EFFECTS OF A PROBLEM-BASED LEARNING APPROACH ON STUDENTS' SPREADSHEET SKILLS

Johannessen, Trond Vegard, NHH Norwegian School of Economics, Helleveien 30, NO-5045 Bergen, Norway, trond.johannessen@nhh.no

Fuglseth, Anna Mette, NHH Norwegian School of Economics, Helleveien 30, NO-5045 Bergen, Norway, anna.mette.fuglseth@nhh.no

Abstract:

This paper reports the preliminary findings from a quasi-experimental project to study the effects on bachelor students' spreadsheet skills after a section on spreadsheet modelling. The spreadsheet modelling section was part of a compulsory 7.5 ECTS introductory course on business data processing.

In the pretest of our quasi-experiment, conducted at the beginning of the course, we asked the students to hand in their answer to a spreadsheet task. The task was voluntary, and 31% of the students (N=117) submitted their model or an empty spreadsheet. The results showed that around 90% of the students lacked basic technical spreadsheet skills, and that they did not know how to structure a spreadsheet so that their model was useful for decision support.

The posttest, conducted at the end of the spreadsheet modelling section, showed that more than 90% of the same students (N=117) had acquired basic spreadsheet skills and managed to model a similar task in the spreadsheet according to principles of spreadsheet design. There was, however, one particularly disappointing exception: 26% of the students (N=30) still had constants in some of their formulas and, thus, hidden assumptions in their model.

It is not surprising that most students had acquired spreadsheet skills after the course. Therefore, we emphasise the weaknesses of the students' answers in order to enhance the understanding of how to improve spreadsheet teaching. Limitations of our study are discussed together with suggestions for further research.

1 INTRODUCTION

The purpose of this paper is to study the effects of a basic course section on spreadsheet modelling. We want to enhance the understanding of how to improve spreadsheet teaching when limited time and resources are available in a study programme.

The spreadsheet is considered an important tool to support decision makers when they analyse data and make decisions (Presthus, 2014; Madahar, 2011; Panko, 2006). Spreadsheets allow end users to take command of their programming needs. However, spreadsheets lack the structure of programming languages enforcing users to follow certain rules in order to avoid errors in specifications and formulas. Evidence exists that critical errors occur in spreadsheets applied in managerial business processes (Panko & Aurigemma, 2010; Morgan Chase, 2013). Furthermore, experience with spreadsheets does not seem to increase expertise considerably (Madahar, 2011).

Therefore, bachelor students should acquire spreadsheet modelling skills during their studies, that is, the students should be taught knowledge of spreadsheet design and programming principles, and they should be trained to apply the principles when developing spreadsheet models.

In order to enhance our understanding of how to plan the lectures and training in an introductory course in business data processing, we designed a quasi-experimental study. We gave the bachelor students a spreadsheet task at the beginning of the course. The task was designed to elicit the students' knowledge of spreadsheet modelling according to principles that are similar to principles of "good spreadsheet skills" applied at other business schools, see, for example, Frownfelter-Lohrke (2017). In the planning of our teaching, we particularly emphasised the weaknesses revealed in the pretest of our study. After the course section on spreadsheet modelling, we gave the students a task that contained the same elements in order to study the effects of the teaching.

The rest of this paper is organised as follows: In the next section, we motivate and position our study. Then we present our research design and procedures for data collection and analysis. In section three, we present our findings emphasising the weaknesses of the students' responses to the posttest. In the final section, we summarise the findings and discuss implications for teaching spreadsheet design and programming. Limitations of the study are discussed together with directions for further studies.

2 LITERATURE REVIEW

In this section, we explain essential concepts, position our study and present our research questions.

A spreadsheet is a large electronic grid sheet. It is easy to enter numbers and make the programme perform calculations, so that individuals with little knowledge of programming principles may use spreadsheets, for example for setting up budgets and calculating results. Even though the spreadsheet helps users avoid *syntactic* errors when entering formulas and applying functions, the spreadsheet does not support the users with regard to the structuring of the spreadsheet, the specification of variables and the entering of formulas that are programmed appropriately (see below). In other words, the spreadsheet does not support users with regard to representing and solving business problems so that the models are suited for considering alternative scenarios or for use by other individuals than the person who has developed the model.

Consequently, there is evidence in the literature that critical errors occur in spreadsheets applied in managerial business processes (Panko & Aurigemma, 2010; Morgan Chase, 2013). From our contacts with business firms, we have an example of a model developed for evaluation of shipping investments. We were asked to take a look at the design of the model, and we detected that the model returned the wrong net present value. The cause of the error was an exchange rate entered as a constant in just one of the formulas deep down in the model. There were no syntactic errors, and neither the programmer nor the other users had detected the error.

Furthermore, experience with spreadsheets does not seem to increase expertise considerably (Madahar, 2011). In his study of 686 managers and consultants' spreadsheet use (Madahar, 2011, p. 117), 62.3% of the respondents reported that they had extensive experience with spreadsheets, but also that they considered that they had only some expertise. 55% of the respondents did not have formal instruction.

