BUILDING STATISTICIANS AT AN EARLY AGE – STATISTICAL PROJECTS EXPLORING MEANINGFUL DATA IN PRIMARY SCHOOL

DANIEL FRISCHEMEIER
University of Paderborn
dafr@math.upb.de

ABSTRACT

Since many decisions in politics, economics, and social sciences are based on statistics, statistical literacy is a key component for an active citizen. Statistical skills and statistical reasoning can already be enhanced in primary school. Therefore, not only in Germany, statistics is an element of the primary-school curriculum, which usually includes issues of posing questions, collecting data, and of drawing displays such as pie charts or bar graphs. But to engage students and to start building future generations of statisticians at an early stage, more challenging incentives should engage students in meaningful statistical inquiries. In this paper, we describe activities that are part of a larger, long-term project that investigates how and if teaching units can be designed for primary schools and what their effect on students at that early age is. We focus on specific activities that have been tested for students at age 10. The empirical results show that – beyond learning techniques and the use of software (TinkerPlots) – the students acquired a way of statistical reasoning about the posed problems. Besides the cognitive development, we also documented the affective impact of our course on the young learners, which was very good. Our studies corroborate that it is possible to lead students to sophisticated statistical activities such as comparing groups with the help of suitable software. Furthermore, a positive attitude towards statistics could be observed as a result of the course, which may be the basis for further success in learning statistics. All in all, we can pave the way to the future generation of statisticians already at an early age in primary school.

Keywords: *Primary school, Statistical projects, Group comparisons, Statistical inquiries with software*

1. INTRODUCTION

In a world of the omnipresence of data and in the era of Big Data, a competent use of statistics is inevitable because many decisions in politics, economics and social sciences are based on statistics and data analysis (Engel, 2017). Data Science nowadays permeates all areas of life and is a powerful instrument in, e.g. industrial and marketing processes. For doing Data Science, as can be read in Manyika et al. (2011), people with deep analytical skills in the sense of competent statisticians, are needed. That is why statistics is an important topic, which should be preferably taught as early as possible to build future generations of statisticians. Burrill and Biehler (2011) have identified fundamental

Statistics Education Research Journal, 19(1), 39–56, http://iase-web.org/Publications.php?p=SERJ

ideas in statistics such as data, representation, and variability. These ideas can already be touched very early in education starting from primary school (Hasemann & Mirwald 2012) so that they can be developed later in a more sophisticated and formal manner.

With the intention to develop statistical reasoning of primary school students, we have designed a teaching unit in which young learners work on statistical projects exploring meaningful data. The design, realisation, and the further development of the teaching unit are embedded in a design-based research setting (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). Design-based research has the intention to connect the processes of teaching and research and to develop a teaching unit in a cyclic approach. The idea is to specify learning goals and contents, develop the design, then to conduct and analyse the design experiments, and finally extend local theories on teaching and learning processes (Prediger & Zwetschler, 2013). The first cycle of this approach in the sense of a pilot study has been realised in 2016 in the frame of a bachelor thesis (Breker, 2016). In this cycle, we gained first experiences of how to design a teaching unit in the primary-school setting. In this article, we present the second cycle of this project, which deals with exploring real data with TinkerPlots in ten lessons in a Grade-4 class. This unit has been jointly designed by Christina Schäfers and the present author in the frame of a further bachelor thesis (Schäfers, 2017). A third cycle is currently on-going. The cycles have the intentions to develop, test, and revise teaching units in the frame of design-based research.

Our project *Statistics with Real Data in the Primary School Using TinkerPlot*s has the aim to building statisticians at an early age. In this paper, we concentrate on group comparisons as an exemplary statistical activity, which we also taught in class. Group comparisons include many of the fundamental ideas listed in Burrill and Biehler (2011); they offer fruitful statistical activities already for primary-school level. Furthermore, the comparison of groups and its generalization with regard to inferential statistics is a fundamental task for professional statisticians. We describe activities that may lead young students to explore real data with digital tools and we report on results from accompanying empirical studies on these activities.

The goal of our empirical studies and the present paper is twofold: The first aim is to investigate whether it is possible to teach group comparisons in statistical projects with real and meaningful data by learning the required statistical techniques in primary school – at least to get an idea of the comparison of groups without too many mathematical details and rather learn to use the software that would produce the diagrams on which the comparisons are based. The second goal is to monitor the affective impact of such a teaching unit on young students. It is important to facilitate learning and to get a positive stance to learning it including the way to learn it (the kind of posing questions, the kind of using graphs, and the kind of using software that supports the process of drawing such graphs).

This article is meant to be a success story, which shows that it is possible to introduce young learners to basic statistical concepts and fundamental activities such as group comparisons and to lay foundations for future generations of statisticians at an early age. It offers insights how to design and implement activities to enhance statistical reasoning at primary school and also reports on first results of experimental classes.

2. COMPARING GROUPS AND EXPLORING DATA WITH DIGITAL TOOLS

We will at first distinguish the several terminologies “statistical literacy”, “statistical thinking” and “statistical reasoning”. In Ben-Zvi and Garfield (2004, p. 7), these three terms are distinguished in the following way:

> „Statistical literacy includes the skills that might be used to understand statistical information or research results. Statistical reasoning is the way in which people reason with statistical ideas and make sense out of statistical information. Statistical thinking involves an understanding of why and how statistical investigations are conducted and the “big ideas” that underlie statistical investigations.”

For a more detailed and further differentiated view on statistical thinking as contrasted to statistical reasoning, see also Batanero & Borovcnik (2016). In this paper, we concentrate mainly on the development of statistical literacy and statistical reasoning of young learners.

