented explanations. The feeling of having one's own opinion as well as being able to more confidently disagree or agree on an explanation, may increase the feeling of participation and contribution. Despite D3 students had very little of the argumentation time, they were still to a large degree active during discussion with commenting, for instance, by asking other students to clarify their explanations. In addition, students might have thought of the same ideas as those being presented in the group and thus have felt that they contributed to the group by agreeing on this explanation: 'yes, I thought exactly the same!', without actually having presented any ideas or explanations of their own.

## 5. CONCLUSION

We have investigated the effects of the initial thinking period with Peer Instruction with regards to student argumentation during group discussions. The data material consisted of surveys and interviews which were compared to video material of students, in groups of three, discussing SRS-quizzes during lectures. The video material supported students' claims of more fruitful discussions, where more arguments were presented, with an increase of 91% of argumentation time medians. However, although students claim the group as a whole contributed more when the initial thinking period was included, the video material indicated more productive discussion only between the most and second most dominant student (dominance based on share of argumentation time among the group) and not among all group members. We argued that this could be due to students sitting in a linear manner which could create a communication barrier towards one of the students. Despite this, students still felt that all members in the group were more inclined to participate with argumentation when they had time to think before discussion. We argued that this could be due to an illusion of group productivity, where the students either had a hard time differentiating between their own contributions and those of other students, or that, by seeing an increase in discussion between two students, feel an increase in discussion in the group as a whole, i.e. among all group members.


[1] Dufrense, R., Gerace, W., Leonard, W., Mestre, J., and Wenk, L. (1996) Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, **7**, 3–47.

[2] Crouch, A. and Mazur, E. (2001) Peer Instruction: Ten years of experience and results. *American Journal of Physics*, **69**, 970–977.

[3] Masikunas, G., Panayiotidis, A., and Burke, L. (2007) The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, **15**, 3–20.

[4] Hake, R. (1998) Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, **66**, 64–74.

[5] Smith, M., Wood, W., Krauter, K., and Knight, J. (2011) Combining peer discussion with instruction explanation increases student learning from in-class concept questions. *CBE-Life Sciences Education*, **10**, 55–63.

[6] James, M. (2006) The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, **74**, 689–691.

[7] Mazur, E. (1997) *Peer instruction: A user's manual*. Upper Saddle River, NJ, Prentice Hall.

[8] Nicol, D. and Boyle, J. (2003) Discussion in large classes: a comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, **28**, 457–473.

[9] Nielsen, K., Hansen-Nygård, G., and Stav, J. (2012) Investigating Peer Instruction; how the initial voting period affects students' experience of group discussion. *ISRN Education*, **2012**.

[10] Kaartinen, S. and Kumpulainen, K. (2002) Collaborative inquiry and the construction of explanations in the learning of science. *Learn. Instr.*, **12**, 189–212.

[11] James, M. (2008) What are they talking about? lessons learned from a study of Peer Instruction. *Astronomy Education Review*, **7**, 37–43.

[12] Stangor, C. (2004) *Social Groups in Action and Interaction*. Psychology Press. New York and Hove.

[13] Stroebe, W. and Diehl, M. (1994) Why groups are less effective than their members: On productivity losses in idea-generating groups. *European Review of Social Psychology*, **5**, 271–303.

[14] Nijstad, B. and Stroebe, W. (2006) How the group affects the mind: A cognitive model of idea generation in groups. *Personality and Social Psychology Review*, **10**, 186–213.

[15] Nijstad, B., Stroebe, W., and Lodeqijkx, H. (2003) Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, **39**, 531–548.

[16] Collaros, P. and Anderson, L. (1969) Effects of perceived expertness upon creativity of members of brainstorming groups. *Journal of Applied Psychology*, **53**, 159–163.

[17] Anderson, C. and Kilduff, G. (2009) Why do dominant personalities attain influences in face-to-face groups? the competence-signaling effects of trait dominance. *Journal of Personality and Social Psykology*, **96**, 491–503.

[18] Patterson, M., Kelly, C., Kondracki, B., and Wulf, L. (1979) Effects of seating arrangement on small-group behavior. *Social Psychology Quarterly*, **42**, 180–185.

[19] Steinzor, B. (1950) The spatial factor in face to face discussion groups. *Journal of Abnormal and Social Psychology*, **45**, 552–555.

[20] Hare, A. and Bales, R. (1963) Seating position and small group interaction. *Sociometry*, **26**, 480–486.

# Paper E

## Teaching with Student Response Systems: Teacher-centric Aspects that Can Negatively Affect Students' Experiences of Using SRS

Nielsen, K.L., G. Hansen-Nygård and J.B. Stav

Preprint 2012. Submitted to *Research in Learning Technology*.

# Teaching with Student Response Systems: Teacher-Centric Aspects that Can Negatively Affect Students' Experience of Using SRS

Kjetil L. Nielsen, Gabrielle Nygård Hansen, and John B. Stav
*Sør-Trøndelag University College, 7004 Trondheim, Norway*

In this article we describe and discuss the most significant teacher-centric aspects of Student Response Systems (SRS) which we have found to negatively affect students' experience of using SRS in their lectures. By doing so we hope to increase teachers' awareness of how they use SRS and how seemingly trivial choices or aspects when using SRS can have a significant negative impact on students' experiences, especially when these aspects are often repeated. We cover areas such as consistency when using SRS, time usage, preparation, the experience level of the teachers with regard to SRS, teacher commitment and attitudes, teacher explanation and how students fear voting results can mislead the teacher. The data is based on three years of experience in developing and using an online SRS in classroom lectures and consists of focused (semi-structured) student group interviews, student surveys and personal observations.

## 1. INTRODUCTION

An SRS can be described as an electronic voting system where students are presented with a multiple choice question, often as a quiz to which they will answer with a small handheld device (commonly referred to as 'clickers'), usually after engaging in peer discussions. Benefits from utilising such a system in lectures include: increased student motivation and engagement [1, 2]; easier clarification of misunderstandings [3]; promotion of active learning [4] and increased student performance [5, 6] including better conceptual understanding [7]. With such a promising record it is easy to forget that SRS is only a tool that, if not used correctly, can even be detrimental to the lectures [8]. Focusing primarily on the technology, with a belief that the technology will automatically improve lectures instead of focusing on how students think and learn, is the single most important reason for failure when implementing new technology into education [9] and SRS is no exception.

There are several publications that give best practice guidelines for using SRS (e.g. [10] in classroom lectures. Most publications on SRS focus on its benefits, often with guidelines for increased learning and effective SRS-use. Some publications also report on teaching aspects of SRS that have been shown to decrease student satisfaction with the system. These include: inefficient use of time with SRS [11, 12] by, for example, using too much time setting up the system and handing out clickers [1, 2] or having discussions drag on for too long [13]; responses from SRS being graded [11, 12]; grading of SRS questions resulting in mandatory attendance [14]; SRS used only to keep attendance [15]; SRS being used just for the sake of using it [16]; teachers having negative attitudes towards SRS [16]; irrelevant clicker questions [17] and lack of teacher explanation after the quiz [18].

Although this seems to be an elaborate list, negative aspects are mostly briefly mentioned and not described in depth in the literature. In-depth studies on these aspects and how they affect students' experience of SRS are, to the authors' knowledge, lacking. To fully understand and appreciate best practice guides on SRS, it is important to understand how and why different aspects of implementation can have a negative impact on the students. In this paper we describe and discuss such aspects after three years of experience (since 2009) in developing and using an online SRS for modern handheld devices, such as smartphones, at the Sør-Trøndelag University College (HiST) in Norway. We start this paper with background information, describing the classes where SRS was used and the implementation choices for the different years of testing. This is followed by a brief description of research methods and a presentation of our results. We conclude the article with a discussion and conclusions.

## 2. BACKGROUND

### 2.1. Description of the classes

Preparatory engineering classes at HiST consist of students who wish to start on an engineering degree, but who do not have the required subjects from senior high school. There are four classes, with approximately 50-70 students per class, and five different courses: physics, mathematics, social science, Norwegian and English. One of the classes, however, only has physics and mathematics. The courses last for a full year and the curriculum is close to that of the second and third year of senior high school, but some changes are made to reflect that the students are to start on an engineering degree.