There is a rich literature on the use of spreadsheets. Much of this literature concerns spreadsheet errors and the categorisation of spreadsheet errors (see, Powell, et al., 2008, for an overview). Most of the literature is descriptive. The purpose is to enhance the understanding of which types of errors occur in spreadsheets, and then – implicitly – also the cause of some errors. For example, Galetta et al. (1993), distinguish between domain errors and device errors. Domain errors refer to errors that are related to the application area (e.g., budgeting), while device errors are related to the application of the spreadsheet (e.g., wrong cell range in a formula). Panko and Halverson (1996) made a distinction between quantitative and qualitative errors. Quantitative errors are errors in the current version of the spreadsheet, whereas qualitative errors refer to risky practices that could lead to errors in subsequent versions (e.g., constants in formulas).

Some of the studies attempt to understand the *process* of developing and debugging a spreadsheet model (Panko & Halverson, 1996; Panko, 1998, 2006). The purpose is to find out how the errors occur when developing a spreadsheet model.

Other studies attempt to evaluate the *effects of training* (Janvrin & Morrison, 1996; Kruck & Sheetz, 2006). These studies are closely related to prescriptive research, presenting guidelines for spreadsheet design and programming. In a recent paper, Frownfelter-Lorcke (2017) presents a plan for teaching spreadsheet design based on "principles of good design" by Kreye and Pendley (1998). Her students are required to complete three spreadsheet assignments in an iterative and repetitive process. According to the author, the students have acquired "good spreadsheet design skills and improved their basic Excel skills" after having completed the three assignments.

To the best of our knowledge, the studies investigating training and effects of training do not control for the students' knowledge at the start of the training. We believe that understanding of the students' previous knowledge and skills is important both when the effects of training are evaluated, and when the training is planned.

Our research is particularly related to the concern that spreadsheet skills do not seem to develop through experience (Madahar, 2011), that is, instruction and training are necessary. Since the spreadsheet is considered an important tool to support decision makers when they analyse data and make decisions (Presthus, 2014; Madahar, 2011; Panko, 2006), it is important that business students learn how to handle a spreadsheet effectively during their studies. As responsible for an introductory course in spreadsheet design and programming, we are, therefore, concerned with the effective organisation of the course.

An essential element in planning a course is to adapt the course contents and the implementation of the course plan to the students' knowledge. Our course is based on the principles of problem-based learning (PBL). PBL was developed by Barrows (1986) as a response to the criticism against courses with traditional lectures, in that they do not prepare the students for their future lives. This criticism against traditional lecturing originated in medical education, but has spread to other disciplines, such as economics (e.g., Fuglseth, Grønhaug & Jörnsten, 2016).

The primary goal of PBL is to enhance learning by requiring that students solve ill-structured, interdisciplinary problems. PBL is based on constructivist assumptions that knowledge is individually constructed and socially co-constructed from interaction with the environment. Since knowledge is an individual construction, knowledge cannot be transferred, but in interactions, we can test the degree to which our individual understandings are compatible (Lave & Wenger, 1991; Wenger, 1998; Savery & Duffy, 2001). The principles of PBL are often misunderstood so that students are presented with problems from the start of the course, and that the teacher plays a minimal role (e.g., Srinivasan et al., 2007). Barrows (1986, p. 484) emphasised, however, that PBL includes an "active, teacher-guided exploration and evaluation of the problem using facilitation or tutorial skills, which directly activate the student's prior knowledge, much of which may otherwise be beyond conscious recall". Barrows (1986, p. 484) elaborated further on the teachers' role as lecturer when it is necessary for the students to receive some initial information, "out of concern that the student must know some vocabulary and definitions, and have an overview of the subject before any meaningful problem-solving can occur".

Thus, this paper is closely related to the paper by Frownfelter-Lorcke (2017), developing a spreadsheet course with problems of increasing complexity to be handled by the students. However, we extend her teaching method in the following three ways: In a pretest, we enhanced our knowledge of the students' spreadsheet skills before we planned the spreadsheet section of the course. We then designed the section with regard to lectures and teacher-guided problem exploration and solving according to the principles of PBL. In a posttest, we checked the effects of our teaching.

Frownfelter-Lorcke (2017) characterised her course as a course on the design of spreadsheets. We prefer to make a distinction between technical skills and design skills. Technical skills are related to the students' ability to enter data and performing calculations. Design skills concern the students' ability to structure and program the spreadsheet so that the models are easy to understand for others than the developer (i.e., transparent), and so that the models are suitable for analysis when assumptions are changed (i.e., dynamic). Thus, we formulated the following research questions for our study:

Having completed the section on spreadsheet modelling,

1. what are the effects on the students' skills with regard to entering data and performing calculations using appropriate cell references and standard library functions?

2. what are the effects on the students' skills with regard to designing a spreadsheet model that is easily understood by other users?
3. what are the effects on the students' skills with regard to designing a spreadsheet model that is suitable for analyses of changed assumptions and uncertainty?

The research questions build on each other, that is, if the students do not possess basic technical spreadsheet skills, they are not able to build transparent and dynamic models.

3 RESEARCH DESIGN

In this section, we describe our research design and procedures.

3.1 Research strategy – an exploratory quasi experiment

In order to answer the research questions, we planned an exploratory quasi experiment with a pretest and a posttest (Cook & Campbell, 1979). The pretest was administered in the first meeting with the students, and since the attendance was not mandatory, we had no possibility to randomise the subjects.