Comparing groups, in the sense of comparing two distributions of a numerical variable (for details see Biehler, Frischemeier, Reading, & Shaughnessy, 2018, pp. 163) can be seen as a fundamental activity since it involves fundamental statistical ideas (as identified by Burrill & Biehler, 2011) such as data, variation, and representation. There are arguments pro and contra the comparison of distribution for the introduction of statistics teaching. Watson and Moritz (1999) argue for the later stage of statistical inference that the comparison of two groups is easier than inference about one group. Transferring this to the early introduction of descriptive statistics at the primary level, our view to put substantial weight on group comparisons in teaching is justified.

Borovcnik (2020) argues that statistical (and probabilistic) concepts are easier to understand if the purpose of these concepts is made explicit through appropriate tasks. In a way, one might argue that the task of comparing two different groups with respect to a specific variable is more natural than to describe the distribution of a variable within one group. The purpose of comparing makes it clear that we should have some criteria and statistical figures for such a comparison. (centre, spread, skewness, or shift; see, e.g., Frischemeier, 2017) or visual properties of a diagram of the distribution. Such a decision to focus on group comparisons is also supported by Konold and Higgins (2003) who consider group comparisons as a fundamental activity, which can also be used at a very early stage to introduce to statistical reasoning.

Group comparisons also bring it to a point that the prevalent local considerations of a distribution (such as specific single values that are extreme, a clustering of values somewhere within the range of data accompanied by gaps with no data, or an extreme skewedness, etc.) are not helpful for a global, a statistical consideration (that uses ideas such as centre, spread, shift, and range between minimum and maximum). With regard to the interpretation of minimum and maximum, we consider them as local properties of a distribution if the minimum or the maximum is very far off the other data.

To shift the focus of learners from a local to a global view on a dataset and to support learners to compare two groups (e.g., given as stacked dot plots), Konold (2002) and Bakker (2004) suggest to use modal clumps. Konold (2002, p. 1) describes modal clumps in the following way:

> “[Modul clumps are defined as] a range of data in the heart of a distribution of values […] these clumps appear to allow students to express simultaneously what is average and how variable the data are […]. modal clumps may provide useful beginning points for explorations of more formal statistical ideas of center”.

To go a step further, hat plots (Watson, Fitzallen, Wilson, & Creed, 2008) can help learners to compare the spread (in the sense of the middle 50%) of two distributions. Hat plots consist of two brims and a crown:

> "[…] the brim is a line that extends to the range for each group; the crown is a rectangle that, […] shows the location of the middle 50% of the data – the Interquartile Range (IQR)" (Konold 2002, p. 1).

These features make modal clumps a suitable tool to introduce even young learners to the concepts of centre and spread. Hat plots can help learners to understand the notions of spread and shift. Modal clumps and hat plots enhance understanding the concepts of centre and spread and support identifying a shift between two distributions.

To analyse real and meaningful datasets, adequate software (Biehler, Ben-Zvi, Bakker, & Makar, 2013) is inevitable if one wants to realise the idea of transnumeration (Wild & Pfannkuch, 1999) and explore data with regard to specific statistical questions. TinkerPlots (Konold & Miller, 2011) is a data-analysis software, which can already be used in primary school. The main idea of the software is that the data are stored in form of data cards and that the construction of statistical displays is realised via the data card-operations stack, separate, order, and tools such as the drawing tool or hat plots. Young students can first analyse data with data cards in hands-on activities in small datasets and then apply the same operations on larger datasets in TinkerPlots. In our teaching experiments, TinkerPlots proved to be an adequate tool for exploratory data analysis already at an early stage. For further details on TinkerPlots, see Konold (2007), or Biehler, Ben-Zvi, Bakker, and Makar (2013).

3. RESEARCH METHODOLOGY

Our main goal is to enhance statistical reasoning of young learners at a primary school level and to design learning trajectories for primary-school students to explore real and meaningful data and to conduct activities such as group comparisons. As we have described in the introduction, we have designed a teaching unit to develop statistical reasoning in primary-school students, which should lead them to conduct group comparisons with digital tools in a design-based research setting (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). We concentrate on the teaching unit in Cycle 2, the second version of our teaching unit here. The second cycle focuses on both the cognitive and the affective level. For the cognitive impact on the young learners, we pursue the following aims:

- to show that it is possible to teach group comparisons successfully at primary level;
- to show that it is possible to introduce TinkerPlots successfully and use it for the diagrams to perform group comparisons at primary level;
- to show that the PPDAC cycle is a useful framework for applied project work also at primary level.

For the affective aspects, the second cycle has the aim to investigate whether the students accepted or liked the various components of the course.

3.1 RESEARCH QUESTIONS

According to our general goal of introducing statistical ideas at an early age and our focus on the cognitive and affective impact in the second cycle of the teaching unit, we formulate two research questions related to our course:

(1) In which way does statistical reasoning with regard to group comparisons of the students improve after the course?

(2) In which way does the course have an impact on the attitudes of the participants?

3.2 METHOD

The cognitive impact of the course was measured in two ways: we analysed the students' documents and posters produced in the teaching unit and the achievement by one specific item in a pretest/posttest scenario. For the analysis of the documents, we used the approach of Bohl (2009). Pretest and posttest were analysed with a qualitative-content-analysis approach (Mayring, 2015). The affective impact was measured by a 3-point-scale in a pre/post survey scenario.

3.3 DATA COLLECTION

For the accompanying research, we have to differentiate between the two different targets of the investigation, namely the cognitive and the affective impact of the course on the students. For the process and the steps in understanding the task of group comparisons, we have collected exercise sheets and TinkerPlots files and written field notes as observations from the lessons from the students. The field notes of the lessons served as an important data base for the discussion of the revision of the lesson between the teacher and the researcher with regard to a re-design of the teaching unit.

