UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Developing early statistical thinking with primary school students and early inferential thinking with preservice teachers for primary school

Special lecture at JCOTS 2020
Institute of Statistical Mathematics, Tachikawa

Dr. Daniel Frischemeier
University of Paderborn

Overview of my lecture

1. Developing thinking about data of primary school students using TinkerPlots
2. Developing thinking about uncertainty of primary school students using TinkerPlots
3. Developing thinking about uncertainty of preservice teachers for primary school (Randomization tests with TinkerPlots)

Part 1

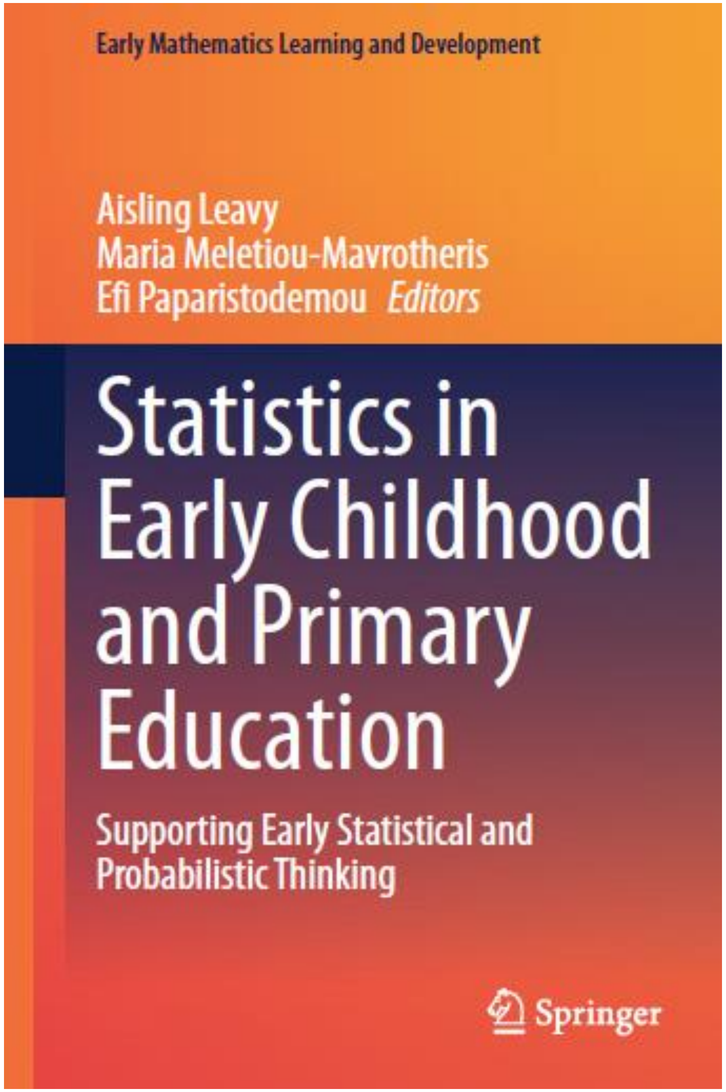
Developing reasoning about data of primary school students using TinkerPlots

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

The need for Early Statistical Thinking

„Today‘s students need to learn to work and think with data and chance from an early age, so they begin to prepare for the data-driven society in which they live.“ (Ben-Zvi, 2018, vii)

Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2018). *Statistics in Early Childhood and Primary Education: Supporting Early Statistical and Probabilistic Thinking*. Singapore: Springer.



Starting point

Proposals to promote statistical thinking at primary level

- Support students in generating adequate statistical questions (u.a. Allmond & Makar, 2010; Arnold, 2013)
- Exploring data with digital tools (u.a. Konold & Higgins, 2003; Harradine & Konold, 2006; Konold, 2007)
- Working with meaningful data (u.a. Leavy & Hourigan, 2018)
- Developing global perspectives on distributions (u.a. Bakker, 2004; Bakker & Gravemeijer, 2004; Konold et al., 2015)
- Comparing groups with regard to preformal concepts like modal clumps (u.a. Konold et al., 2002; Bakker, 2004; Fielding-Wells, 2018; Allmond & Makar, 2018)
- Conducting real statistical projects, experiencing data analysis cycle (z.B. Wild & Pfannkuch, 1999)

Paradigm in statistics education – also in primary school

Data analysis cycle like PPDAC-Cycle (Wild & Pfannkuch, 1999)

Different phases of PPDAC

• **P**roblem	Developing statistical question
• **P**lan	Plan of data collection
• **D**ata Collection	Data collection (survey, observation, experiment)
• **A**nalysis	Data analysis and exploration

Is it possible to go beyond this in primary school?

for example

Exploring large and real datasets? (u.a. Garfield & Ben-Zvi, 2008)

Using digital tools for data exploration? (u.a. Konold, 2006)

→ Conducting statistical projects with real and meaningful data?