### 2.2. Implementation

The lecture format with SRS during the testing was the traditional teacher-style lectures (using digital blackboards) normally used on the courses, as well as the teacher presenting the students with 1-4 quizzes during the session. The SRS-sequence could vary; for instance, by using the Peer Instruction method [19], which consists of an individual thinking period, followed by a vote, group discussions and revote. Other examples could be to exclude the first voting session and go straight into the discussion of the alternatives, or have the students discuss, but not show the alternatives until students are ready to vote. In none of these three years of testing did we have SRS responses affect students' grades in that SRS participation did not give extra credit, nor did answering correctly on quizzes. Below we give a brief overview of the implementation choices (technical choices not included).

*Semesters 2009-2010*

In the autumn semester of 2009, SRS was used in one class in physics, taught by one of the authors (KLN) for four weeks. In the spring semester, 2010, SRS was used for eight weeks in all four classes in physics. The eight weeks of testing were part of a study to compare different SRS-sequences [20]. We prepared around 50 quizzes which the teachers used for the duration of the study. All teachers used the same quizzes to minimize variables.

*Semesters 2010-2011*

SRS was implemented in two classes and three courses (physics, mathematics, and society studies) for the whole duration of the 2010-2011 semesters. Teachers were given no restrictions on the use of SRS in their courses, or on how much they wished to use it. Two teachers were assigned to make a pool of quizzes to be used in mathematics and physics. However, all teachers were free to ignore this pool and make their own questions. Before and during the semester, we had regular meetings and seminars with the teachers where they were given the opportunity to use SRS, discuss methodological choices and so on. One of the authors (GHN) acted as a contact person whom the teachers could present questions to about SRS or report bugs in the software.

*Semesters 2011-2012*

In the last year of testing, 2011-2012, SRS was made available to all classes and teachers on the preparatory engineering courses. Seminars were held to discuss methodological choices with SRS before the start of the semester, but there were no SRS meetings during the semester due to time constraints among the researchers.

## 3. METHOD

The data consists of student interviews and surveys from three years of SRS-testing. In total there were 13 focused (semi-structured) group interviews consisting of around four students per group. The interviews were conducted by one of the authors (GHN). During the first two years of testing, GHN was present as an observer in the lectures where SRS was used. There were nine interviews from the 2009-2010 period, two of which were conducted during the autumn semester of 2009, and the rest during the spring semester of 2010. Four interviews were conducted at the end of the spring semester of 2011. Surveys were conducted during all three years of testing. Many of the questions from the surveys were constructed based on information that emerged from the interviews. Most survey questions or statements used a 1-5 Likert scale ranging from, for instance, 'totally agree/very strong degree' to 'totally disagree/very little degree'.

The interviews from the different semesters were part of larger studies to investigate various aspects of SRS implementation. Interview guides were constructed to explore both students' positive and negative experiences with SRS, and later interview guides were modified based on results from previous interviews and feedback from teachers. Even though the different studies had different focuses, the interviews covered a broad spectre of themes, including various aspects on how the teachers used SRS. The analytical tools consisted for the most part of a three-step coding and analysing scheme adapted from Charmaz [21, 22], which is a method for identifying experiences, feelings and attitudes. The results shown in this article are in large measure an accumulation of the most significant negative aspects that emerged from the interview analysis.

## 4. RESULTS

Our students experience SRS in general as a very positive element in their lectures and a valuable tool for both the teacher and students. In the 2011-2012 survey, 95.6% of the students answered that they would have wanted SRS to be used in their future education if they had the chance (3.8% answered 'don't know' and 0.6% answered 'no', n = 160). Although the students are in general positive to SRS, there are several aspects that can negatively affect students' experiences of the system. Some of these problems are due to technical difficulties, such as voting devices not being able to connect to the service. This was a noticeable problem during the first year of testing and a source of frustration among the students. As technical difficulties became negligible, and teachers could use SRS as they saw fit, most criticism from students shifted towards the way SRS was used by different teachers, especially when they were in the position to compare different teachers.

The results are divided into six different sections: 1) *Consistency when using SRS*, 2) *Difference in teachers' SRS experience level*, 3) *Time usage*, 4) *Commitment: a two-way street*, 5) *Preparation of the questions* and 6) *Voting results: a false image of understanding.* The last section also contains two subsections on the importance students place on the *teacher's explanation* and *including 'don't know' as an alternative.*

### 4.1. Consistency when using SRS

Before starting to use SRS in their lecture, teachers should have a clear goal of why they want to use SRS and be consistent in their choice. Inconsistency when using SRS gives students the impression that the teacher does not know why (s)he wants to use it and more or less uses SRS for the sake of using it. The students value using SRS to receive instant feedback, and give feedback to the teacher, or to have a pause in the middle of the lectures where they can engage in peer discussions. Inconsistency, such as only using SRS now and then, creates an uncertainty among the students in not knowing how long they have to wait for these benefits and thus creates a source of irritation and frustration. According to the students, inconsistency is also a sign of low teacher commitment (more on teacher commitment below). Students emphasise that if a teacher is going to use SRS, (s)he should use it regularly or not at all. The following student quotes are on why they did not value SRS in a specific course:

***Student 1:*** Because it is so rarely used that... it should have been more... that he either used it more or not at all.

***Student 2:*** I feel that it is used so little that there's like no point.

***Student 3:*** Yes, I agree with you on that. I feel that I don't get much out of it [using SRS] in that course because it is used so little.

If SRS is rarely used, it will also feel less integrated as a natural part of the lectures. As well as being expressed in the interviews, it is also reflected in the surveys from the 2011-2012 classes, table 1, question/statement (QS) 6, where the majority of students agree with this statement. The teacher will not receive enough practice using SRS which in turn results in more fumbling with the software and SRS use feels less 'smooth' to the students.

### 4.2. Difference in teachers' SRS experience level

The difference in experience level of the teachers is also apparent from the survey. Some questions on the survey were organised so that students could give different answers based on different courses. We separated the responses into a group consisting of teachers with previous experience with SRS and teachers who used SRS for the first time during the 2011-2012 period. Teachers with previous experience used SRS to a greater extent (QS4) and SRS was also regarded as a more natural part of their lectures (QS1). Students also felt that SRS helped them learn the curriculum better when it was used by experienced SRS-teachers (QS3). A Mann-Whitney U showed a statistical significant difference between teachers with and without previous SRS-experience for QS1, QS3 and QS4 ($p < 0.001$ in all cases). There was no statistical significant difference ($p > 0.05$) between the experience levels of SRS on students feeling that SRS compromised the lectures (QS2): in both cases the majority of students did not regard SRS as compromising.

### 4.3. Time usage

Having SRS as a natural and integrated part of the lecture is important as students are wary of how teachers distribute time during lectures, especially regarding SRS. This includes all aspects of SRS, such as how long it takes to set up the system, present the quiz or even how long the students have to vote. During the voting session there is a 30-second timer during which students can deliver a vote. Several students complained about 30 seconds being too long because 'it doesn't take more than a few seconds to vote', especially when they have already made up their mind, showing that even a few seconds of 'dead time' can have a negative impact. Distribution of time is especially important during discussions. The students emphasise that relevant peer discussion is usually only maintained until they come to an agreement about the correct answer. If the discussion time drags on after a consensus is reached, it increases impatience and irritation among the students.