In addition, we have collected the posters as final products of the data-analysis projects of our students. For students' direct progress in the experimental teaching, we have used one specific task of comparison, and applied a pretest-posttest comparison of the achievement of the students. We collected all students' notes of pre- and posttest. For the affective impact of the experimental class, we administered an evaluation at the end of the teaching unit, using a three-point scale of qualitative judgment again in comparison before and after the course.

3.4 PARTICIPANTS

Twelve Grade-4 students (seven male, five female, aged 10–11) from a primary school in a rural area of North Rhine-Westphalia in Germany, have participated in the teaching unit. For the data analysis on the cognitive and the affective aspects, all twelve students have been selected. All students had little knowledge of statistics before participating in our experimental class: in Grade 3, they have collected data, displayed them in tallies and pie graphs, and they have dealt with reading pie graphs. However there cannot be made further assumptions and conclusions on the mathematical abilities of the participants because we did not conduct a standardised test in this respect.

4. EXPLORING REAL DATA BY DIGITAL TOOLS IN PRIMARY SCHOOL: A BEST PRACTICE EXAMPLE

The major idea of the teaching unit was to enhance statistical reasoning of primary-school students; we introduced them into data analysis and group comparisons with TinkerPlots to enable them to conduct statistical projects and to explore real and meaningful data. One basic idea of the design of the teaching unit was that the students should get acquainted to different phases in a data-analysis process Problem, Plan, Data, Analysis, and Conclusions (see also Wild & Pfannkuch, 1999), and orientate their work according to these phases. These phases are described in the PPDAC cycle of Wild and Pfannkuch (1999) and are also mentioned in the requirements of primary-school students in statistical reasoning (Hasemann & Mirwald, 2012). There are other cycles (with roots in mathematics education going back to the 1980s (see, e.g., Blum & Kirsch, 1989; Blum & Leiss, 2007) or with roots in quality control of the 1940s (see, e.g., Tsubaki, 2018), which may have a stronger emphasis on modelling which is also a fundamental aspect in data analysis. For our purposes and for learning and teaching data analysis in primary school, we decided to use the PPDAC cycle. As a further source for the design of our teaching unit, we used a guideline on principles of Statistical Reasoning Learning Environments (Garfield & Ben-Zvi, 2008). Garfield and Ben-Zvi (2008, p. 48) suggested the following ideas on the design of statistical-reasoning learning environments (SRLE):

"1. Focuses on developing central statistical ideas rather than on presenting a set of tools and procedures.
2. Uses real and motivating data sets to engage students in making and testing conjectures.
3. Uses classroom activities to support the development of students' reasoning.
4. Integrates the use of appropriate technological tools that allow students to test their conjectures, explore and analyze data, and develop their statistical reasoning.
5. Promotes classroom discourse that includes statistical arguments and sustained exchanges that focus on significant statistical ideas.
6. Uses assessment to learn what students know and to monitor the development of their statistical learning as well as to evaluate instructional plans and progress."

In our teaching unit we used real data, we used digital tools (TinkerPlots) for analysing the data, we implemented peer and group collaboration in class, and we focused on central statistical ideas. In the following, we describe the content of the teaching unit and then analyse its cognitive and affective impact.

4.1 THE TEACHING UNIT

The teaching unit consisted of ten lessons (of 45 minutes each). The sequence of the lessons and the contents of each lesson can be seen in Table 1. At the beginning of the teaching unit, our Grade-4 students were introduced into the basics of descriptive statistics and then continued to work in peers and groups on little statistical projects like "My class in numbers". In a next step, the students were introduced into the different phases *Plan, Problem, Data, Analysis, and Conclusions*. In these phases, the students have generated their own statistical questions, collected data of their classmates and schoolmates, and were then introduced to the analysis of the class data (drawing by hand and reading bar graphs, pie charts, and stacked dot plots). Then, in a further step, they were introduced to use TinkerPlots to be enabled to explore larger datasets such as the dataset of their entire school.

Table 1. Overview of the lessons of the teaching unit (Schäfers, 2017)

Lesson	Content of the lesson
	Introduction empirical work
1	Getting to know about variables, values, etc.
2	Data analysis cycle (PPDAC), learning to generate statistical questions
3	Collecting data
	Analysing data about the class
4	Drawing and reading bar graphs and pie charts
5	Drawing and reading stacked dot plots
	Introducing TinkerPlots and working with this software
6	Analysing larger univariate datasets
7	Group comparisons I
8	Group comparisons II
	Project work
9	Group comparison project, preparing posters
10	Poster presentation of the results of the group-comparison project

Significant in the introduction was that the students have been introduced to data-analysis concepts and representations on different representation levels (enactive, iconic, and symbolic). Major parts of the introduction were teacher-centred, but there were also interactive parts when the students worked with data cards and small datasets and then applied their data-card operations (stack, separate, and order) within small datasets to larger data sets using the same operations in TinkerPlots. The students worked in pairs; in the working phase, the teacher went around and supported them when they had technical problems with the software. All in all, the students did not need much help and were able to proceed on their own.

Also, the idea of modal clumps and hat plots has been introduced first within small data sets on the board (the teacher demonstrated how to identify the modal clump in a distribution) and later with larger datasets using the drawing tool and the hatplot feature in TinkerPlots. The teacher introduced the median on an enactive level and asked the students to row up by their height. Then the teacher discussed with them interactively the location of the median in the setting of an odd and an even number of students. In a next step, the teacher used TinkerPlots to identify the median as a formal method to characterise the centre of a distribution in the modal clumps. We did not introduce the arithmetic mean as a measure of centre because we wanted our students to use the steps of modal clumps, medians, and hatplots in the reading and interpreting of distributions.