Learning data analysis with digital tools

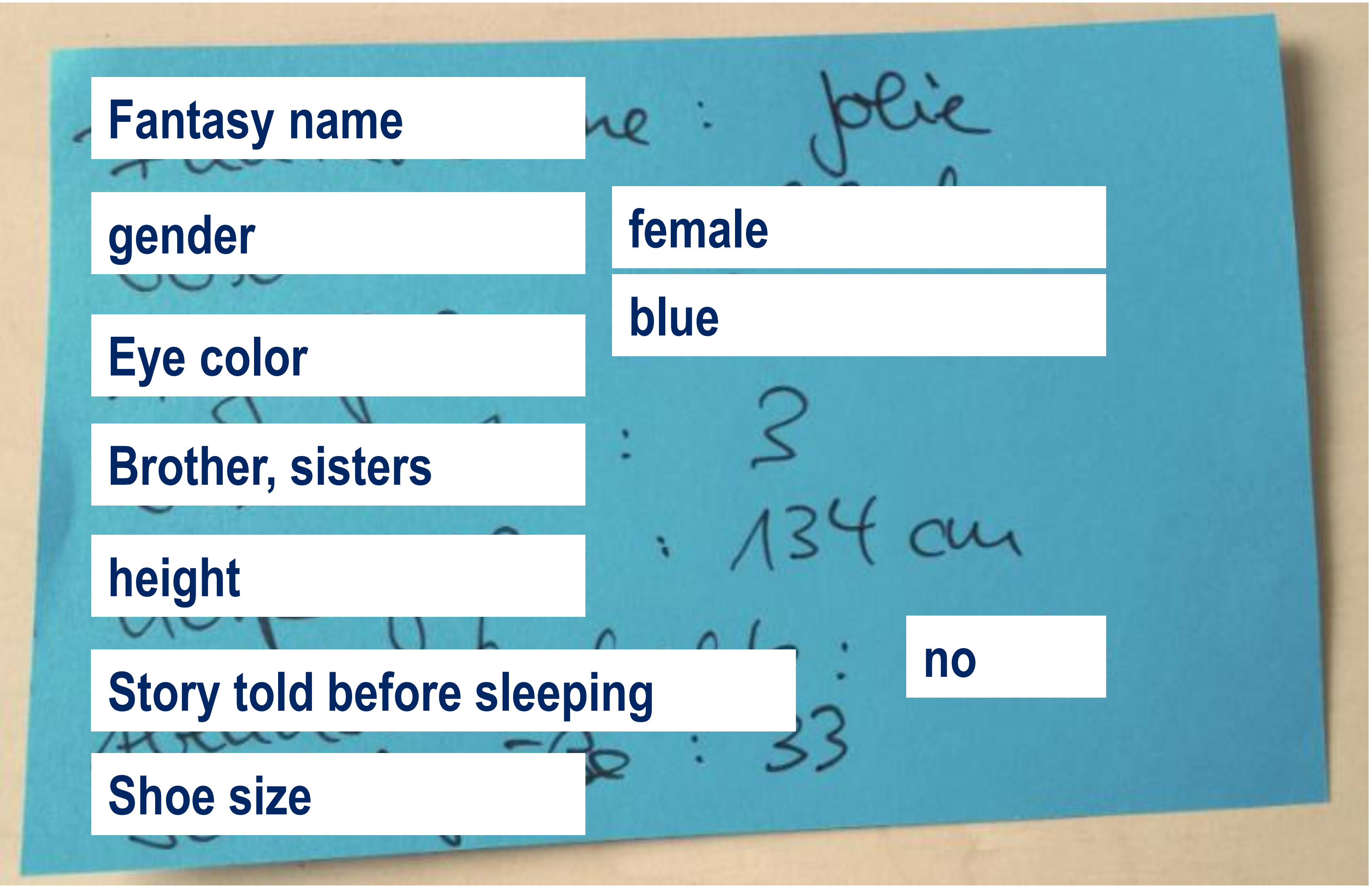
1. Building understanding for fundamental data operations (like Separate and Stack) on the example of small data sets

→ **Data analysis with Datacards** (c.f. Harradine & Konold, 2006)