***Student 1:*** It gets boring when it takes too long.

***Student 2:*** Yes, it should have been more effective [in a particular course].

***Student 1:*** It is much worse to sit doing nothing rather than not having enough time to finish.

What constitutes as proper use of time during SRS-quizzes, greatly depends on the nature of the questions. Factual questions can be beneficial for quick repetition, but are otherwise regarded as inefficient use of time with little to no learning gains. According to the students, with such questions they either know the answer or they do not. Consequently, including an individual thinking period before group discussion, which students' normally regard as an important factor for good discussions, has little value since there are no reasoning skills involved other than knowing where to look for the answer in the textbook. The discussions themselves also have little learning gains for the same reason. The students therefore emphasise conceptual questions, where the answer

TABLE I: A selection of answers from the 2011-2012 survey (160 responses). The table shows the mean value and standard deviation (in parenthesis) for the responses. 5 = totally agree/very strong degree etc., 1= totally disagree/very small degree etc. *'Experience with SRS' translates to teachers who used SRS in at least one of the previous years of testing, while 'not experienced' translates to teachers using SRS for the first time in the 2011-2012 classes.

| Question/Statement | Teacher experienced with SRS* | Teachers not experienced with SRS* |
|---|---|---|
| 1. To what degree do you feel SRS is a natural and integrated part of the lecture? | 4.24(0.77) | 3.34(1.19) |
| 2. I feel SRS compromises the lectures. | 1.62(0.89) | 1.60(0.84) |
| 3. To what degree does SRS help you learn the subject matter? | 3.64(0.90) | 3.19(1.12) |
| 4. Do you feel SRS is used too much/too little in the lectures? | 2.90(0.52) | 2.39(0.80) |
| 5. How important is the commitment of the teacher with regards to your experience of SRS? | 4.21(0.82) | |
| 6. If SRS is rarely used it will also feel little integrated with the lectures. | 3.62(0.75) | |

is not obvious and a deeper level of reflection is required (making the initial thinking period more valuable), as more beneficial to their learning and therefore a better use of SRS-time.

### 4.4. Commitment: a two-way street

The teachers who are regarded as the most proficient in the use of SRS are also those who use it most frequently, are well prepared and have a clear purpose for using it. SRS is then regarded as a tool in their lectures. It is also these teachers whom students perceive as most committed and enthusiastic. Commitment and attitudes are extremely contagious; if the teacher shows low commitment and enthusiasm towards using SRS, for example, by regarding it as a chore or by being insufficiently prepared, SRS is experienced as more or less meaningless. Commitment is explained by the students as a 'two-way street', in that SRS is dependent on two parts: the teacher and the students. These two are dependent on each other in order for a lecture with SRS to function properly. If the students sense that the teacher does not care or does not put an effort into making SRS work, their motivation towards participation during SRS-quizzes decreases. Students find it motivating to see teachers using SRS actively as a tool to improve the quality of their lectures.

***Student 1:*** The way [teacher] uses SRS, it makes me feel that it is something that [teacher] really wants to use to become better. Yes, both as a teacher and for us to understand it better. [Teacher] uses it really in a way that is supposed to benefit us.

***Student 2:*** Yes, and [teacher] seems very committed and seems very certain of... yes like he said, prepares herself more. And then you as a student come on the same level as a teacher, really, in the way that when you see that someone who is supposed to teach you something, prepares for the lectures, then it has to be much more fun to be a teacher. And for you as a student, it goes both ways as well. If you as a student see that the teacher comes unprepared, then you lose a lot of motivation around it, especially such things as SRS. If you see that the teacher doesn't care about it or hasn't prepared enough for it, for instance, if there are a lot of mistakes in the questions and such.

***Student 3:*** Plus that it is very motivating to see that [teacher] uses SRS as a tool and not something that she has to use. That is really good.

One student summarised his view of the importance of teacher commitment and how it affects students by drawing a parallel to a working environment, and the authority of the supervisors:

I think that most of us regard this [the preparatory engineering course] as our workplace. We have struggled and worked for this and then we also take it very seriously. And of course at such a workplace, the attitudes of your colleagues and bosses spread to the other employees. And it is like this here as well. So if you picture the teacher as your boss. And then they are, at least in my eyes, looked upon as an authority who we are supposed to look up to and learn from. So if they then, maybe without really meaning it, give us a feeling that they don't really care about learning, or that they don't care about the result of their use of the SRS, then this will spread to us.

### 4.5. Preparation of the questions

There are several factors that give students an impression of low teacher commitment, insufficient preparation being one of them. Perhaps the most important aspect of SRS preparation is the actual quizzes: an SRS-question can pique students' interest in the subject matter and motivate them to engage in peer discussion, but it can also be a source of irritation and frustration. According to the students, the requirements for having a good SRS-question are higher than other questions or problems. As one student put it:

The questions have to be done properly; it is much easier to regard an SRS-question as ridiculous in comparison to other problems.

Questions have to be implemented as a natural part of the lecture. Even if teachers have not made the questions themselves, but have availed of those in a question bank or from literature, they still have to make them their own. Students familiarise themselves with the teacher's and textbook's way of explaining and therefore it can cause confusion if they are suddenly presented with a question with a very unfamiliar presentation, i.e. different wording or use of illustrations. The confusion can be further increased if the teachers fail to properly explain the context of the quiz because of lack of preparation. However, this should not be confused with presenting students with questions with an unfamiliar context or setting, which can be an important part of conceptual quizzes.

Also, students emphasise that there has to be a rationale or motivation behind the questions. In other words, the quiz has to have a clear purpose. Is the question testing newly acquired theory or is it checking if the students remember yesterday's lecture? Moreover, the questions themselves have to be clear. What is actually being asked? If the questions are not clear, students use a lot of time and energy trying to interpret the questions rather than working towards finding the solution. This can be a major source of frustration.

***Student 1:*** Yes, like yesterday... lately there have been much of... so much interpretation of the questions that I don't even want to bother. I want pure... I mean, I want questions that I can read and understand immediately.

***Student 2:*** Yes, that it is clear... a clear question that has an answer, not where you have to sit and try to figure out what the questions are asking you. You are supposed to mull over the solution, not the question in itself!

### 4.6. Voting results: a false image of understanding

One of the benefits of SRS is the opportunity for the teacher to see the level of understanding among the students and use this information to tailor the lectures. However, there are several factors that affect students' decision-making during a discussion. One such factor is opinions from 'strong' students (i.e. A-students), which are valued higher than those of lesser skills. If students feel uncertain, it is very easy to agree on a solution presented by stronger students. This can result in the rest of the group accepting explanations whether they have understood them or not, and the voting results are thus not a correct representation of the level of understanding among the students.

*Teacher explanation*

Consequently, the students emphasise the importance of the teacher explaining all alternatives, both why the correct alternative is correct and why the incorrect is wrong, even though the majority of students have answered the correct alternative. The teacher's explanation is in fact regarded by the students as the most important aspect of SRS with regards to their learning. This was shown on the 2011-2012 survey where 62% of students answered 'teacher's explanation' on this question (the other alternatives being 'group discussion before voting' (7%), 'experiencing increased engagement and motivation' (9%), 'the immediate feedback from the voting results' (14%), and 'other' (8%), N = 160).

*Including 'don't know' as an alternative*

Students emphasise that they do not want to guess blindly or vote on an alternative they do not understand as they fear this can mislead the teacher and result in an insufficient explanation of the solution. Therefore, there is one aspect of the questions that has been repeatedly brought up: the inclusion of an alternative called 'don't know'. The students regard it as a consolation that the teacher can see that there is a large number of the students who are uncertain. Having 'don't know' as an alternative in the questions is regarded as a tool with which they can 'push' the teacher into giving more in-depth explanations following the discussions. For some students having this tool is regarded as such an important factor that they will refrain from answering if 'don't know' does not feature among the alternatives.

***Student 1:*** I think that people are more inclined to avoid voting if 'don't know' is not present.

***Student 2:*** Yes, we often see that there is a difficult question when there are a lower number of people voting.

***Student 1:*** It is natural that it becomes like this when we don't have the opportunity to answer 'don't know'. We don't want to answer just for the sake of answering, what is the point of that?

## 5. DISCUSSIONS AND CONCLUSIONS

We have presented the most significant teacher-centric aspects which we have found negatively affect students' experience of SRS after three years of testing at the Sør-Trøndelag University College in Norway. The preparatory engineering course students have a tight time schedule, having to take approximately two years of senior high school curriculum in one year, resulting in their being extremely focused on what kind of activities lead to effective learning. While students can easily skip exercises from the textbook, they do not have the same opportunity with SRS-quizzes. When a quiz is presented, there is a certain amount of the lecture time used on this quiz irrespective of the quality of the questions. The requirements for the SRS-questions to be positively received, compared to traditional textbook problems, probably increase for this reason.

The tight time schedule also results in any activity that is ineffective or where unnecessary time is spent with SRS, for instance, by fumbling with the software, being prone to cause irritation. Effective use of SRS does not, however, translate to 'using less time on each quiz'. During the first year of testing we compared the Peer Instruction method with and without the initial thinking period and voting session. Even though including this initial session results in more time spent on each quiz, 70% of the students (N=109) still preferred having this initial thinking period (with 19% preferring not having it) since this would increase the quality of the discussions.