In both the pre- and the posttest, we handed out a budgeting assignment. The assignments contained the same technical and design elements, but the contexts were different. In the pretest, the students were asked to assist a fellow student who was going to decide financially on his housing situation, see Johannessen and Fuglseth (2017). In the posttest, the students were asked to analyse the profitability of a planned investment in a go-cart track. The assignments were designed so that the students were expected to demonstrate their knowledge and skills with regard to the development of a transparent and dynamic spreadsheet model. The assignments should be solved using the spreadsheet programme MS Excel 2016.

In the pretest, the tasks were specified. In the posttest, the students were expected to apply the principles taught in the course without being explicitly asked to do so.

The mandatory course on business data processing is placed in the third semester of the bachelor programme. Around 400 students attend the course each year. We controlled that Excel had not been taught or used in other courses of the bachelor programme when we performed the experiment. In the first year of their study, the students have an introductory course in business economics. Thus, the domain of the assignments should be well-known to the students, so that our evaluation of the students' spreadsheet skills could focus on the application of the *device*, cf. Galetta et al. (1993).

3.2 Procedures for data collection

Before the first meeting in the course, the students had been told to bring their PCs with MS Excel installed. The pretest assignment was handed to all students present in the second part of the introductory lecture of the course. The students were asked to solve the problem and send their Excel workbook by email to the teacher after class, that is, they had 45 minutes to solve the assignment.

The students were informed that the assignment was voluntary, that their performance had no influence on their grade, and that the purpose of the assignment was to improve our understanding of their spreadsheet skills so that we could adapt the course to the students' knowledge and skills. We urged all students to send in their answers. Students that did not manage to solve even part of the problem using a spreadsheet, were asked to send an email explaining why they were not able to attach a spreadsheet model. As an incentive to participate, we announced that ten gift cards of NOK 1 000 would be randomly assigned to those who sent in their answers.

At the end of the spreadsheet modelling section, the posttest assignment was administered through the learning platform to all the students signed up for the course. The answer to this assignment was mandatory, that is, each student had to submit an answer in order to be allowed to sit for the exam. This requirement was according to the usual procedures for the course. With regard to the mandatory assignment, the students were not allowed to collaborate, and they were not allowed to get help from the course assistants or the user support personnel except for technical computer problems. Since the answers were submitted to the teacher by email, we had the possibility to compare the answers to the

posttest with the same group of students who had submitted the answer to the pretest. Thus, we had a within-group design (Frankfort-Nachmias & Nachmias, 2007).

3.3 The "treatment": The contents and teaching activities of the spreadsheet section

In this section, we present the contents of our course section and the teaching activities. As mentioned above, the course is delivered in the third semester of the bachelor programme. Thus, the students have experience with the university system; they are grown-up people and expected to take responsibility for their own learning. Therefore, the participation in teaching activities is voluntary.

3.3.1 Course contents

In line with the European Qualifications Framework (EQF), we distinguish between knowledge and skills[1] in the course. Knowledge is related to the students' understanding of spreadsheet concepts and design principles. Skills concern the students' cognitive and practical abilities to solve problems. Regarding the contents of the course, we distinguish between technical and design knowledge and skills. The technical knowledge and skills are related to entering data and performing calculations in the spreadsheet. The design knowledge and skills are related to the development of a spreadsheet model so that it is suitable for decision support.

The pretest revealed that 90% of the students lacked basic technical spreadsheet skills, and that they did not know how to structure a spreadsheet so that their model was useful for decision support. For example, we expected that bachelor students were well acquainted with concepts such as constants and variables from lectures in mathematics. However, the pretest indicated that the students did not seem to recognise these concepts in a spreadsheet setting. Thus, according to the principles of PBL (Barrows, 1986), the students needed to develop a conceptual system for spreadsheet programming and design principles.

Concerning technical knowledge, we therefore decided to start the course at the basic level with explanation of concepts and components, that is, how to enter text and numbers into spreadsheet cells, how to construct formulas, how to search for and use the built-in functions, and how to format a spreadsheet.

With regard to spreadsheet design principles, we have developed a set of principles based on our knowledge of system development and programming principles, and on our experience from developing decision support systems using spreadsheets for the user interface (partly documented in Fuglseth & Grønhaug, 2003; Fuglseth & Strandenes, 1997). Our principles are similar to the design principles applied by Frownfelter-Lorcke (2017, pp. 70-71), that are based on the advice from Kreye and Pendley (1998). In the following, we present a subset and an abbreviated version of these principles, emphasising the principles of transparency and dynamics (see, Fuglseth, Håtuft & Johannessen, 2011, for a detailed description of the principles):

Transparency

The principle of transparency reflects that a spreadsheet model shall be clearly laid out so that it is apparent what the model and the variables represent, and how the model may support the users in their decision processes.

Guidelines for modelling according to this principle are that all assumptions of the model must be visible so that they can be inspected by the user. Furthermore, results (outdata) and input data (indata) should be clearly separated. Formulas should be easy to understand, that is, long formulas should be avoided and intermediate variables defined. Data should be grouped into purposeful categories. Purposeful categories depend on the type of model. For a budgeting task, a purposeful grouping of indata may be the distinction between decision variables and non-controllable variables. For example, managers can decide on the price of their products, but they cannot determine the price of crude oil.