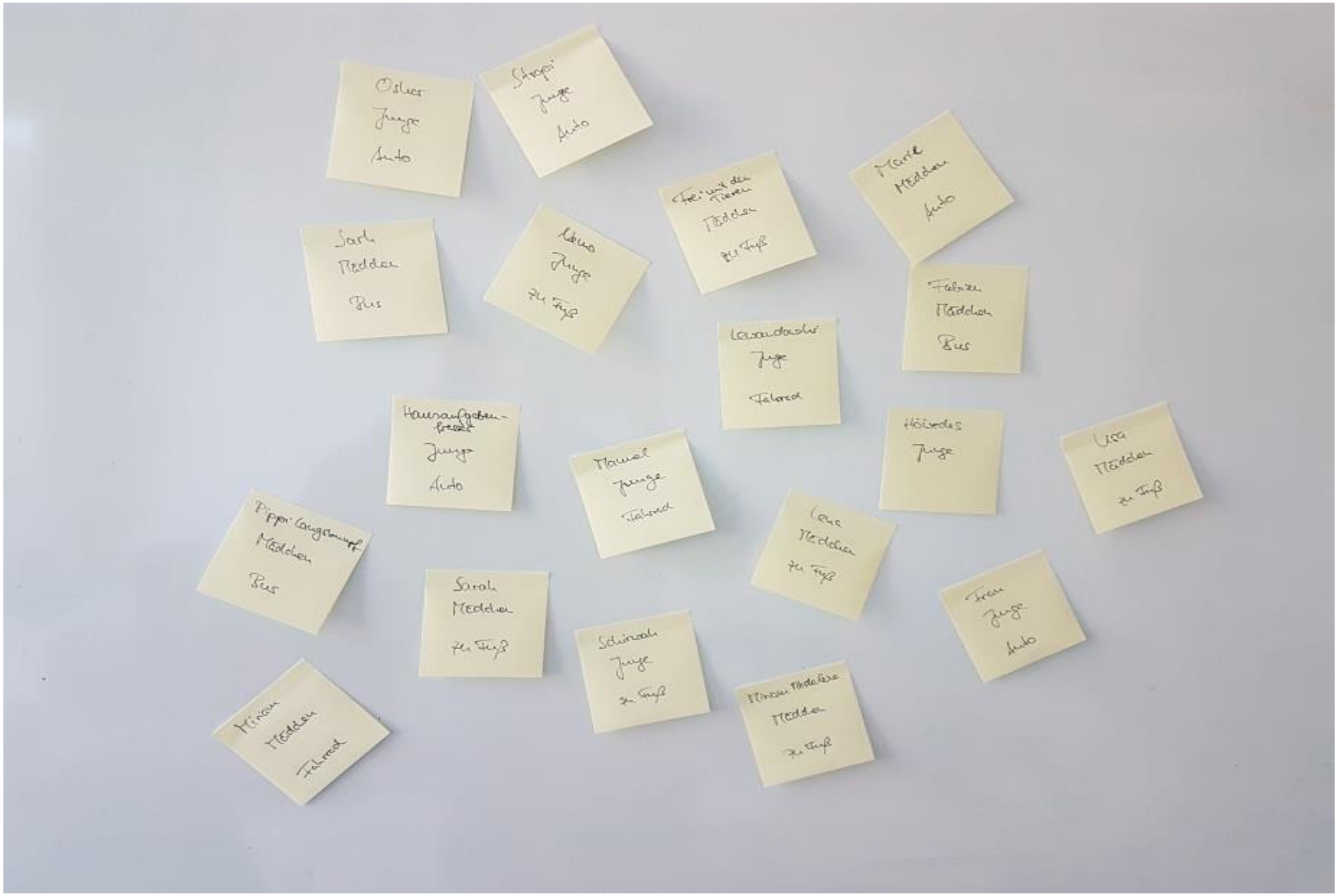
2. Transferring activities from 1. to the analysis of larger datasets

→ **Data analysis with TinkerPlots** (c.f. Konold, 2006)

Data cards (Biehler & Frischemeier, 2015)