Students emphasise teacher commitment as the most important factor for successful SRS-implementation. Teacher commitment does not directly translate to teachers being enthusiastic when using SRS, though this is a major benefit. Poor preparation, inconsistent use of SRS, not having a clear goal of SRS use, and bad question design, all give the students an impression of low teacher commitment. This can result in a 'vicious circle' where motivation and participation decreases among the students, which, in turn, further decreases the teacher's motivation for using SRS during class - the two-way street.

Teachers with more SRS-experience, either by having previous experience with SRS or by simply using it more, were rated higher by the students in our study. However, although the 2011-2012 survey showed those students found SRS as a less natural and integrated part of the lectures with inexperienced teachers, they did not regard SRS as compromising to the lectures. This may seem contradictory to the attitudes of 'use it right or not at all' which were prominent during the interviews of the previous year. A possible explanation for this result could be that SRS was used to a lesser extent in these courses (QS4 in table 1) and as a result not used enough to compromise the lectures. Also, most students who answered the survey had participated in courses with and without teachers with previous SRS-experience. Since the students have seen the benefits of using SRS 'the right way', it is possible that this has influenced the survey results from the courses with inexperienced teachers.

The difference between teachers with and without previous SRS-experience may have been amplified by the omission of the forum used in the previous year. During the 2010-2011 testing, teachers had regular meetings with one or more of the researchers where all aspects of SRS could be discussed and teachers shared experiences in using it in their lectures. The new SRS-teachers of 2011-2012 did not have this opportunity, due to time constraints, and consequently did not have the same level of guidance as the previous year. This could have hampered the improvement of their SRS-skills, both on a pedagogical and technical level. In addition, lack of technical support could be a factor that contributes towards decreased teacher commitment [10]. We received feedback from experienced teachers that they missed having the forums of the previous year and that it was difficult to follow up inexperienced teachers in an effective manner without such an arena for discussion.

During the testing periods, some teachers reported a noticeable amount of students not voting during quizzes. One possibility could be that this was a result of technical difficulties, but the students expressed omission of 'don't know' as an alternative as a major reason for less participation for fear of misleading the teacher by guessing. There was some discussion among the researchers (and to some extent the teachers) on whether such an alternative should be included. Although students have high regard for such an alternative, some of the researchers feared that including 'don't know' would give students an 'easy way out'. Being uncertain about the answer is an important aspect of conceptual SRS-questions. The answer should not be obvious and students have to challenge themselves to use their knowledge about the curriculum in an unfamiliar setting, a task that can be daunting for those students who often rely on memorisation of familiar situations. Including the 'don't know' alternative could result in students not making enough effort to challenge themselves to find the solution since they have the option to 'give up'.

Another possible scenario, however, could be that, when presented with a difficult quiz, the omission of 'don't know' decreases students' motivation to challenge themselves. The more difficult the question, the higher the probability of students not finding the correct solution or at least not understanding it. Since the students do not wish to guess, they might refrain from trying to find the solution because of the high probability of their not voting at that session and thereby seeing the effort of working out a solution as a waste. The act of giving a vote can be a major motivation factor for participation during SRS as it gives both the participation and group discussion a meaning [23], but much of this motivation factor can be lost if students know they are likely to refrain from voting.

Having 'don't know' as an alternative is briefly discussed in the study by Draper and Brown [1] where they reported that few students chose 'don't know'. The researchers argued that the anonymity of answering with SRS seemed to make students more inclined to choose a definite answer despite being very uncertain, which is contradictory to the reports of the students in our study. The effects of including 'don't know' as an alternative are not clear and future research could entail comparison of learning gains (or students' experiences) from two groups where the only difference in the questions is the inclusion or omission of 'don't know'.

The problem of participation is most likely not an issue if the students are rewarded for their participation in the form of points on their grade. Participation in

SRS-quizzes at HiST during the testing periods did not have any effects on students' grades, and thus there was no 'punishment' for not participating. However, as we mentioned initially in this article, having SRS participation graded can result in less student satisfaction [11, 12]. There have even been results indicating that having extra points for correct quiz answers can lower the quality of the group discussions [24] as opposed to only being given points for participation. Some commercial SRS brands give students the opportunity to follow up their vote with a rate of confidence level [25]. Such an option could negate the necessity for a 'don't know' alternative as the students have the opportunity to state their uncertainty.


[1] Draper, S. and Brown, M. (2004) Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, **20**, 81–94.

[2] Masikunas, G., Panayiotidis, A., and Burke, L. (2007) The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, **15**, 3–20.

[3] Rice, R. and Bunz, U. (2006) Evaluating a wireless course feedback system: The role of demographics, expertise, fluency, competence, and usage. *Studies in Media and Information Literacy Education*, **6**, 1–32.

[4] Boyle, J. and Nicol, D. (2003) Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal*, **11**, 43–57.

[5] Hake, R. (1998) Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, **66**, 64–74.

[6] Rao, S. and DiCarlo, S. (2000) Peer Instruction improves performance on quizzes. *Advances in Physiology Education*, **24**, 51–55.

[7] Crouch, A. and Mazur, E. (2001) Peer Instruction: Ten years of experience and results. *American Journal of Physics*, **69**, 970–977.

[8] Barber, M. and Njus, D. (2007) Clicker evolution: seeking intelligent design. *CBE Life Sciences Education*, **6**, 1–8.

[9] Mayer, R. (2005) *The Cambridge handbook of multimedia learning*. Cambridge University Press.

[10] Caldwell, J. (2007) Clickers in the large classroom: current research and best-practice tips. *CBE Life Sci. Educ.*, **7**, 9–20.

[11] Reay, N., Li, P., and Bao, L. (2008) Testing a new voting machine question methodology. *American Journal of Physics*, **76**, 171–178.

[12] Knight, J. and Wood, W. (2005) Teaching more by lecturing less. *Cell Biology Education*, **4**, 298–310.

[13] Dufrense, R., Gerace, W., Leonard, W., Mestre, J., and Wenk, L. (1996) Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, **7**, 3–47.

[14] Graham, C., Tripp, T., Seawright, L., and Joeckel, G. (2007) Empowering or compelling reluctant participators using audience response systems. *Active Learning in Higher Education*, **8**, 233.

[15] University of Wisconsin. *Student Response Systems.* `http://www4.uwm.edu/ltc/srs/faculty/best_practices.cfm` (last checked 15 July 2012).

[16] Draper, S. (2002), Ensuring effective use of prs: results of the evaluation of the use of prs in Glasgow University, October 2001- June 2002. `www.psy.gla.ac.uk/~steve/ilig/papers/eval.pdf` (accessed 9 January 2012).

[17] Barnett, J. (2006) Implementation of personal response units in very large lecture classes: student perceptions. *Aust. J. Educ. Techn.*, **22**, 474–494.

[18] Smith, M., Wood, W., Krauter, K., and Knight, J. (2011) Combining peer discussion with instruction explanation increases student learning from in-class concept questions. *CBE-Life Sciences Education*, **10**, 55–63.

[19] Mazur, E. (1997) *Peer Instruction: A user's manual*. Upper Saddle River, NJ, Prentice Hall.

[20] Nielsen, K., Hansen-Nygård, G., and Stav, J. (2012) Investigating Peer Instruction: how the initial voting period affects students' experience of group discussion. *ISRN Education*, **2012**.

[21] Charmaz, K. (2001) *Rethinking methods in psychology*, chap. Grounded Theory, pp. 27–49. London: Sage Publications.

[22] Charmaz, K. (2003) *Qualitative Psychology; A practical guide to research methods*, chap. Grounded Theory, pp. 81–110. London: Sage Publications.

[23] Hansen, G. (2008) *'Men pausen tilhørte liksom faget'. Studenters opplevelse av bruken av studentresponssystem på forelesningen*. Master's thesis, Norwegian University of Science and Technology.

[24] James, M. (2006) The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, **74**, 689–691.

[25] Bruff, D. (2012), Clickers. `http://cft.vanderbilt.edu/teaching-guides/technology/clickers/` (last checked 15 July 2012).