---

[1] See, for example, https://ec.europa.eu/ploteus/en/content/descriptors-page

Another guideline pertaining to the principle of transparency is that model variables should be thoroughly explained. For example, a label describing the price of a product should also contain the unit applied, for example, whether the price is for 1000 pieces.

Dynamics

The principle of dynamics reflects the characteristic of a spreadsheet model that the users can change the assumptions (indata), and that the consequences of the changes are immediately transmitted throughout the model. The principle implies that the users can study the effects of their changes through the intermediate results to the final results. In other words, the model should be designed so that it is suitable for exploration of changed assumptions and uncertainty through "What if" analyses.

The most essential guideline for implementation of this principle is closely related to basic technical skills and the principle of transparency: Formulas must only contain cell references, not constants. The only exception to this guideline is a constant that is considered a fixed element of a general calculation formula, such as the formula for calculation of percentages, for example that a cell contains the formula "$= 1 + \$B6$", where cell $\$B6$ contains the rate.

Another guideline is that the designer should consider whether indata should be entered as time series data. For example, the price of a product may at first be considered to be fixed, but it may be purposeful to allow not only changes of the price, but also different changes of the price over time.

We developed an evaluation form for scoring the answers. The statements in the form should be scored as true or false. The statements measured technical spreadsheet skills and certain aspects of the students' design skills. Three experienced teaching assistants helped scoring the answers using the evaluation form with regard to the students' technical skills. These results were coded and analysed in Excel. The authors then analysed the students' spreadsheet models in order to evaluate the design skills, and in this process, we elaborated on and extended the evaluation form.

3.3.2 Teaching activities

As mentioned above, the pretest revealed that most students lacked basic spreadsheet skills, but also that the students' knowledge of spreadsheet modelling varied. Therefore, we planned the teaching activities so that the students could adapt their participation in the activities according to their various knowledge and skill levels.

According to the principles of problem-based learning (PBL) (Barrows, 1986), the course section was organised to help the students develop their conceptual understanding of spreadsheeting and prepare them for real-life spreadsheet modelling through a series of problem-solving sessions of increasing complexity. However, in a large group of around 400 students, it is not possible for a course facilitator to have the extensive contact with the students assumed in the PBL approach. Therefore, we complemented the PBL approach with aspects from the flipped-classroom approach (e.g., Herreid & Schiller 2013):

Instead of plenary sessions, the course facilitator developed a set of video sessions emphasising essential concepts and principles. The purpose of the first two videos was to explain essential concepts and illustrate the use of the concepts in examples that followed the design principles. The third video was explicit on spreadsheet design principles.

- The first video lasted for 24 minutes and introduced the students to the basics of spreadsheeting. Basic concepts and components were explained, and fundamental operations were illustrated, that is, how to enter text and numbers into cells, how to edit, copy and paste the contents of cells, and how to format a spreadsheet. Finally, the three spreadsheet layers were explained, that is, the formula layer, the value layer and the formatting layer.
- The second video lasted for 39 minutes and addressed how to construct formulas using relative, absolute and mixed cell references. Furthermore, the video showed how to search for and to use the built-in standard functions.
- In the third video, which lasted for 17 minutes, the students had an introduction to spreadsheet design principles, such as the principles of transparency and dynamics. The video also illustrated

the advantages of using the design principles with regard to analysis of changed assumptions and uncertainty.

The three videos were available right after the introductory lecture, that is, after the pretest. For each video, the students could download the PowerPoint presentation and the spreadsheet with the examples used in the video.

Related to each video, a computer lab was offered. Each lab lasted for two hours. The labs were organised as problem-solving sessions in which the students worked on assignments related to the issues explained in each video, thus, there were three lab sessions. The students were required to read the prepared texts on spreadsheeting and to watch the videos before they came to class. The labs were organised with 20 – 40 students and administered by experienced student assistants, one assistant per around 20 students. At the end of each lab session, the assistants had a debriefing period discussing and challenging the students' solutions. The student assistants had been trained for the lab sessions by the course facilitator before the course started.

In order to help the students develop their knowledge and skills further, an oracle service was offered during the semester. The oracle service was open two to four hours per week and was administered by two or more teaching assistants depending on demand. Students could drop by the oracle service to get help on the spreadsheet assignments (except the posttest) and to discuss/get an explanation of concepts or principles that they did not understand. The course facilitator was also available for the students during office hours.

At the end of the spreadsheet section, the course facilitator had two debriefing sessions, each lasting two hours. Before the first session, the students were required to solve the pretest assignment anew. In the session, the students were asked to compare their new solution with their original solution and to ask questions related to their problems and reflections related to solving the assignment. The students responded by raising their hands or by anonymously posting their comments to a Padlet, (i.e., a virtual bulletin board, see, https://padlet.com), using their phones, tablets or computers. The Padlet and the posts were visible on the projector screen so that the students' comments/questions served as input to a class discussion on how to handle particular technical challenges, and how and why to design the spreadsheet model in accordance with the design principles presented in the videos. An example of a Padlet post related to technical spreadsheet aspects was: "Why do you use intermediate calculations in your solution?" The course facilitator then illustrated the benefits of splitting up complex calculations into several steps in a spreadsheet shown on another projector screen.