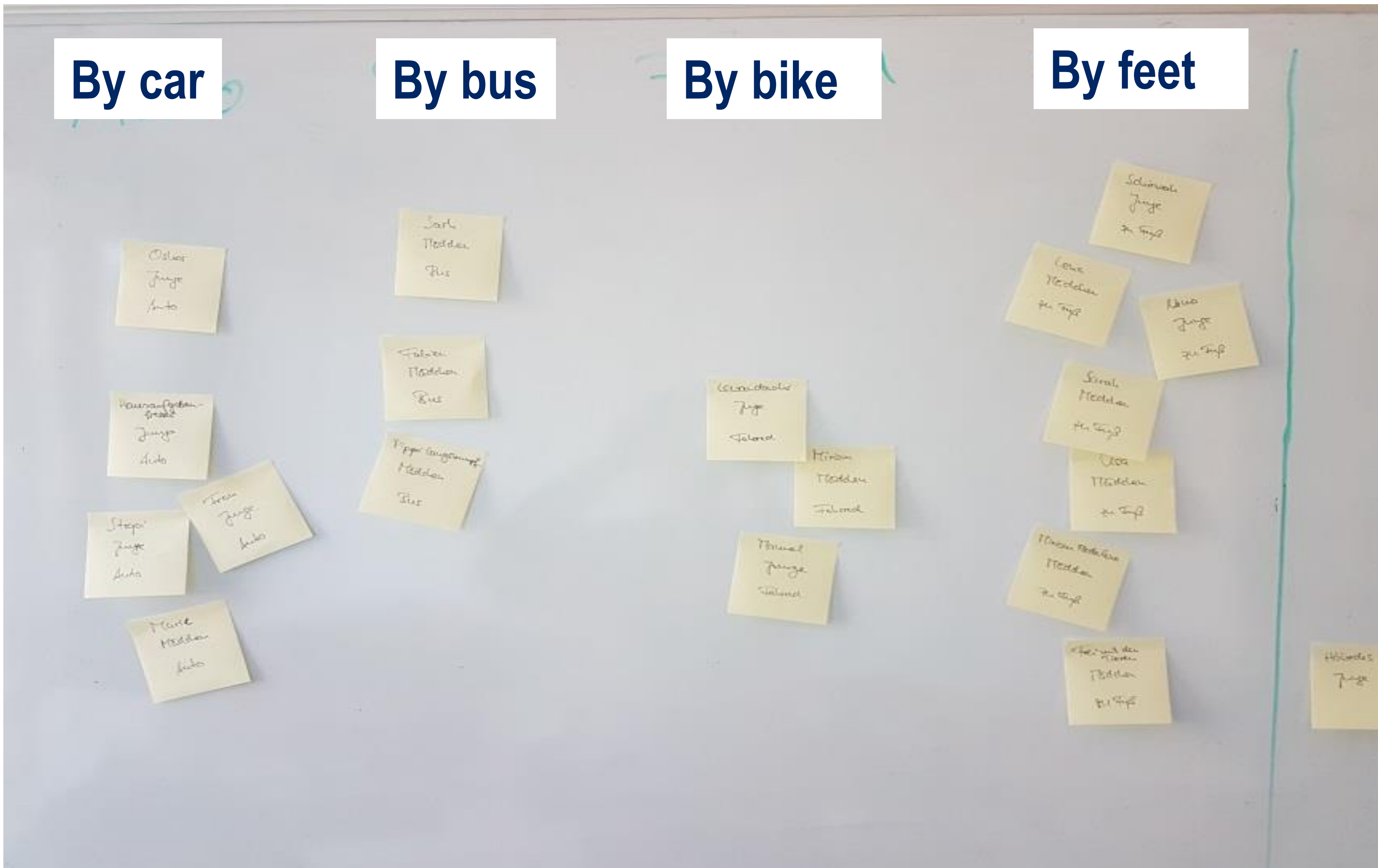


Data cards without any order



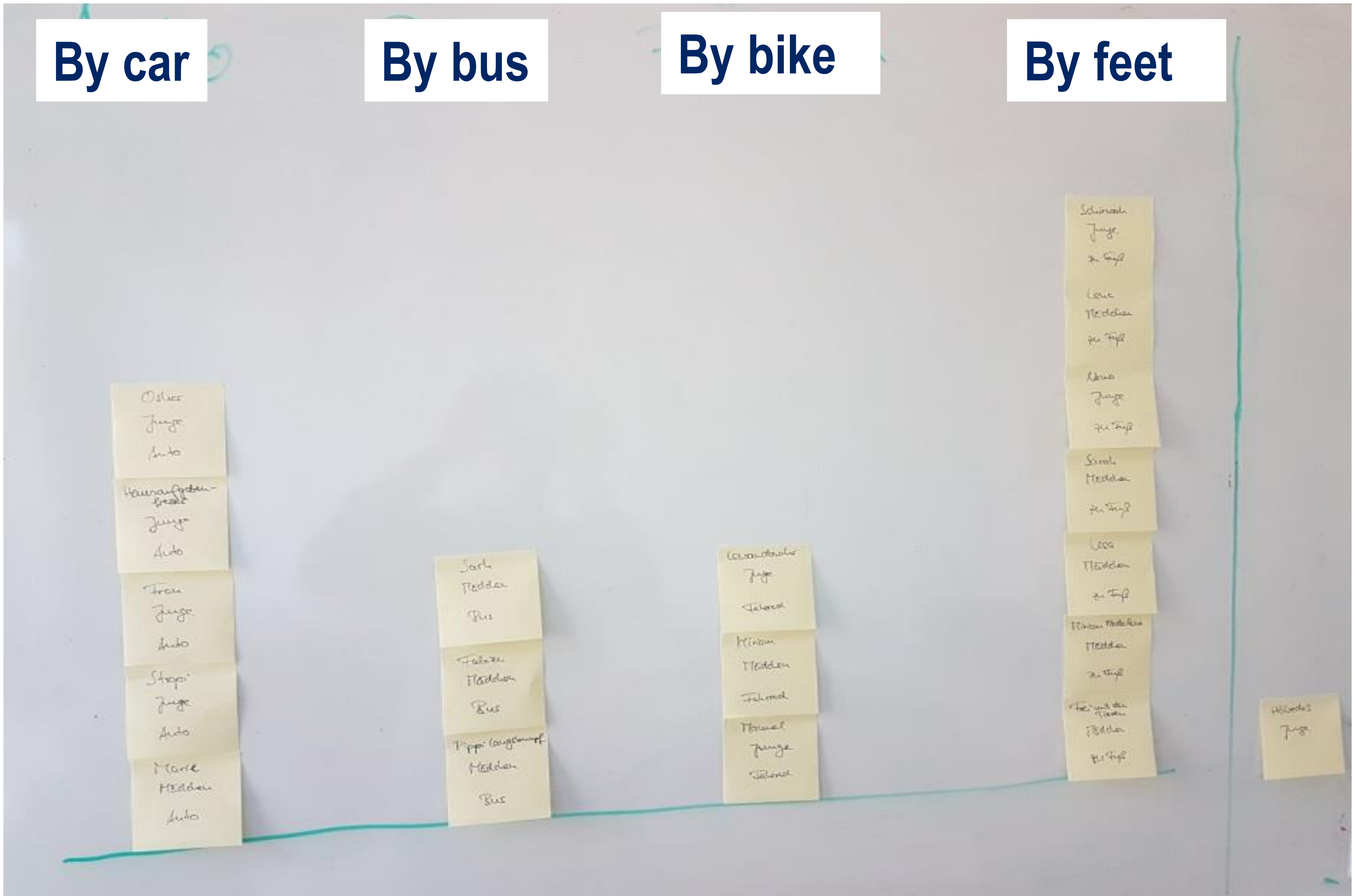
Data: (Primary school Willich, class 4, 2017)