# Paper F

## Student Response Systems in Physics Lectures: Does Voting Results Represent a Correct Image of Student Understanding?

Nielsen, K.L., and J.B. Stav

Preprint 2012. Submitted to *Research in Science & Technological Education*.

# Student Response Systems in Physics Lectures: Do Voting Results Represent a Correct Image of Student Understanding?

Kjetil Liestøl Nielsen and John Birger Stav
*Sør-Trøndelag University College, 7004 Trondheim, Norway*

***Background:*** Student Response Systems (SRS) or 'clickers' are often used with quizzes during lectures to identify misunderstandings and assess students' understanding in order to tailor the lectures accordingly. However, the reliability of SRS as an assessment tool is rarely addressed in research on SRS.

***Purpose:*** We have investigated the group discussions of two quizzes in physics which have different contexts and require different levels of knowledge, despite both having solutions being dependent on very similar aspects. In addition to identifying conceptual misunderstandings about diffraction and refraction, the aim of the paper is to highlight aspects of using SRS as a tool for assessing student understanding and addressing misunderstandings.

***Sample:*** The sample consisted of video clips of 21 small student groups during the first quiz and 22 during the second, a total of 69 different students in four different classes at a preparatory engineering course in physics at a University College in Norway.

***Design and methods:*** The paper includes a qualitative analysis of group discussions. Each utterance was coded, and the ideas and arguments presented were categorised. Additionally, voting results from the two quizzes are presented in order to evaluate the impact of the arguments presented.

***Results and conclusions:*** Arguments presented by students revealed flawed understanding of physical principles during both quizzes. While this led to the majority of students voting incorrectly during the second quiz, 90% of the students voted correctly during the first quiz despite their arguments showing conceptual misunderstandings of the underlying physics. Consequently, the findings illustrate an example of how voting results with SRS can provide misleading feedback to both the teacher and students on students' level of understanding. The findings also emphasise the importance of teachers having a critical awareness of limitations with the quizzes presented if SRS is to be used as an assessment tool.

## 1. INTRODUCTION

Physics can be a very challenging subject for novice engineering students. They are expected to learn to reason qualitatively and logically about physics, but students often end up relying on memorisation and equation hunting [1]. Students are rarely given time to reflect upon the subject matter in traditional lecture formats [2] and it can be a challenge for teachers to assess if the students have understood what is being taught [3]. Using a Student Response System (SRS) in class is a popular way of dealing with these challenges [4]. Teachers can present students with multiple choice questions to which they answer with a small voting device, often after engaging in peer discussion [5]. Students' knowledge of physics can be challenged with SRS by teachers presenting students with conceptual problems where students cannot rely on mere memorisation, but have to use their knowledge about the subject matter (Mazur 1997). Thus, SRS is used to give both the students and teacher feedback about the level of understanding among the students [6] by revealing misunderstandings and gaps in students' knowledge [7, 8]. Using such systems has been shown to increase both student motivation and engagement (e.g. [2, 9]) as well as increased student performance (e.g. [10, 11].

One of the key features of using SRS in class, is the ability to use the feedback from voting results to tailor the lectures [3], for instance by using less time on topics where a vast majority of students answer correctly on quizzes [12]. This raises an interesting question: do voting results from SRS represent a correct image of student understanding? There have been studies that have shown, for instance, that students might accept explanations presented by dominant students without fully understanding them [13]. Even if students feel that they have understood the solution, and the vast majority of students vote correctly on a quiz, does this mean that there are few to no flaws on their understanding of the solution?

The purpose of this paper is to investigate the group discussions of two quizzes which have different contexts and require different levels of knowledge about physics, despite both having solutions being dependent on very similar aspects. This way we aim to increase knowledge about using SRS as a tool for revealing misunderstandings and assessing student understanding, as well as identifying some conceptual misunderstandings and confusions about physics, in this case diffraction and refraction. The paper starts with background information, where the investigated quizzes are described as well as research methods. This is followed by the presentation of our results, a categorisation of students' arguments during the group discussion. These results are discussed with two different focuses: first, we discuss students' misconceptions and misunderstandings about the subject matter and how these might have affected students' decision making; last, we discuss our results with regard to implications for teaching with SRS.

## 2. BACKGROUND

This paper consists of an analysis of students discussing two multiple choice quizzes, in groups of three, presented during the same physics lecture in four different preparatory engineering classes at Sør-Trøndelag University College (HiST) in Norway. These group discussions were part of a larger study to investigate how the inclusion of the individual thinking period before discussion, a method commonly referred to as 'Peer Instruction' [14], affected students' argumentation during the discussions [15]. Consequently, the data consist of groups which had and groups which did not have the time to reflect before discussion.

During the study we could examine the voting results from the different quizzes. Two quiz questions piqued our interest in particular: despite the answer of both questions in the end being dependent on the same aspect of the same equation, a vast majority of the students were able to find the correct answer in the first quiz, but were more divided in the second. An additional factor was that both questions were presented during the same lecture, which inspired further investigation of the ideas and arguments presented by the students in these discussions.

### 2.1. The quizzes

The preparatory courses include a curriculum close to that of senior high school. The optics part of the curriculum start with the geometrical model of refraction: the change in direction when waves enter new mediums. This is followed with a description of the basic nature of waves before introducing the concept of interference and diffraction: how waves can 'bend' or 'spread out' when passing small obstacles or narrow slits, respectively. The quizzes, shown in Figure 1, deal with interference of light that passes through small slits. The first quiz (Q1) requires students to identify whether the red or the yellow light has the longest wavelength and how this affects the interference pattern. This can be determined by using the grating equation, defined in the textbook as:

$$d \sin\theta_m = m\lambda, \tag{1}$$

where $d$ is the distance between the slits, $\theta_m$ is the angle between the centre and $m$'th maxima, and $\lambda$ is the wavelength of the light. Since red light has a longer wavelength than yellow, the distances between centre and the maxima will be larger (alternative D). In the centre there will be an orange spot due to the central maximum ($m = 0$) being independent of wavelength (red + yellow = orange). The second quiz (Q2) describes a double slit experiment, but where the whole apparatus is submerged in water. In addition to using the grating equation, students have to find out how water affects the wavelength. The wavelength in water is determined by the refractive index, $n_{water}$:

$$\lambda_{water} = \frac{\lambda}{n_{water}} \tag{2}$$

Since the refractive index in water, $n_{water}$, is equal to 1.33, the wavelength will be shorter than that of air/vacuum, and the distance between the interference maxima will decrease (alternative B).

## 3. RESEARCH METHOD

Students discussing the quizzes were captured on video and audio. Prior to filming, students signed permission waivers which also explained that the video material was going to be used for research purposes. We encouraged students to sit in groups of three, but otherwise did not interfere with how students formed their groups during discussion. There were 21 groups captured during the first quiz and 22 during the second. Apart from one group during the first quiz and two during the second, the groups were identical during both quizzes. In order to find a possible explanation for the differences in voting results in the two quizzes, we wanted to investigate and categorise the conceptual content of students' argumentation.

In the comparison study described earlier, the video clips were subjected to a coding process where each utterance was placed in pre-determined categories adapted from Kaartinen and Kumpalainen [16] and James [17]. The categories were divided into two main sets: *comments* and *argumentation*. The latter was defined as utterances that included conceptual information which was presented with the intention of being related to the solution of the question, i.e., utterances that included an idea, explanation or simply an argument. It included three sub-categories. The first described the purpose of the argument, for instance, justifying an opinion or contradicting a presented argument. The second described the extent of the idea being new, an expansion or rephrasing of an idea that had previously been presented or simply a repetition. The last sub-category described the nature

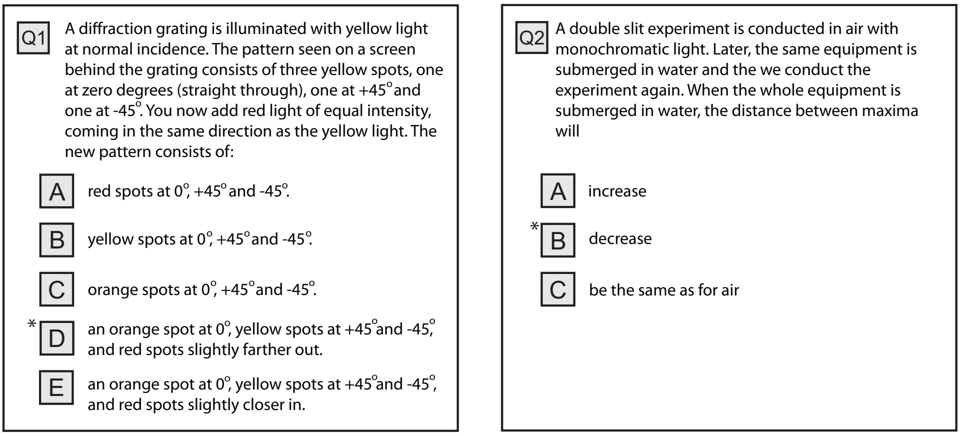



FIG. 1: The investigated quizzes. The quizzes were presented to the students in Norwegian (the word 'diffraction' was not present in the Norwegian translation). Quiz 1 (Q1) is adapted from Mazur [14]. * The correct alternative.

of the argument, for instance if it included formal scientific language, more casual language or if it was mostly based on gestures such as hand movements. The comments category included sub-categories such as 'asking for opinion', 'agreeing/disagreeing', 'stating answer preference', 'reading the question out loud', and so on.