At the end of the first debriefing session, the students were given an extension of the pretest assignment to work on in the session. The purpose of the extension was to apply the design principles to the pretest assignment in order to illustrate how to add input and output variables, and how to analyse the problem with changes of the assumptions. The students had approximately 45 minutes to work on the extension. The course facilitator and two student assistants were available for the students while working on the extension.

3.4   Procedures for data analysis

In our evaluation of the students' skills, we distinguished between technical spreadsheet skills and design skills. The technical skills comprised evaluation of whether the students performed calculations in the spreadsheet that returned the correct result, whether they used appropriate cell references, and the degree to which they used standard functions, for example, the SUM() function instead of several plus operations. The design skills were evaluated according to the students' abilities to represent the assignment problem and the solution according to the design principles described above.

For the analysis of the pretest, we prepared a list of items we wanted to evaluate. The pretest task was constructed to cover these items. The analysis of the pretest was presented in detail by Johannessen and Fuglseth (2017).

In the analysis of the posttest, we used the same procedures, that is, we used the same list of items, and we constructed the task to cover these items. Then, we compared the differences between the answers to the pre- and the posttest in order to study the effects of the teaching.

4 FINDINGS

As mentioned above, the course on business data processing was mandatory, and 383 students was supposed to attend the course. Only 374 students completed the assignment that was required to sit the exam. The 117 students that had sent an email response to the voluntary pretest were among the 374 students that completed the course. In the following, we present our analysis of the 117 students. A summary of the results is shown in Table 1:

Table 1   Summary of the results from the pretest and the posttest

	Pretest N=117		Posttest N=117	
Technical spreadsheet skills				
Students can perform calculations in the spreadsheet	80	68 %	117	100 %
All the calculations are carried out in the spreadsheet	60	51 %	116	99 %
The results of the calculations are correct	75	64 %	106	91 %
The students have not used constants in the formulas	2	2 %	87	74 %
Formulas are created using cell references	78	67 %	117	100 %
Absolute cell references ($) are used when appropriate	4	3 %	114	97 %
Relative cell references are used when appropriate	59	50 %	117	100 %
The students can use standard functions	42	36 %	115	98 %
The IF() function is used for handling conditions specified in indata	0	0 %	106	91 %
Design skills				
The students distinguish between indata and outdata	2	2 %	112	96 %
There are no hidden assumptions	16	14 %	87	74 %
Variables are grouped into purposeful categories	37	32 %	109	93 %
All formulas are easy to understand and transparent	48	41 %	107	91 %
Purposeful labels are used to describe variables	69	59 %	115	98 %
Purposeful units are used in the labels	66	56 %	50	43 %
Indata are perceived as time-series data	5	4 %	113	97 %

4.1   Addressing research question 1

After the course section, what were the effects on the students' skills with regard to entering data and performing calculations using appropriate cell references and standard library functions?

In the pretest, only 80 students performed calculations in the spreadsheet. In the posttest, all students demonstrated that they were capable of performing calculations in the spreadsheet. Compared to the pretest, the students performed *all* calculations in the mandatory assignment in the spreadsheet with one exception: One student had entered the instalment of the loan as a data value, but performed the other calculations in the spreadsheet. This student had to hand in a new answer to the posttest in order to sit the exam.

Eleven students had errors in their calculations. However, these errors were mainly due to erroneous handling of the loan and were as such not related to the technical handling of the spreadsheet.

In the pretest, 78 students used cell references in their formulas, but only two students had no constants in their formulas. In the posttest, all the students used cell references in their formulas. However, 26% of the students (N=30) had still constants in some of their formulas. Five students used constants extensively. The remaining 25 students typically used constants when they entered data they did not expect to change (often), such as the repayment period of the loan.

In the pretest, only four students used absolute cell references when it was appropriate, that is, they used the dollar sign to lock row and/or column references. In the posttest, 114 students used absolute cell references when purposeful. The remaining three subjects used only relative cell references, which is the default way of referencing in Excel. Several students used absolute cell references when not necessary, see, for example, Figure 1 line 13, the references to cells $C$13, $D$13 and $E$13. Using absolute cell references when not necessary may give correct answers in the present version of the

model. However, improper use of absolute references implies that the model does not satisfy the principle of dynamics.

115 of the 117 students used the SUM() function in Excel to summarise data. Seven of these 115 students, however, entered the function arguments as cell references instead of using ranges, making the formulas unnecessary long, for example, SUM(D42; D43; D44; D45; D46) instead of SUM(D42:D46). The remaining two students summarised data using plus or minus signs in their formulas.

In the pretest, only 13 students calculated the net present value of the project, but they entered the mathematical formula, that is, they did not use the NPV() function. In the posttest, 116 students calculated the net present value. One student ignored this part of the assignment. Six students calculated the net present value entering the mathematical formula, and 110 students used the NPV() function. However, 59 of the 110 students using the NPV() function added input values as cell references instead of referring to a range, as described above for the SUM() function.