As data set for the introduction to group comparisons, we used a data set containing the data of pack weights of school students across different grades (Grades 1 and 3). Together with our students we investigated whether Grade-three students tend to have heavier backpacks than first graders. This introduction to group comparisons was realised in Lessons 7 and 8 in four steps (1)-(4) in a teacher-centred way using TinkerPlots:

(1) Draw a two-stacked dot plots for the group-comparison (see Figure 1);
(2) Use the drawing tool to mark the modal clumps in both distributions (Figure 2):
(3) Calculate the median in both distributions (Figure 3);
(4) Display hat plots in both distributions (Figure 4).

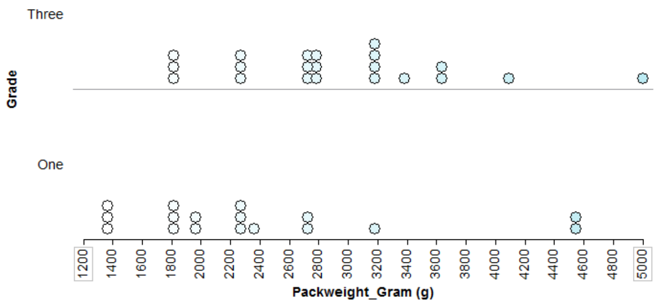

Figure 1. Group-comparison by stacked-dot plots in TinkerPlots

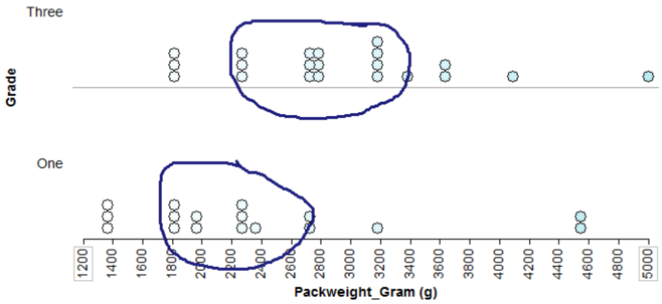

Figure 2. Group comparison by stacked dot plots and modal clumps in TinkerPlots

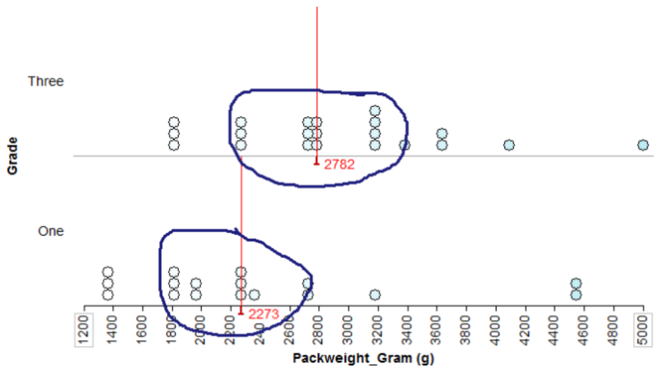

Figure 3. Group comparison by stacked dot plots, modal clumps, and medians

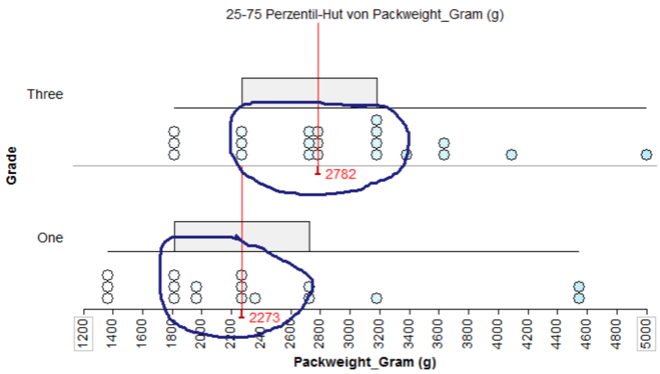

Figure 4. Group comparison: stacked dot plots, modal clumps, medians, and hatplots

In the group-comparison process, the teacher explained that the comparison of the location of the modal clumps (Figure 2) provides a first impression that Grade-3 students tend to have heavier backpacks then Grade-1 students. As a next step (3), the teacher demonstrated how to calculate the median in TinkerPlots. Accordingly, the teacher formulated a more precise comparison statement and stated that the median of PackWeight is approximately 500 g larger in Grade 3 than in Grade 1 (Figure 3). Finally, the teacher used hat plots and showed that the crown of the upper hat plot is shifted more to the right in Grade 3 as compared to Grade 1 (Figure 4). We have to admit that this was a teacher-centred process, and that for a re-design it might be better to realise a more student-centred way to discuss, e.g., the identification of modal clumps, to make sure that the students understand the underlying concepts.

The conclusion of the teacher was that Grade-3 students tend to have heavier backpacks than Grade-1 students. This was corroborated by comparing groups in larger datasets before the students started to work on the projects (see below in the next section) to explore and analyse data (which they have collected in their school) and compare groups on their own (Lesson 9). The wording “tend to have” was used by the teacher in the group-comparison setting of Lesson 9 but there was no class discussion about this expression, so there is no hint in how far the students have understood this wording. In a re-designed teaching unit, it would be appropriate to discuss this wording in class. Finally, the students were asked to summarise their results in form of posters and present these posters to their classmates in the final lesson (Lesson 10).

4.2 STUDENTS’ WORK ON PROJECTS AND POSTERS WITH DIAGRAMS

As mentioned in the previous section, the students work on different topics in groups. One group worked on the “Height differences between boys and girls”, another group worked on the “Shoe size differences between boys and girls” and a third group worked on the “Differences in time playing on the smartphone, tablet (in minutes) between boys and girls”. Thus, all three groups worked on group-comparison activities.