Data cards separated with regard to „how students come to school“



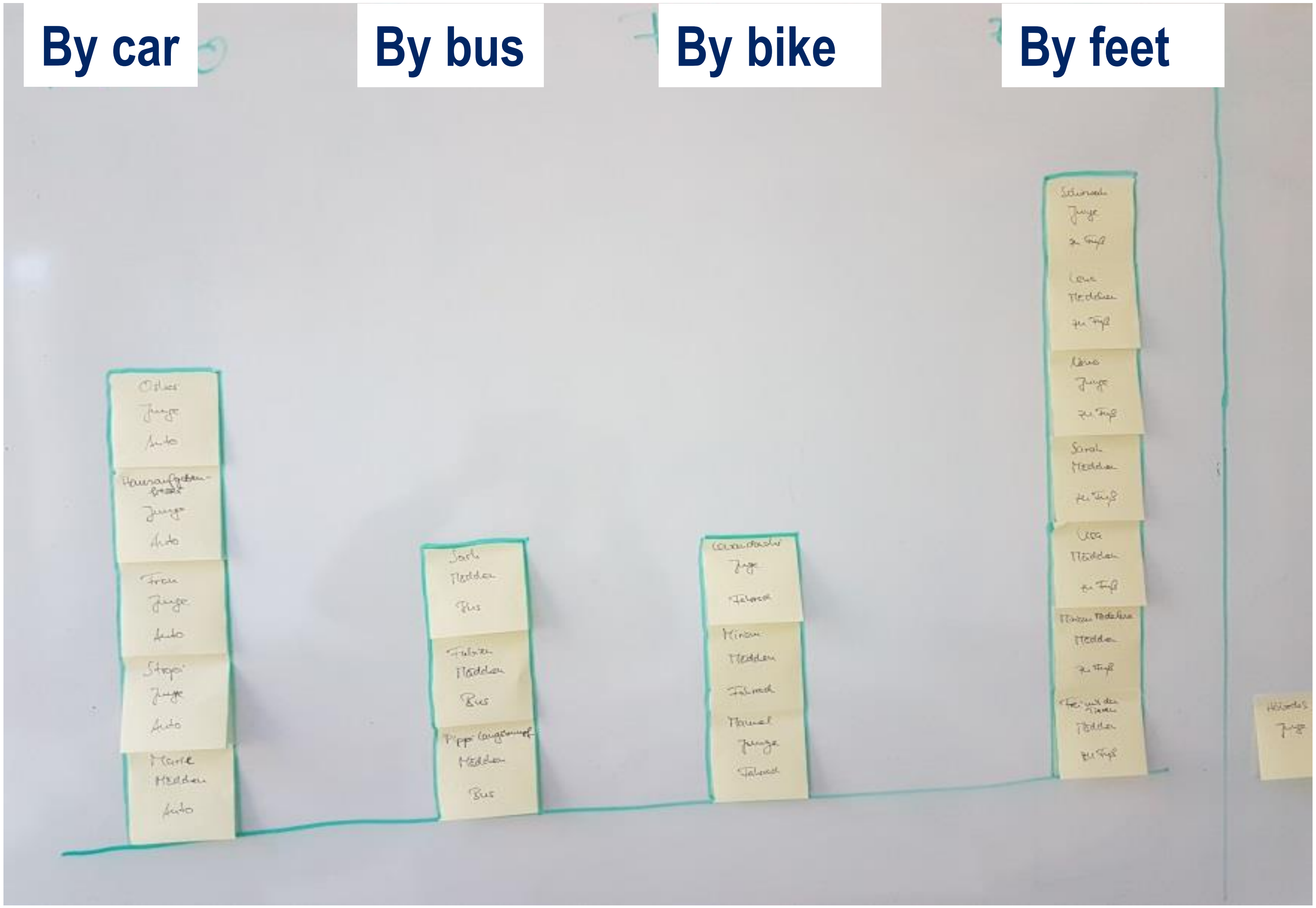
Data: (Primary school Willich, class 4, 2017)