Although the initial coding scheme did not include the actual conceptual content of the arguments, it simplified locating utterances that included such information. Each utterance which included an argumentation-code was recoded with in-vivo codes based on the conceptual content of the argument. In-vivo codes, i.e., codes based on students' own words, were used to minimise the risk of misinterpretation of students' arguments at this early stage of analysis. The codes were reviewed and categorised based on the nature of the content. An example of such categories is 'blending of the lights' which includes arguments for the lights being 'mixed' or 'blended' to explain why there is orange light in the middle in the first quiz.

The next step was to refine the categories to make sure they were a correct representation of students' arguments during the discussions. Each category was summarised (the number of utterances included) and described in more detail, including possible connections to other categories. In addition to the conceptual codes, we also coded larger sections of each discussion based on the discourse codes detailed earlier in order to describe each discussion as a whole with a smaller set of codes. These codes were used to write a short summary of each group discussion, emphasising both the conceptual information presented as well as the dynamics of students' argumentation, for instance, how presented arguments were received and followed up in the group discussions.

These summaries had two purposes. The first was to simplify identification of connections between the categories. Through evaluation of the summaries, for example, we could identify that some categories were used in combination with other categories or as an introduction leading to other categories, which could result in their being merged into a more general category. Second, in combination with the voting results and number of utterances, we could easier evaluate the impact or 'importance' of each category. Larger categories were more likely candidates for the final set of categories and were thus often used as a starting point. Though an iterative process, constantly reviewing and redefining, splitting and/or merging categories, we ended up with a set of general categories representing students' arguments during discussion of the two quizzes.

## 4. RESULTS

The discussions during the first quiz were usually divided between arguments for red light having a wider interference pattern and explanations for there being orange light in the middle. While some students used *the grating equation* to justify the interference pattern, a more common argument was the notion of *more refraction with red light*, despite refraction not being relevant in this quiz. The orange light in the middle was explained by a *blending of the lights*. There were also a few students who argued for there being *no difference in interference pattern*, resulting in a small category. The discussions would often be heavily focused on one of the categories described above, but they were not mutually exclusive with the most potentially inclusive categories being the argumentations of wider interference pattern and explanations of orange light in the middle.

Although students provided several different arguments during the second quiz, they can be generally separated into two different categories. In the largest category, *no refraction - no difference*, students argue that there is no difference to the interference pattern since the whole apparatus is submerged in water resulting in no refraction. In the second category, *maintaining equality of the equations*, students manipulate equations to find out the speed of light in water or how the refractive index of water affects the wavelength and thus how they affect the angles to the maxima.

In the description of the categories we present quotes from students during discussion as examples, where *stud m*, *stud l* and *stud r* indicate a student in the middle of a group, a student to the left or a student to the right respectively.

### 4.1. Quiz 1

Students were quick to identify the wavelength as being vital to the solution, and the majority also successfully identified the longer wavelength resulting in a wider interference pattern for red light. This is also indicated in Figure 2, where 90% of students voted for the correct alternative (alt. D). Figure 2 also shows the voting results from the classes using Peer Instruction (right histogram), where we see an increased consensus after discussion. While some students would simply state a wider interference pattern for red light, without further explanations other than the longer wavelength, others would be more specific by providing a connection between wavelength and the width of the interference pattern.

*More refraction with red light*

A few students would use the more correct argument of longer wavelength resulting in light being bent more (note that the Norwegian textbook uses the word 'bending' instead of 'diffraction') through the same opening, but a far more common argument was to argue that red light is subjected to more refraction resulting in red dots appearing farther out than the yellow. With some exceptions where students referred to a 'rainbow effect', or directly referred to the refractive index, refractive angle or the chapter in the textbook about refraction, arguments of refraction during the first quiz were not usually explained or elaborated on other than simply stating that red light refracts more.

The red light has to have more refraction, right? More wavelength, more refraction.

While longer wavelengths result in a wider pattern, there was also a broad consensus that some light always goes 'straight forward' irrespective of wavelength. This was commonly presented as a fact without further justification. However, the example below shows a group more specifically differentiating between refraction 'occurring at the angles' while not at zero degrees. The example also shows how a student, while initially a bit unsure about the explanation, gradually becomes convinced.

***Stud r:*** They [the red and yellow light] only blend at zero degrees, but not at 45 degrees?

***Stud m:*** No, because there they will refract.

***Stud l:*** They refract differently. You get a rainbow effect.

***Stud m:*** There will be no refraction at zero degrees. There they will meet, you can say.

***Stud r:*** [repeats after stud m] There they will meet... Yes.

***Stud m:*** The red light has a higher refraction than the yellow.

***Stud r:*** So that the first dot with red light starts farther out than the yellow dot?

***Stud m:*** Yes.

***Stud r:*** [nods] Ok, I agree.

*Blending the lights*

The quotes from the example above show the main, and practically only, argument for why there will be orange light at zero degrees. Since wavelength does not matter at zero degrees, both light waves hit the same spot and 'blend' together to form orange. Although most students simply used the word 'blends' without elaboration, some students tried to explore why. For instance, one group argued that the two lights interfere to form orange. The student quotes below show an example of a group exploring different explanations of how the light blends.

***Stud r:*** Two different lights hit the same spot.

***Stud m:*** Yes.

***Stud l:*** So it is anyway... I was thinking... I was mulling over, it has to be D or E, they [the red and yellow light] have to be separated between orange. I'm thinking like... two different wavelengths meet, then it becomes [gestures a wavelike motion], I was thinking that it automatically becomes orange. Like when you are painting, you know? When you blend red and yellow, it becomes orange, but I'm not sure if the same applies here.

***Stud m and r:*** Yes.

***Stud r:*** It will become blended when it comes in the middle.

***Stud l:*** Yes.

***Stud m:*** Yes, it is only there that they blend.

***Stud r:*** Yes.

***Stud l:*** It has to be like that. [Pause] Maybe you see both the red and yellow light at the same time, but they are so close that...

***Stud r:*** Yes, that it is the same as with white light, you see all colours at the same time.

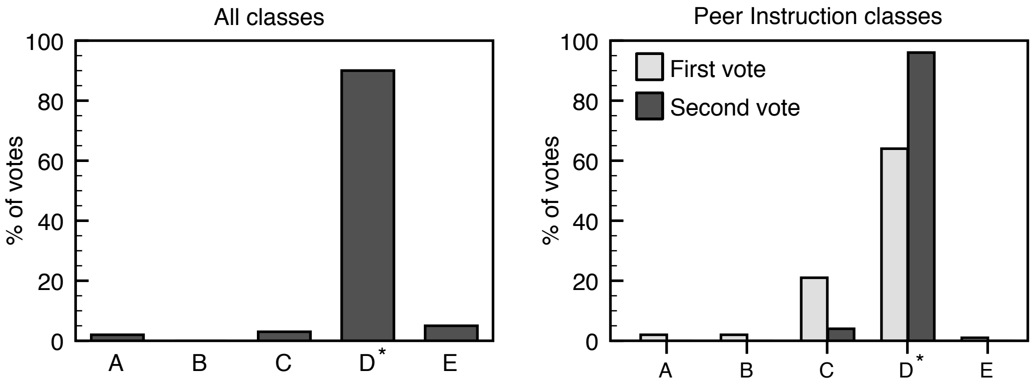



FIG. 2: Voting results from the first quiz. The histogram to the left shows the collected N=114 responses from all four classes while the histogram to the right shows the collected responses from the two classes using Peer Instruction (N=47 for the first vote and N=52 for the second). * The correct alternative.