To get an impression on the products of the students in their projects in Lessons 9 and 10, we show the poster of the group “Blue” (Tim, Titus, Elton, and Noel). This poster has been chosen as convenience sample. The posters of the other groups are on a similar level. The investigation deals with the height differences between boys and girls in their school (all students of their primary school have been surveyed). In the poster, they present a graph with stacked dot plots; they identify the modal clumps, calculate the medians, and display the hats of both distributions (see Figure 5).

Underneath the graph in Figure 5, the students have written the report of their investigation:

> “The students of the fourth class of the Alme School Wewer were asked how tall they are. […] Most girls are 140 and 148 cm tall and most boys are between 149 and 155 cm tall. Many values with the boys accumulate at 149-155 cm to a hill. For the girls, the modal clump is between 140-148 cm. The modal clump of the boys is concentrated in higher numbers. That means most [of] the boys are taller than the girls. […] The crown of the hat shows the middle half of the values. The range [of the crown] for boys is 140-152 cm. The girls’ crown is much smaller and stretches from 140 to 148 cm. This means that the girls are smaller than the boys. The median shows the centre of the distribution. The median for boys is 146.5 cm and the median for girls is 144 cm. The mean value of the boys is larger than that of the girls. That means the boys tend to be taller than the girls.”

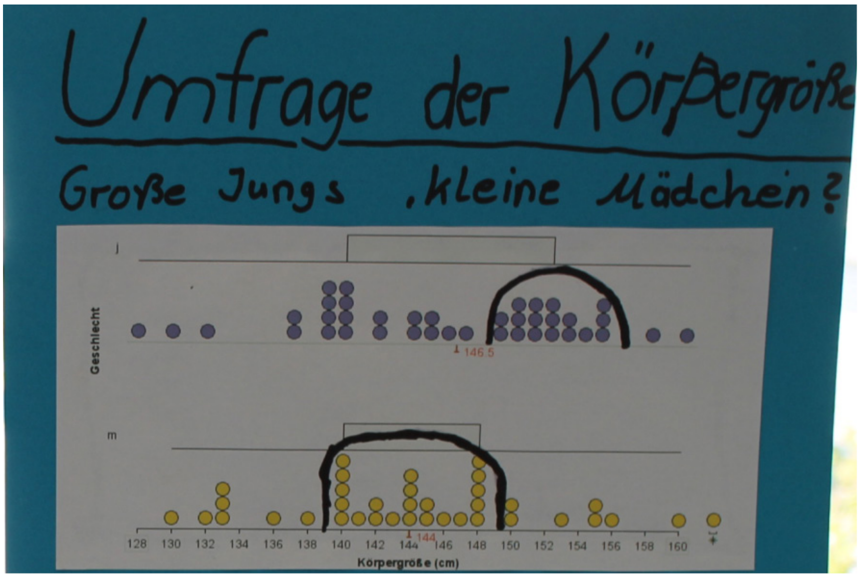

Figure 5. Poster "Survey on Heights – Tall Boys, Little Girls?" from the group Blue

At first the group Blue stated the aim of their data investigation:

> "The students of the fourth class of the Almeschule were asked how tall they are."

They then identified the modes of the distributions and the modal clumps in both distributions:

> "Most girls are 140 and 148 cm tall and most boys are between 149 and 155 cm tall."

Although the modal clump in the distribution of the boys is not identified correctly, they described both modal clumps:

> "For the girls the modal clump is between 140-148 cm. The modal clump of the boys is concentrated in the higher numbers."

They mis-located the modal clumps in the distribution of the boys, which might follow a confirmation bias (the students might think that boys are taller, they identify something – wrongly – that shows that boys are taller than girls). Yet, they made the correct conclusion by comparing the shift of modal clumps:

> "That means most [of] the boys are taller than the girls."

Then they compare the distributions with regard to hat plots and describe their locations:

> "The crown of the hat shows the middle half of the values. The range for boys is 140cm-152cm. The girls' one is much smaller and goes from 140 to 148cm."

Here they do not refer to spread issues but concentrate on the shift of the hats and state that comparing the hats means that

> "the girls are smaller than the boys."

Finally they compare the medians of both distributions to come to the conclusion that

"the boys tend to be taller than the girls."

We have to state that the expression "tend to" was introduced by the teacher and that it is not clear at this stage whether the students have fully understood the situation and the wording "tend to". Nevertheless, in this short description of the result of an exemplary project, we can see that, in their analysis, these students used adequate tools such as modal clumps, medians, and hat plots to compare both groups. Although there are some shortcomings such as the incorrect identification of the modal clump in the distribution of the boys or neglecting the spread differences between both groups when comparing the hats, we can say that the group Tim, Titus, Elton and Noel has made adequate and interesting investigations and showed a sophisticated reasoning when comparing the distribution of the height of boys and girls at Alme School using a software. This example and the posters of the other groups, which are elaborated on a similar level, give clear indication that statistical project work and the elaborated analysis of data using software such as TinkerPlots can be already realised at an early age in primary school.

5. RESULTS OF ACCOMPANYING EMPIRICAL STUDIES

We have collected several data to evaluate this second cycle of the teaching unit (see also Section 3.3): exercise sheets, TinkerPlots files, and posters (final products at the end of the teaching unit); and we have collected written field notes as observations from the lessons. Furthermore, we have conducted a pretest and a posttest to investigate the improvement of the students. An evaluation of the affective impact of the course completed our analyses. We have the aim to investigate the (1) cognitive impact and (2) affective impact of the course. Thus, in the following, we investigate the development of statistical reasoning with regards to the skill of group comparisons in our students before and after the teaching unit. Furthermore, we study the affective impact on the students; that means how they liked the components of the teaching unit.