„Data cards“ bar graph



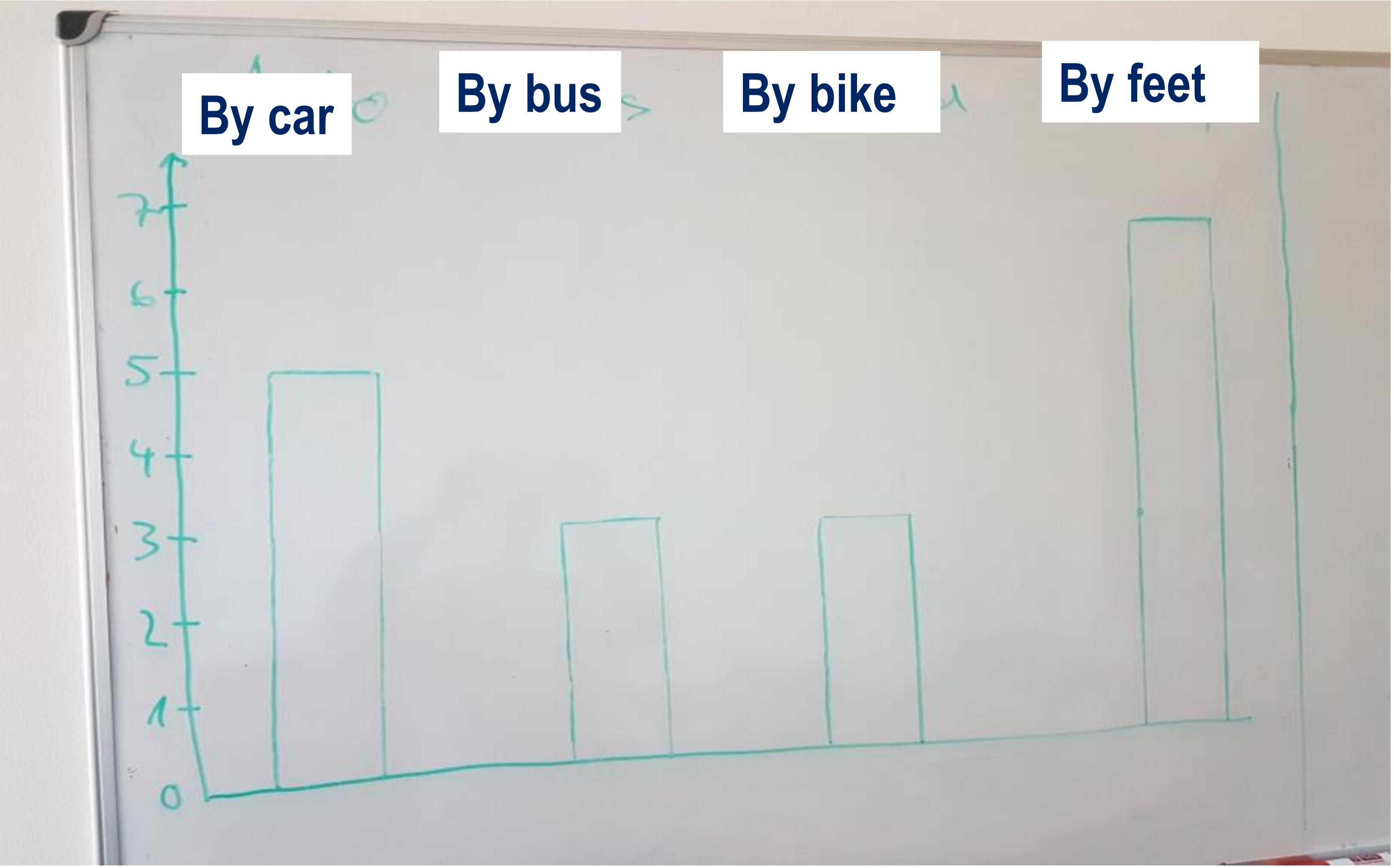
Data: (Primary school Willich, class 4, 2017)

„Data cards“ bar graph (framed bars)



Data: (Primary school Willich, class 4, 2017)

Conventional bar graph



Data: (Primary school Willich, class 4, 2017)

From data cards to TinkerPlots

→ To build on the data analysis with data cards and on well-known data operations (separate, stack)

→ Enabling young students to explore large datasets

TinkerPlots

- Developed by Clifford Konold and Craig Miller on the basis of current findings in statistics education didactics for use in classes 3-8

- TinkerPlots is suitable to provide early access to statistical and probabilistic ways of thinking of students (Konold, 2007) - the special issue about it:
  - Takes up the work with data cards
  - Creating the graphics using three operations: "Separate“, “Stack” and “Order”

Demo

TinkerPlots

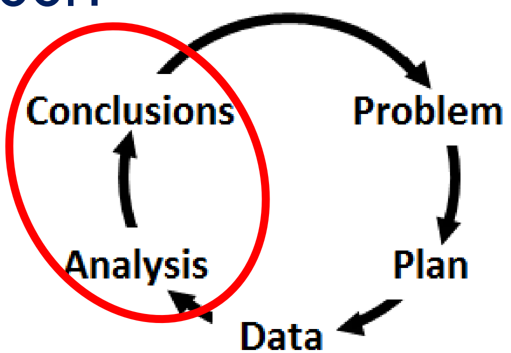
At which stages can TinkerPlots support young learners in their data analysis activities?

- Transferring the data operations (Separate, Stack, etc.) to larger datasets and creating conventional diagrams in larger datasets (Konold 2006)

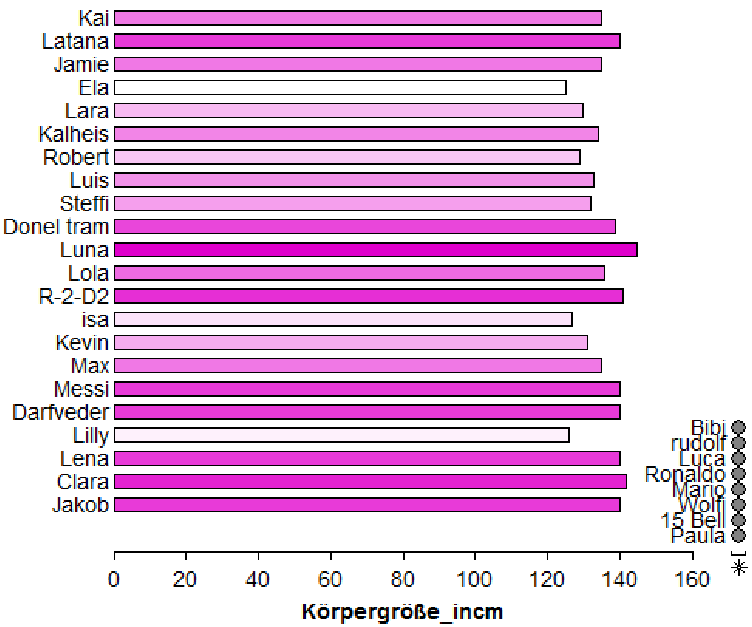
- Creating pie charts
- Paving the way to stacked dot plots (u.a. Cobb 1999, Bakker 2004) → distribution of a numerical variable
- Using preconcepts like modal clumps (Konold et al. 2002, Bakker 2004) and hatplots (Watson et al. 2008), to read and interpret distributions of numerical variables

Stacked dot plots

How can we visualize the distribution of a numerical variable (like height) in primary school?