Table 1 reveals that none of the students performed a required logical test using the IF() function in the pretest. The posttest was designed so that the number of employees was dependent on the number of go-cart trips per year. The initial demand of the project required three employees per year, but if the demand for trips reached 35 000 or more per year, an extra employee was required. In order for the spreadsheet model to handle the number of employees dynamically, the students had to create a logical test using the IF() function. 106 students managed to create such a test. The remaining 11 students entered the number of employees as constants, see Figure 1 line 18, or they ignored this aspect of the assignment.

4.2 Addressing research question 2

After the spreadsheet section, what were the effects on the students' skills with regard to designing a spreadsheet model that is easily understood by other users?

In the pretest, none of the students distinguished between indata and outdata. In the posttest, 112 students *attempted* to make a distinction between the assumptions underlying the model and the calculations leading to the results, and they presented the assumptions in a section above the calculations. The remaining five students had no such distinction, see Figure 1 for an example. The figure shows part of one of the inadequate answers that illustrates several typical errors. The student had used more than one spreadsheet, which was not necessary and reduced the transparency of the model in this case. She had categorised the variables. However, she had mixed the assumptions and the calculations and made it difficult to get an overview of the assumptions and possible relationships among the assumptions.

For a spreadsheet model to be understood by others – and by the developer after some time, all assumptions must be specified. As revealed for research question 1, only two of the subjects managed to create a model without constants in their formulas in the pretest assignment. In the posttest assignment, 87 students had specified *all* assumptions in an indata section. This means that 30 students had still constants in their formulas, that is, their models had hidden assumptions. The student that delivered the model shown in Figure 1 had not understood the principle. She had constants in her formulas in line 18. Furthermore, she had first calculated the investment in cell C8. Then she entered the loan amount (which is the same amount as the investment) as a constant in cell C13, and she referred to this constant when she calculated the interests in line 14.

In the pretest, 69 students created formulas that were difficult to read because the students did not understand how to use ranges, and they did not think of specifying intermediate variables. In the posttest, 107 students had understood how to construct rather transparent formulas. The remaining ten students created formulas that were unnecessarily complex. For example, one student calculated the income in the last year of his budget using the following formula:

$$=E17*(E18*(1+F18)*(1+G18)*(1+H18)*(1+I18)*(1+J18)*(E19*(1+F19)^5)).$$

The student tried to incorporate annual change rates for both the demand and the price in the same formula. If the students had defined variables for price per year and demand per year, the formula could have been simplified considerably.

	A	B	C	D	E	F
3		**År**	**0**	**1**	**2**	**3**
4		**Investering**				
5		Innkjøpspris biler	540000			
6		Sikkerhetsutstyr	390000			
7		Tomteinvestering	6000000			
8		**Sum Investering**	**=SUM(C5:C7)**			
9						
10		**Finansieringskostnader**				
11		Avdrag		=$L$10/$L$11	=$L$10/$L$11	=$L$10/$L$11
12		Renter i %		0,045	0,055	0,06
13		Restbeløp per 31.12	6930000	=$C$13-D$11	=$D$13-E11	=$E$13-F11
14		Renter i kr		=D$12*C$13	=E$12*D$13	=F$12*E$13
15		**Sum finansieringskostnader**		**=D11+D14**	**=E11+E14**	**=F11+F14**
16						
17		**Lønnskostnader**				
18		Antall ansatte		=IF(Inntekter!D6<35000;3;4)	=IF(Inntekter!E6<35000;3;4)	=IF(Inntekter!F6<35000;3;4)
19		Antall timer		=$L$17*$L$18*$L$19	=$L$17*$L$18*$L$19	=$L$17*$L$18*$L$19
20		Årlig økning timelønn			0,03	0,03
21		Timelønn		120	=D21*(1+E20)	=E21*(1+F20)
22		**Sum lønnskostnader**		**=D18*D19*D21**	**=E18*E19*E21**	**=F18*F19*F21**
23						
24		**Markedføringskostnader**				
25		Kostnad		750000		
26		Økning/redusering i %			0	-0,25
27		**Sum markedsføringskostnader**		**=D25**	**=D27*(1+E26)**	**=E27*(1+F26)**
28						
29		**Vedlikeholdskostnader**				
30		Vedlikeholdskostnader		550000		
31		Økning i %			0,025	0,025
32		**Sum vedlikeholdskostnader**		**=D30**	**=D32*(1+E31)**	**=E32*(1+F31)**
33						
34		**Driftsutgifter**				
35		Driftskostnader		420000		
36		Økning i %			0,025	0,025
37		**Sum driftsutgifter**		**=D35**	**=D37*(1+E$36)**	**=E37*(1+F36)**
38						
39		**Faste kostnader**				
40		**Sum faste kostnader**		**65000**	**65000**	**65000**

Figure 1   An example of part of a model illustrating several typical errors

Proper labeling of variables is necessary for a model to be easily understood and to prevent confusion. In the pretest, only 69 subjects labelled variables in their models appropriately. In the posttest, 115 students used appropriate variable labels. The two subjects who lacked proper labelling, used only keywords, such as "Cars" as a variable label instead of, for example, "Purchase price go-carts" grouped under "Investments". For example in Figure 1, the student had not specified a proper label for the number of hours in line 19. It is not clear whether the label "Number of hours" refers to number of hours per day, per month or per year.

A disappointing finding from the posttest is that 67 students had not specified units at all or had not specified units adequately. A consequence of not specifying units is that ambiguity is introduced, for example, whether a variable value represents NOK per month or NOK per year.