5.1 IMPROVEMENT OF THE COGNITIVE ACHIEVEMENT

In the pretest and the posttest, we wanted to investigate, in which way the teaching unit (especially Lessons 7–10) enhanced statistical reasoning with regard to comparing groups; the skills are measured by achievement, the capability to reason is checked by the justifications of their solutions. That is why we provided a group-comparison task before attending to the teaching unit and the same task was handed out to them at the end of the teaching unit. As a task, we asked our students to compare the distributions of the variable "pack weight" of first and third graders. The concrete task was (see for the TinkerPlots display in Figure 6):

"Do first graders or third graders tend to have heavier backpacks? Explain your answer."

We used qualitative content analysis (Mayring, 2015) to categorise and explore, which explanations our students used to state that there is a difference between the two groups. In both groups there are 17 values, which makes the task much simpler as counts rather than proportions may be used to solve it.

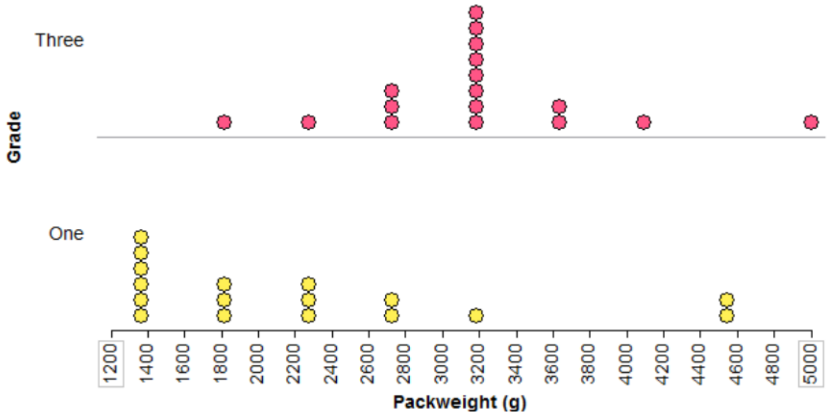

Figure 6. Distribution of the variable PackWeight (g) in Grades 1 and 3

There are several strategies that can be used:

- Total sum (or average): Since the groups are equal-sized, the students could solve it by adding up all values in each group and then compare the sums. Sums are somewhat equivalent to averages of groups, which have the same sample size. In addition, we have to mention that our focus in the teaching unit was on the median, not on the mean or on sums. We call this procedure “total sum” (see, e.g., Liv in Table 2). If the groups have a different size, the strategy has to be adapted to the average, which is more complicated.
- Shift of points: Another possibility could be to identify the shift of the points to the right in the group of third graders in comparison to the group of first graders (exemplary coding: “Because the points are located further behind”, see Maria, Figure 7). We call this kind of comparison strategy “shift of points”.
- Interval strategy: Furthermore students could use cut-off point strategies (count the number of dots in a specific interval and compare third and first graders) to compare both groups. Cut-off points are thresholds (below or above). We call this kind of group comparison strategy p-based (this was not found in our data). However, the strategy “to count the number of dots in a specific interval” can also be misleading in some cases.
- Use statistical concepts: Finally our students could also use more sophisticated concepts (also with a view of comparing un-equal-sized groups in the future) like modal clumps, hat plots or medians to compare both groups (see Figure 8, example of Maria in posttest).

The result of all 12 participants (they have worked on their own) in the pretest and posttest is shown in Table 2. Information is displayed on whether the student’s answer is correct and which explanation they give.

Comparing the correct statements, 8 of 12 students provided a correct statement (66.7%) in the pretest while all 12 students gave a correct answer in the posttest. This at a first sight is not a big improvement. Of course, students should have improved after the teaching unit; yet, the improvement from pretest to posttest becomes clearer by the explanations, the students used. In the pretest, only 4 of 12 (33.3%) students used adequate explanations. All others (66.7%) did not provide any explanation. In the posttest, 11 of 12 students (91.7%) used adequate group-comparison elements (such as modal clumps, medians, or hats) to describe that third graders tend to have heavier backpacks than first graders.

Table 2. Overview of results of pretest and posttest

	Pretest		Posttest	
Name	Correct statement	Explanation	Correct statement	Explanation
Tim	Yes	None	Yes	Median, Hat
Titus	Yes	None	Yes	Hat
Elton	No	None	Yes	Median, Hat, Shift, MC
Noel	No	None	Yes	Median
Maria	Yes	Shift of Points	Yes	Median, Hat
Johannes	Yes	None	Yes	Median
Finnja	No	None	Yes	None
Emilie	No	None	Yes	Hat
Oliver	Yes	Shift of Points	Yes	Hat, Median
Liv	Yes	Total sum	Yes	Hat
Hannah	Yes	Shift of Points	Yes	Hat
Loris	Yes	None	Yes	Hat

This is corroborated by Maria's comparison statement of (Figure 7).

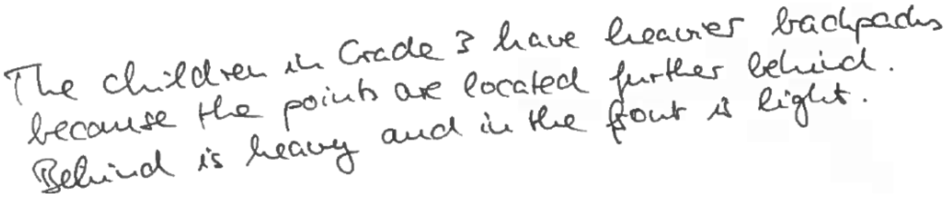

Figure 7. Group-comparison statement of Maria in the pretest – English translation of the handwritten statement of Maria

Maria writes:

> "The children in Grade 3 have heavier backpacks because the points are located further behind. Behind is heavy and in the front is light."

We see that Maria already uses "shift of points" for describing the difference between Grade-1 and Grade-3 students in the pretest. In the posttest, Maria uses more formal arguments such as hat plots (brims & crowns), and medians (see Figure 8).