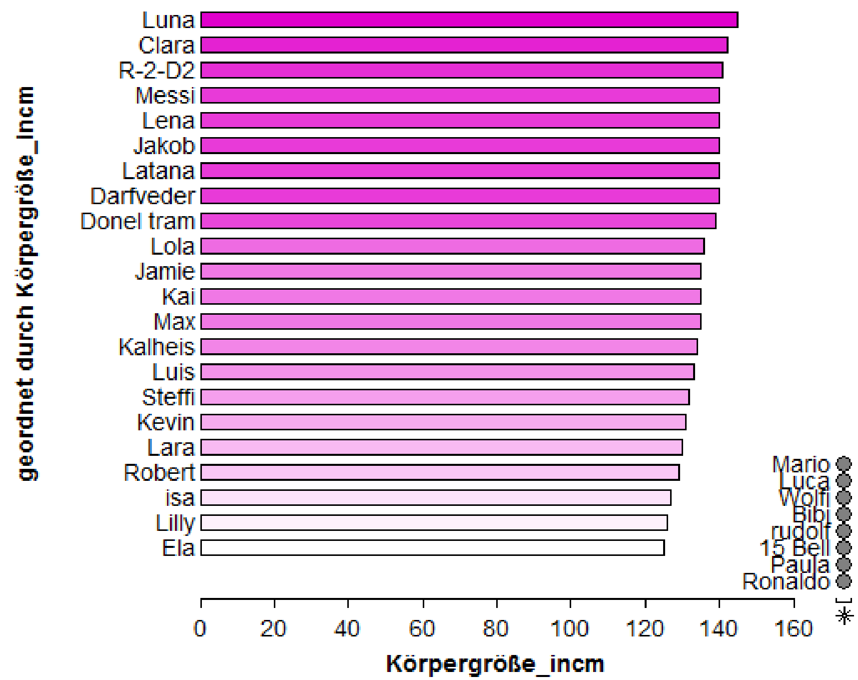
From case value plot to a stacked dot plot in four steps



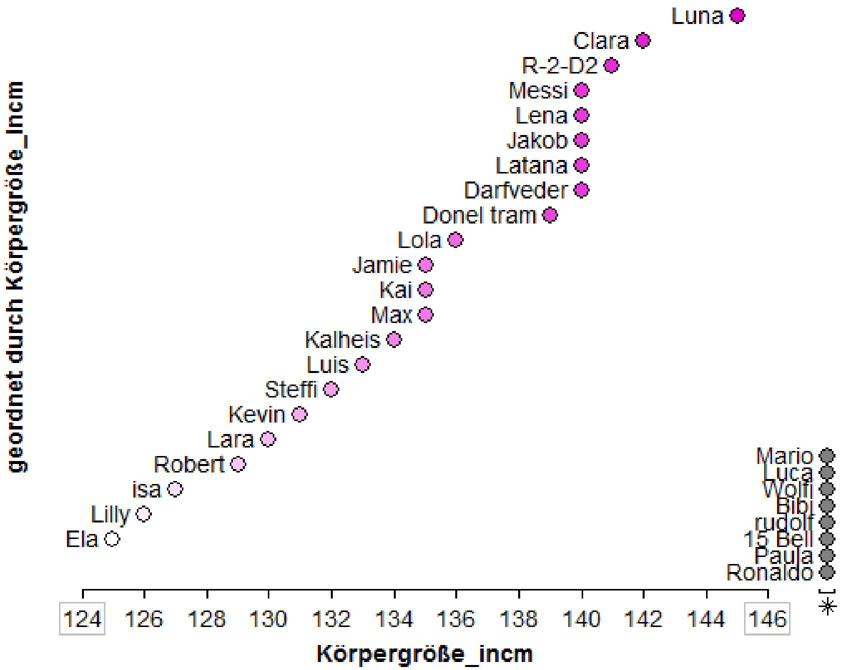
Step 1: Value bars unordered

See Cobb (1999), Bakker (2004)