*The grating equation*

While the use of refraction represented the more physical (and most common) explanation of the connection between wavelength and wider interference pattern, some students concentrated on a mathematical explanation by examining the grating equation and how different wavelengths affected the angles from the slits to the interference maxima. Higher wavelengths result in a larger sine or 'a larger number' resulting in an increased angle. The example below shows how a student used the mathematical explanation to both explain the solution and contradict another student's argument.

***Stud m:*** I voted B. I'm thinking that the yellow light has more energy than the red, and then it overrides the red.

***Stud l:*** So then it overrides the other?

***Stud m:*** Yes, that is how I saw it.

***Stud l:*** But if we look at the formula, then we see that the yellow light has [a wavelength of] 570-590 [nanometers] and the red light has 690-800, right? So if you use the formula, they get different angles. If the opening of the slits is the same [inaudible] if they are different. So, then according to the formula the red dots will be farther out than the yellow dots. They will not hit the same spot.

*No difference in interference pattern*

The contradicted argument shown above introduces a small category representing arguments of there being no differences to the interference pattern. The arguments either consisted of one of the colours 'overriding' the other, as seen in the example above, or that there is no difference in the width of interference pattern resulting in the two lights always hitting the same spots, blending to form orange. As we can see in figure 2, these opinions were in a small minority.

### 4.2. Quiz 2

While students widely agreed on the correct alternative in the first quiz, students were more divided in the second quiz. Figure 3 suggests that the students were mostly divided between there being no difference to the interference pattern with the apparatus submerged in water and the interference pattern being narrower. Figure 3 shows the results from the classes using Peer Instruction (right histogram). We see an increase in votes in the last alternative, 'there being no difference', suggesting that arguments for this alternative had a higher impact during group discussion.

*No refraction - no difference*

Arguments for there being no difference to the interference pattern were based on the absence of refraction. Since all of the equipment is submerged in water, the light does not cross a surface between two media. Students acknowledged that a difference in interference pattern would have occurred if the light emerged from air, but with no refraction between air and water there will be no change since water in itself has no effect on the wavelength. Below are a few typical examples of such arguments.

***Stud r:*** Won't it just be the same?

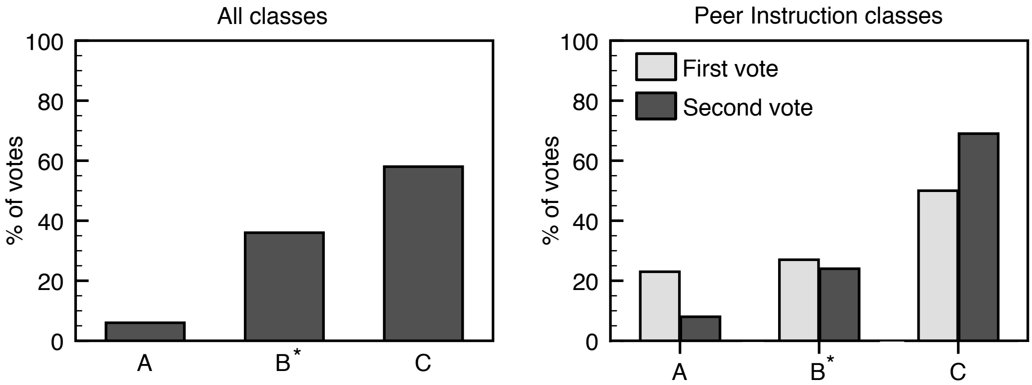



FIG. 3: Voting results from the second quiz. The histogram to the left shows the collected N=117 responses from all four classes while the histogram to the right shows the collected responses from the two classes using Peer Instruction (N=52 for the first vote and N=51 for the second). * The correct alternative.

***Stud l:*** Do you think so?

***Stud r:*** Yes, it just goes in and out of the same material.

—

***Stud m:*** The wavelength isn't changed.

***Stud l:*** Yes, I thought the same thing, there is nothing that changes.

***Stud m:*** Yes, there is no refraction because you are in water the whole time. You don't go from air into water, or water into air. You are always in water.

—

There will be nothing that happens with the wavelength. Because I would think that... yellow light that you send from air down into water is yellow under water as well.

The reasoning behind students' claim of water having no effect on the wavelength becomes more apparent by students discussing the relationship between refraction and the refractive index. Since there is no refraction between an air-to-water surface, the refractive index is irrelevant. Furthermore, the refractive index does not appear in the grating equation, which was also used as an argument for its irrelevance. The first example below shows students in consensus about the irrelevance of the refractive index while the second example shows how students would use it to contradict other group members presenting the refractive index as vital to the solution.

***Stud m:*** Won't it just be the same? Because it is the same refraction. You don't get any refractive index when everything goes through water. You don't have that, right? You don't get any refractive index when everything goes through water, right?

***Stud l:*** No, it doesn't go from air to water, no.

***Stud m:*** It is just water-water and then it should be the same.

—

***Stud r:*** So, if you have a higher refractive index then...

***Stud m:*** Yeah, but you know, that is if you go from one medium to another. Here it is in the same medium.

***Stud l:*** [agrees]

***Stud r:*** Yes, but it still has to be as when you do calculations on light in water, that you divide by the refractive index.

***Stud l:*** Yeah, but then it is from air to water.

***Stud m (to stud r):*** No, but here the whole setup is in water, so there will be no refraction. It is kind of in the same medium.

*Maintaining equality of the equations*

As shown in Figure 2, there was also a large number of students who voted for the interference pattern becoming narrower. Nevertheless, arguments towards this alternative were not as 'direct' as for alternative C, i.e., a direct cause and effect between the argument presented and how it related to the solution. The arguments had two starting points: the relevance of the refractive index or, more commonly, the speed of light in water. Students would often start with a claim of the speed of light being lower in water. The process towards reaching a conclusion about the interference pattern included students trying to identify and use the relation between speed, frequency and wavelength to find how the speed of light affected the wavelength. The emphasis was constantly on how they could manipulate the equations while still maintaining equality, i.e., that both sides of the equality sign 'stay the same'. With the change in wavelength identified, students would either turn to the grating equation or simply state that a lower wavelength results in a more narrow pattern. The example below is one of the more complete arguments of equality.

What I'm thinking is that the frequency of the waves is constant. Since the frequency given by the speed divided by lambda [the wavelength], then... The speed will be decreased, it goes slower. Then lambda has to decrease in order for the frequency to stay the same, and in order to decrease [inaudible]... If it decreases in this side [points to the notebook] then it has to decrease the same on this side [points to the notebook] in order for there to be equality. Then, I think that, since this decreases [points to the notebook], then it has to be, it has to become less.

Figure 3 shows that there were also students who voted for a wider interference pattern (alt. A). All arguments towards this alternative, which were found in the video material, used the same logic of equality described above, but where students would wrongly manipulate the equations. Below is an example of how a student is corrected when making such a mistake.

***Stud m:*** When you decrease the speed on one side of the sign, then you have to increase on the other side - the frequency is constant, so then the wavelength has to increase. Then you get bigger width.

***Stud l:*** No, that is not entirely correct. If you decrease something on one side of the equality sign, then you have to decrease it on the other side as well. The question is... the thing is where you are relative to the fraction line. Because if it is beneath the fraction line and you increase, then the total will decrease.

## 5. DISCUSSIONS AND CONCLUSIONS

### 5.1. Misconceptions

The results indicate that many students misunderstood the concept of diffraction, despite being able to find the correct alternative in the first quiz. Arguments presented suggested a mental image and understanding of diffraction, a physical optics phenomenon, being based on the geometrical model of refraction. One could argue that students' use of the word 'refraction' likely was a result of students being novices in using scientific language instead of actually thinking of the concept of refraction. Previous studies have shown that students often use scientific terms wrongly or without understanding their meaning [18]. The argument of a simple word confusion seems at first glance to be strengthened by students' claim of more refraction with longer wavelength (which is not the case).