> "The children in Grade 3 have heavier backpacks. I can see it, because brim and crown are located further behind. The median in the groups of the Grade-3 students [is] 3182 g and in the group of the Grade-1 students [it is] 1818 g."

We see that Maria uses the components of the hats (brims, crown) and the value of the median to make precise the difference between the two groups. All in all, with regard to pretest and posttest, we can say that the quality of explanations of our students has increased considerably.

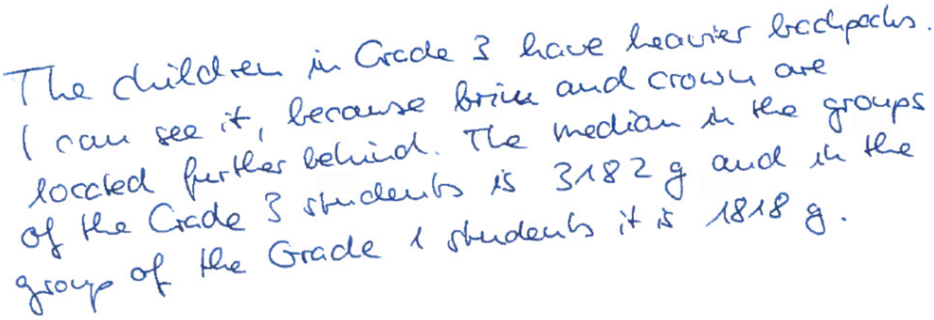

Figure 8. Group-comparison statement of Maria in posttest – English translation of the handwritten statement of Maria

5.2 AFFECTIVE IMPACT OF THE COURSE

We have also asked our students to rate the teaching unit as a whole and we wanted to evaluate the attitudes towards each of these components of our students. Since we want to evaluate primary-school students, we decided not to use a common Likert scale but a 3-step-scale they are used to in school. The laughing smiley is meant to display a positive attitude, the neutral smiley is meant to display a neutral attitude and the negative smiley is meant to display a negative attitude with regard to the specific item. All twelve students have participated in this post-course evaluation. The items of the evaluation scheme and the results of the evaluation can be seen in Table 3.

The main message of the results of the evaluation sheet is that the students liked the project and working with TinkerPlots. As we can see in Table 3, all twelve students liked the project. Furthermore, all students liked the collection of data and the basics of data analysis. Eleven of the twelve students liked comparing groups with TinkerPlots and also the work on the posters. Ten of the twelve students liked working with the drawing tool and with hats in TinkerPlots and the presentation of the posters. The only component which was not evaluated equally positive was the drawing of stacked dot plots (only 8 participants liked this component).

Table 3. Students' attitudes towards specific components of the teaching unit

Data analysis with TinkerPlots

Please choose:

	☺	😐	☹	No answer
I liked the project work.	12	0	0	0
I liked learning about the basics of data analysis.	12	0	0	0
I liked the collection of data.	12	0	0	0
I liked drawing stacked data plots.	8	3	0	1
I liked TinkerPlots.	12	0	0	0
I liked comparing groups with TinkerPlots.	11	1	0	0
I liked working with the drawing tool and the hats in TinkerPlots.	10	2	0	0
I liked the work on the posters.	11	1	0	0
I liked the presentation of the posters.	10	2	0	0

All in all, we can say that the primary students liked applying the PPDAC cycle and to collect their "own" data. They also liked exploring their "own" data and conducting group comparisons with TinkerPlots. So this is an indication that this teaching unit can also help to enhance a positive stance towards statistics for primary-school students to bring them to the next level and to prepare their stage to become a part of a future generation of statisticians.

6. CONCLUSION AND DISCUSSION

The goal of our project on *Statistics with Real Data in the Primary School Using TinkerPlots* is to lay foundations for future generations of statisticians. This education in statistics should preferably start as early as possible, that means, in primary school. We have designed and implemented a teaching unit with ten lessons, which enhances statistical reasoning and lead primary-school students to compare groups in large and real datasets using TinkerPlots. In this paper, we have described this teaching unit and the rationale behind. The new aspects are exploring real data and using software such as TinkerPlots. Furthermore, the primary-school students have been introduced to the comparison of groups, a fundamental activity in statistics.

The final products – the posters of the teaching unit – indicate a gain in cognitive capacity; after the teaching unit, the primary-school students seem to be capable of analysing data in the context of a real-life problem with TinkerPlots and conduct fundamental statistical activities including group comparisons by means of simple statistical tools that are suitable already at this early age (modal clumps or hat plots).

The analysis of pretest and posttest shows that there is already a basis to compare two groups before any formal teaching, since 66.7% of our students provided a correct comparison statement already in the pretest. This also shows that the children are ripe to be introduced to the topic of group comparisons. Yet, they failed to provide explanations for the perceived differences between the two groups. With a look at the results of the posttest, we find empirical evidence that the teaching unit on concepts such as modal clumps and hat plots helped our students to use adequate tools to compare the groups and to identify the differences in a more precise manner. Furthermore, whereas in the pretest only four students gave an explanation for the group difference, in the posttest all but one student used adequate explanations (modal clumps, medians, hats) to describe the differences between the groups.

The evaluation of the affective impact of our experimental teaching shows that our students overall liked the teaching activities; all students liked the project work. Also, 11 of 12 students specifically liked the activity of comparing groups with TinkerPlots.

In summary, the accompanying empirical study of our teaching unit corroborates that statistical project work and an elaborated analysis of data using software such as TinkerPlots can be realised already at an early stage. For the transition to the secondary school and later, teachers may pick up and build on the concepts of modal clumps and hat plots and lead students to more formal concepts such as boxplots and quantiles to compare groups.