4.3   Addressing research question 3

After the spreadsheet section, what were the effects on students' skills with regard to designing a spreadsheet model that is suitable for analyses of changed assumptions and uncertainty?

In order for a model to be suitable for analyses of changed assumptions and uncertainty, the model must satisfy the principles of transparency. In the posttest, 87 subjects created a model without hidden assumptions. These students had also taken into consideration that the assumptions could change over time. Consequently, it was rather easy to extend the time horizon of their models, for example, if they changed the repayment period of the loan to eight years.

5   DISCUSSION AND CONCLUSION

The analysis has shown that the teaching activities had an effect on the students' abilities to represent and solve capital budgeting problems in the spreadsheet – as was to be expected. We can summarise the effects in the following way: Before the course section on spreadsheet modelling, around 90% of the students did not know how to handle a spreadsheet effectively. After the course section, more than 90% had acquired the knowledge and skills to handle at least the basics of spreadsheeting. It is not surprising that some students did not grasp all the aspects of spreadsheet modelling taught. The posttest was a requirement to sit the exam. It was only to be expected that some students delayed some of their learning efforts to the final exam.

Still, it is rather disappointing that 26% of the students had not understood the importance of specifying all indata as variables. Most of the errors were related to the students' interpretation of the assignment. They seemed to use constants when they did not expect variable values to change, such as the repayment period and the employee capacity. In the assignment, the employee capacity was not specified clearly as an input variable. The students were required to reflect on how to represent that parameter in the model. However, if the employee capacity is not specified as an input variable, the model contains a hidden assumption, that is, the model does not satisfy the principle of transparency. The repayment period is not expected to change often, but if the repayment period is not specified as a variable, the model does not satisfy the principles of transparency and dynamicity. Errors related to employee capacity and repayment period were discussed in a debriefing session. If the students had applied constants in other cases, their answer to the mandatory assignment was not approved.

It is difficult to understand why not all students had understood how to apply absolute cell references, and how some students could miss the use of ranges, for example, when they applied the SUM() function. However, in our sample of 117 students, five students were "responsible" for most of the violations of the principles taught in the course section. With the rejection of the answers to the mandatory assignment that contained the above-mentioned errors together with a debriefing session for the posttest assignment, we hope that most students have acquired the *basic* technical knowledge and skills to apply the design principles effectively.

One tentative explanation of why not all students had acquired knowledge on basic technical spreadsheeting, is that they have mainly watched the videos. This explanation is supported by an analysis of the students' papers submitted in connection with their final exam. In their papers, the students mainly referred to the videos instead of the syllabus. Is it possible that the video sessions and the recording of the lectures have contributed to less reflection and less exploration of the prescribed texts?

Acquiring detailed technical spreadsheet knowledge is a prerequisite for exploiting the potential of the spreadsheet and for building models that satisfy the principles of transparency and dynamics. The advantage of the videos is that the course facilitator can illustrate *how* an experienced programmer performs certain operations in the spreadsheet and builds a transparent and dynamic model. The disadvantage of the videos is probably that the students are not able to catch the details of oral explanations and illustrations in a video lasting around 20 minutes.

The course had been designed so that the students were expected to watch the teaching videos *and* to read and work with the syllabus. The syllabus explains the importance of specifying all input data as variables and gives examples of how also rather stable indata should be represented as variables. In addition, the syllabus has examples of powerful functions to support managerial decision making, and encourages via several hints the students to explore the potential of the spreadsheet as a decision-support tool.

One of these hints encourages the students to use the Microsoft help systems. The systems explain the syntax of using the standard functions and provide examples. However, the help systems may have contributed to some inadequate and erroneous use of the functions. For example, the NPV() function is illustrated with cell references instead of indications of the use of ranges. The SUM() function is even illustrated with the use of numbers instead of references, as shown in Figure 2:

Figure 2   Example of a ScreenTip in Excel

The findings from the analysis of the posttest have implications for the teaching activities. We believe that the flipped classroom approach has been a success. Traditional lectures with one-way communication to around 400 students are probably not very effective. In the students' course evaluation, the video sessions were well received by the students. We believe, however, that we should provide more videos of a shorter duration and with a title explaining the contents so that the students could easily look up a video illustrating a particular issue.

The lab sessions were also highly evaluated, but we should probably introduce the students to more demanding tasks in the second and third lab sessions challenging the students to search for solutions in the syllabus. The presentations and discussions of the task solutions should include illustrations of the correct use of absolute cell references in order to build dynamic models. Furthermore, the teaching assistants should illustrate how to use the Microsoft help systems appropriately for exploring and applying the potential of the spreadsheet.

The debriefing sessions with the course facilitator worked well with regard to the representation of the problem and the development of the model and the solution. However, working with the extension of a problem in a plenary session was not effective even with two assistants to help the students. This part of the sessions should have been organised as group sessions, similar to the labs.

In a third of a mandatory 7.5 ECTS course it seems that we have managed to help bachelor students develop basic spreadsheet knowledge and skills related to capital budgeting. However, the reported results from the analysis of the sample of 117 students may be too positive. The sample may represent students that were more motivated for learning spreadsheet modelling than the rest of the class. Only five students of the sample, that is 4.3%, did not get a pass in the first attempt to submit the mandatory assignment, whereas 24 students from the rest of the class, that is 9.3%, had to resubmit their answer.