However, this could be explained by examining the curriculum and textbook used. The syllabus textbook on the preparatory physics course, at the time of this study, used the second edition of the Norwegian textbook 'Rom, Stoff, Tid' (Space, Matter, Time) by Jerstad, Sletbak and Grimenes [19]. Students are presented with the basics of refraction at a single surface, but dispersion (how refraction varies with wavelength) is only addressed (but not referred to by name) in a brief paragraph at the end of the refraction chapter. The only reference to which wavelengths have a higher degree of refraction is a figure of different wavelengths refraction through a prism. In other words, students' knowledge of dispersion was minimal. When students were introduced to the concept of diffraction or 'bending' of different wavelengths, many students might have mistaken this for dispersion. The claim of students actually using a mental model of refraction is further indicated by some students directly referring to refraction concepts.

The second quiz also shows a confusion of the relationship between refraction and the refractive index, i.e., that the refractive index is not relevant since there is no air-to-water surface where refraction occurs. It might seem paradoxical to use the argument of no refraction because the light 'goes in and out of the same material' while still regarding refraction occurring at the slits. One possibility is that when discussing the quiz students thought of actual, physical diffraction grating slits, which often consist of a transparent material. Thus, the students might have thought of diffraction as light changing direction because of refraction in the transparent material, not considering the fact that any refraction occurring when light enters the material would be 'cancelled out' when the light leaves on the opposite side. There was one student who argued for there being a glass material at the slits. Although no other students used this argument, it could be possible that more students shared this view, in particular since students have seen and handled such diffraction gratings in previous lectures.

Refraction because of a transparent material does not, however, explain why the effect is stronger with more narrow slits, a concept which is clearly emphasised in the textbook with figures of diffraction through different slit widths as well as being emphasised during lectures, and therefore very likely a concept that the students remember. Consequently, a second possibility is that students simply regarded light as rays changing direction as it passes through narrow openings (i.e., using a geometrical model). In both cases, there is no difference when the apparatus is submerged in water because there is no *additional* refraction, only the refraction that already occurs at the slits. Previous studies have also found conceptual misunderstandings with regard to geometrical and physical optics, either by students not being able to separate one from the other [20], or by using a hybrid model of both physical and geometrical optics to explain physical optics such as diffraction [21].

An example of this hybrid model was a differentiation between light going through the middle of the slit, which formed a geometrical image, and light being bent to form other bright areas on the screen when striking the edges of the slit. We see a similar tendency in our findings with arguments of 'no refraction at zero degrees' suggesting that students might believe that refraction (or diffraction) only occurs for light striking at the edges, while light hitting the centre goes unchanged. Another possibility is that they think of incoming light striking the slits at different angles and that light striking 90 degrees on the slits forms an incident angle of zero degrees resulting in no refraction, i.e., no change in light direction.

### 5.2. Implications for teaching with SRS

Our results show an important aspect regarding teaching with SRS. Many students were able to argue for the correct answer in the first quiz, but most arguments were not based on a correct understanding of the physical phenomenon. Therefore, teachers should be careful interpreting voting results from SRS as a correct representation of students' understanding of the subject matter. Students in our study could identify the relevant equation, and how to use it, but their arguments show that they did not fully understand the subject matter. A similar result can be found in the study by Hrepic, Zollman and Rebello [18] where students with an incorrect mental model of sound waves in some instances even received higher test scores on the subject matter than students with the correct mental model. Moreover, if the teacher explanation following the discussions does not address the misunderstandings, the whole SRS-session can give both the teacher and students a wrong impression of the level of understanding. For instance, in our case, the teacher explanations of the first quiz were likely to focus on the grating equation and how it is affected by different wavelength. As this does not conflict with the misconception of more refraction with red light, it could have resulted in students becoming more confident in this argument.

Students conceptual misunderstanding of diffraction could have affected their decision making during the second quiz. Refraction was established in the first quiz by students as a vital part of the solution. This could have increased the impact of arguments of no refraction during the second quiz and Figure 3 shows that more students were convinced of the incorrect alternative after discussion. Although the second quiz probably cleared the confusion between refraction and refractive index for many students, their conceptual misunderstanding of using refraction to understand diffraction was probably not cleared after the SRS-session. McDermott [21] argues that 'certain conceptual difficulties are not overcome by traditional instructions. Persistent conceptual difficulties must be explicitly addressed in multiple contexts' (p.165), and our results are a clear example of this. If SRS is going to be used as a tool to identify and clear up misconceptions, it is important to present students with several quizzes with significantly different settings or contexts as this increases the probability of identifying misconceptions. If students are not presented with new and unfamiliar settings, it also increases the possibility of SRS misleading both students and teachers about the level of understanding of the subject matter. Nevertheless, teachers should be careful of presenting too many quizzes during lectures as this can result in much of the curriculum being dependent on self-study [8]. Unless the lecture format is centred on SRS-quizzes, teachers should strive for a balance between the numbers of quizzes to present each lecture.


[1] Knights, R. (2002) *Five Easy Lessons. Strategies for Successful Physics Teaching*. Addison Wesley.

[2] Masikunas, G., Panayiotidis, A., and Burke, L. (2007) The use of electronic voting systems in lectures within business and marketing: a case study of their impact on student learning. *Research in Learning Technology*, **15**, 3–20.

[3] Kay, R. and Lesage, A. (2009) Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers and Education*, **53**, 819–827.

[4] Duncan, D. (2006) Clickers: A new teaching aid with exceptional promise. *Astronomy Education Review*, **5**, 70–88.

[5] Beatty, I. (2004) Transforming student learning with classroom communication system. *Research Bulletin*, **3**, 2–13.

[6] Wood, B. (2004) Clickers: A teaching gimmick that works. *Developmental Cell*, **7**, 796–798.

[7] Rice, R. and Bunz, U. (2006) Evaluating a wireless course feedback system: The role of demographics, expertise, fluency, competence, and usage. *Studies in Media and Information Literacy Education*, **6**, 1–32.

[8] d'Inverno, R., Davis, H., and White, S. (2003) Using a personal response system for promoting student interaction. *Teaching Mathematics and its Applications*, **22**, 163–169.

[9] Draper, S. and Brown, M. (2004) Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, **20**, 81–94.

[10] Hake, R. (1998) Interactive-engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, **66**, 64–74.

[11] Crouch, A. and Mazur, E. (2001) Peer Instruction: Ten years of experience and results. *American Journal of Physics*, **69**, 970–977.

[12] Hatch, J., Jensen, M., and Moore, R. (2005) Manna from heaven or "clickers" from hell. *Journal of College Science Teaching*, **34**, 36–39.

[13] Nicol, D. and Boyle, J. (2003) Discussion in large classes: a comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, **28**, 457–473.

[14] Mazur, E. (1997) *Peer Instruction: A user's manual*. Upper Saddle River, NJ, Prentice Hall.

[15] Nielsen, K., Hansen-Nygård, G., and Stav, J. (2012) Investigating Peer Instruction; how the initial voting period affects students' experience of group discussion. *ISRN Education*, **2012**.

[16] Kaartinen, S. and Kumpulainen, K. (2002) Collaborative inquiry and the construction of explanations in the learning of science. *Learn. Instr.*, **12**, 189–212.

[17] James, M. (2006) The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics*, **74**, 689–691.

[18] Hrepic, Z., Zollman, D., and Rebello, N. (2010) Identifying students' mental models of sound propagation: The role of conceptual blending in understanding conceptual change. *Phys. Rev. Special topics - Physics Education Research*, **6**.

[19] Jerstad, P., Sletbak, B., and Grimenes, A. (2003) *Rom, Stoff, Tid*. J.W. Cappelens Forlag A.S., 2 edn.

[20] Ambrose, B., Chaffer, P., Steinberg, R., and McDermott, L. (1998) An investigation of student understanding of single-slit diffraction and double-slit interference. *American Journal of Physics*, **67**, 146–155.

[21] McDermott, L. (2000) Bridging the gap between teaching and learning: the role of physics education research in the preparation of teachers and majors. *Investigações em Ensino de Ciências*, **5**, 157–170.