The teaching unit presented here can be seen as an example how to design and implement activities to enhance statistical reasoning, even as early as in the primary school. The design and the material can be adapted to other activities such as exploring the relationship between two categorical variables or exploring the relationship between two numerical variables in order to enhance other statistical ideas as well. One experience we want to share is that the PPDAC cycle of modelling real and meaningful data and analysing them with digital tools such as TinkerPlots is a key element of the design of such a course.

For future research and for the further design of activities, we plan to conduct a qualitative study to get more insight into the cognitive *processes* of primary-school students when comparing groups with TinkerPlots, in order to be able to develop improved materials and tasks to support statistical reasoning of young students and to ground in them a positive stance towards statistics so that "the way for their future careers as statisticians is paved."

ACKNOWLEDGEMENTS

The author is very grateful to Rebecca Breker and Christina Schäfers who have collaborated with him in the design of the teaching units and who have realised the experimental teaching of the statistical projects in the class. In addition, the author is also very grateful to the editors Manfred Borovcnik, Ayse Bilgin, Peter Howley, and the reviewers for very helpful und constructive comments and annotations on a former version of this article.

REFERENCES

Bakker, A. (2004). *Design research in statistics education–On symbolizing and computer tools* (Unpublished doctorial dissertation). University of Utrecht, The Netherlands.

Batanero, C. & Borovcnik, M. (2016), *Statistics and probability in high school*. Rotterdam: Sense Publishers.

Ben-Zvi, D., & Garfield, J. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht: Kluwer Academic Publishers.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clements, A. J. Bishop, C. Keitel-Kreidt, J. Kilpatrick, & F. K.-S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643–689). New York: Springer.

Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, M. (2018). Reasoning about data In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 139–192). Cham, Switzerland: Springer International.

Blum, W. & Kirsch, A. (1989). The problem of the graphic artist. In W. Blum, J. S. Berry, R. Biehler, I. D. Huntley, G. Kaiser-Messmer, & L. Profke (Eds), *Applications and modelling in learning and teaching mathematics.* (pp. 129–135). Chichester: Horwood.

Blum, W., & Leiß, D. (2007). How do students and teachers deal with modelling problems. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modeling: Education, engineering, and economics.* (pp. 222–231). Chichester: Horwood.

Bohl, T. (2009). *Prüfen und Bewerten im Offenen Unterricht*. Weinheim und Basel: Beltz.

Borovcnik (2020). Contexts, analogies, and tasks that expose the purpose of the key concepts of probability. *Zetetiké, 28*(1), 1–25.

Breker, R. (2016). Design, Durchführung und Evaluation einer Unterrichtseinheit zur Entwicklung der Kompetenz „Verteilungen zu vergleichen“ in einer 4. Klasse unter

Verwendung der Software TinkerPlots und neuer Medien. (Design, implementation and evaluation of a teaching unit for the development of the competence "Comparison of distributions" in a 4th grade class using the TinkerPlots software and new media) University of Paderborn: Unpublished Bachelor thesis.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education* (pp. 57–69). Netherlands: Springer.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal, 16*(1), 44–49.

Frischemeier, D. (2017). *Statistisch denken und forschen lernen mit der Software TinkerPlots*. Wiesbaden: Springer Fachmedien.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning. Connecting research and teaching practice*. The Netherlands: Springer.

Hasemann, K., & Mirwald, E. (2012). Daten, Häufigkeit und Wahrscheinlichkeit. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer, & O. Köller (Eds.), *Bildungsstandards für die Grundschule: Mathematik konkret* (pp. 141–161). Berlin: Cornelsen Scriptor.

Konold, C. (2002). *Hat plots?* Unpublished Manuscript. University of Massachusetts.

Konold, C. (2007). Designing a data tool for learners. In M. Lovett & P. Shah (Eds.), *Thinking with data: The 33rd Annual Carnegie Symposium on Cognition* (pp. 267–292). Hillside, NJ: Lawrence Erlbaum Associates.

Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–215). Reston, VA: National Council of Teachers of Mathematics.

Konold, C., & Miller, C. (2011). TinkerPlots 2.0. Emeryville, CA: Key Curriculum Press.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). *Big data: The next frontier for innovation, competition, and productivity*. [Online: www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation]

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education* (pp. 365–380). Dordrecht: Springer.

Prediger, S. & Zwetzschler, L. (2013). Topic-specific design research with a focus on learning processes: The case of understanding algebraic equivalence in grade 8. In T. Plomp & N. Nieveen (Eds.), *Educational design research: Illustrative cases* (pp. 407–424). Enschede: SLO, Netherlands Institute for Curriculum Development.

Schäfers, C. (2017). Durchführung und qualitative Evaluation einer redesignten Unterrichtsreihe zur Entwicklung der Komptenz "Verteilungen zu vergleichen" in einer Jahrgangsstufe 4 unter Verwendung der Software TinkerPlots (Implementation and qualitative evaluation of a redesigned teaching series for the development of the competence "Comparison of distributions" in a grade 4 class using the software TinkerPlots). University of Paderborn: Unpublished Bachelor thesis.

Tsubaki, H. (2018, July). *Education of the grammar of science for sciences and society*. Plenary lecture at the Tenth International Conference on Teaching Statistics (ICOTS 10). 8–13 July, Kyoto, Japan.

Watson, J., Fitzallen, N., Wilson, K., & Creed, J. (2008). The representational value of HATS. *Mathematics Teaching in Middle School, 14*(1), 4–10.

Watson, J. M., & Moritz, J. B. (1999). The beginnings of statistical inference: Comparing two data sets. *Educational Studies in Mathematics, 37*(2), 145–168.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223–248.
[Online: doi.org/10.1111/j.1751-5823.1999.tb00442.x]

DANIEL FRISCHEMEIER
University of Paderborn, Institute of Mathematics
Warburger Straße 100
33098 Paderborn, Germany