However, most of the students had not developed a construct system so that they could transfer their knowledge to new settings. In the final exam, in the data modelling section, we asked the students to design a spreadsheet template for entering data into a relational database. Most students did not recognise that this task required application of the principles from the spreadsheet section.

We believe that business schools should extend the teaching of spreadsheeting, not necessarily in courses in business data processing, but in applying the principles from such a course in other parts of the curriculum. Such extension requires, however, that the instructors in other courses have a good command of the principles. Another approach could be a collaboration between instructors in, for example, finance or business economics and data processing, so that comprehensive assignments were used in both courses

6   REFERENCES

Barrows, H. S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, 20, 481 – 486.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Frownfelter-Lohrke, C. (2017). Teaching good Excel design and skills: A three spreadsheet assignment project. *Journal of Acconting Education*, 39, 68-83.

Fuglseth, A. M. & Grønhaug, K. (2003). Can computerised market models improve strategic decision-making? An exploratory study. *Journal of Socio-Economics*, 32, 503-520.

Fuglseth, A. M., Grønhaug, K. & Jörnsten, K. O. (2018). Students' Ability to Apply Their Knowledge in a Gaming Exercise: An Exploratory Study. *Scandinavian Journal of Educational Research,* 62(1) 88-108.

Fuglseth, A. M., Håtuft, J. V. & Johannessen, T. V. (2011). *PC-bruk 1 for høyskoler og universiteter*, 6th ed. Bergen: Fagbokforlaget.

Fuglseth, A. M. & Strandenes, S. P. (1997). NorshipS: A Decision Support System for Analyses of Bulk Shipping Markets. *Computers in Human Behavior*, 13(4), 517-541.

Frankfort-Nachmias, C. & Nachmias, D. (2007). *Research Methods in the Social Sciences,* 7th ed. New York: Worth Publishers.

Galetta, D. F., Dolphy, A., El Louadi, M., Lekse, W., Pollalist, Y. & Sampler, J. L. (1993). An Empirical Study of Spreadsheet Error-Finding Performance. *Accounting, Management and Information Technologies*, 3(2), 79-95.

Herreid, C. F. & Schiller, N. A. 2013 Case Studies and the Flipped Classroom, *Journal of College Science Teaching*, 42(5), 62-65.

Janvrin, D. & Morrison, J. (1996). Factors Influencing Risks and Outcomes in End-User Development, *Proceedings of the 29th Annual Hawaii International Conference on System Science*, 346-355.

Johannessen, T. V., Fuglseth, A. M. (2017). Bachelor students' spreadsheet skills and consequences for teaching in business schools, *NOKOBIT, Bibsys Open Journal System*, ISSN 1894-7719, 25(1).

Morgan Chase & Co., Management Task Force (2013). Report of JP Morgan Chase & Co. Management Task Force Regarding 2012 CIO Losses. 124.

Kreye, J. & Pendley, J. (1998). Spreadsheet safety: Designing the perfect file. *Journal of Accountancy*, 186(5), 31-51.

Kruck, S. & Sheetz, S. (2001). Spreadsheet accuracy theory. *Journal of Information Systems Education*, 12, 93–108.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participantion*. New York, NY: Cambridge University Press.

Madahar, M. (2011). *Spreadsheet use for strategic decision-making: An analysis of spreadsheet use and associated risk*. Cardif: Cardiff School of Management, University of Wales Institute Cardiff.

Panko, R. & Halverson, R. (1996). Spreadsheets on trial: a survey of research on spreadsheet risks, *Proceedings of the 29th Annual Hawaii International Conference on Systems Sciences*, 326–335.

Panko, R. (1998). What we know about spreadsheet errors. *Journal of End-User Computing*, 10, 15–21.

Panko, R. (2006). Spreadsheets and Sarbanes-Oxley: Regulations, risks, and control frameworks. *Communications of the Association for Information Systems*, 17(1), 29.

Panko, R. & Aurigemma, S. (2010). Revising the Panko-Halverson taxonomy of spreadsheet errors. *Decision Support Systems*, 49, 235–244.

Powel, S. G., Baker, K, R. & Lawson, B. (2008). A critical review of the literature on spreadsheet errors. *Decision Support Systems*, 46, 128-138.

Presthus, W. (2014). Breakfast at Tiffany's: The Study of a Successful Business Intelligence Solution as an Information Infrastructure. *Proceedings of the European Conference on Information Systems (ECIS) 2014, Tel Aviv, Israel*.

Savery, J. R. & Duffy, T. M. (2001). *Problem*

Based Learning: An instructional model and its constructivist framework. CRLT (Center for Research on Learning and Technology) (Technical Report No. 16-01). Bloomington, IN: Indiana University.

Srinivasan, M., Wilkes, M., Stevenson, F., Nguyen, T., & Slavin, S. (2007). Comparing problem-based learning with case-based learning: Effects of a major curricular shift at two institutions. *Academic Medicine*, 82(1), 74-82.

Wenger, E. (1998). *Communities of practice: learning, meaning and identity*. New York, NY: Cambridge University Press.