Step 2: Value bars ordered

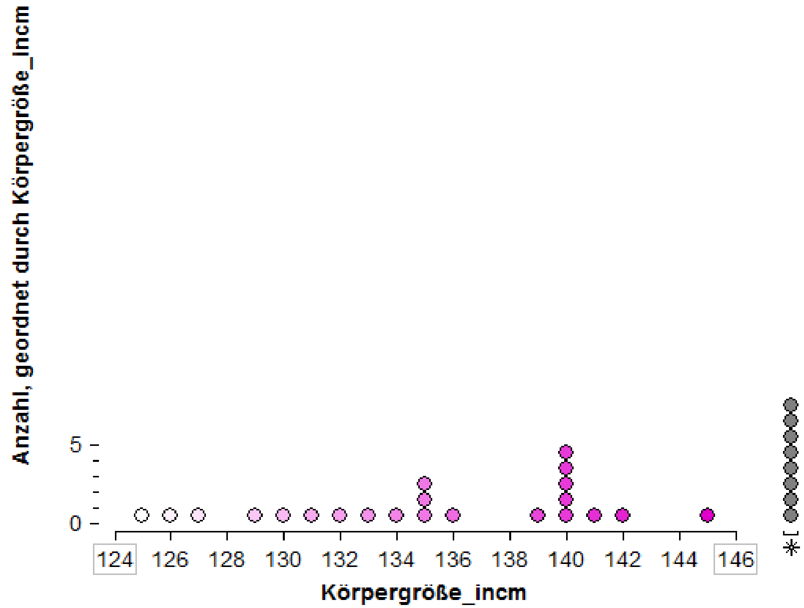


From case value plot to a stacked dot plot in four steps

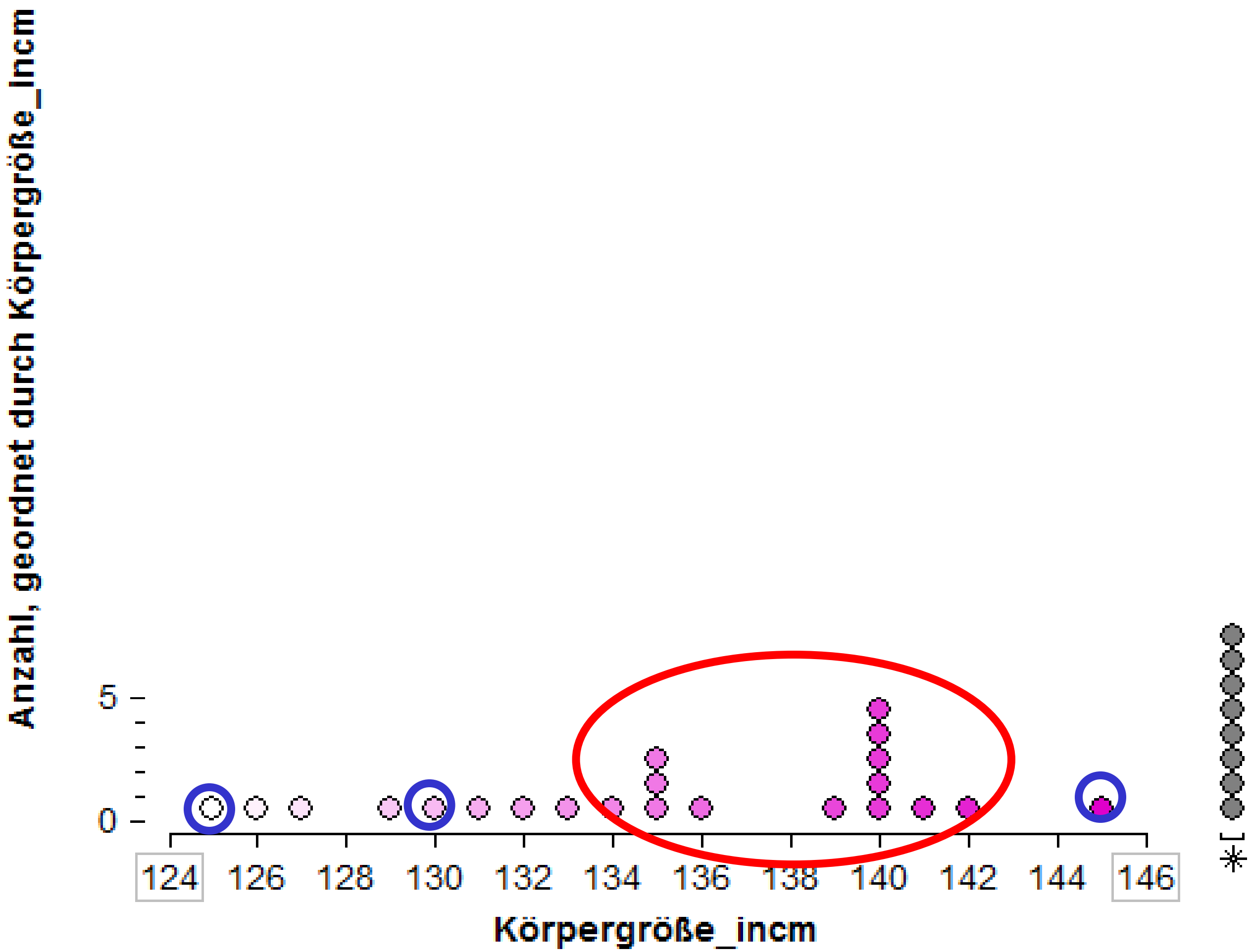


Step 3: Replacing value bars with dots

Step 4: Stack dots



Local vs. global view on distributions (see Bakker & Gravemeijer, 2004)



Local vs. global view on distributions

(see Bakker & Gravemeijer, 2004)

Distribution (global view)

↕

Shape
Density

Measures for
center & spread

Informal
preconcepts
e.g., modal clumps

↕

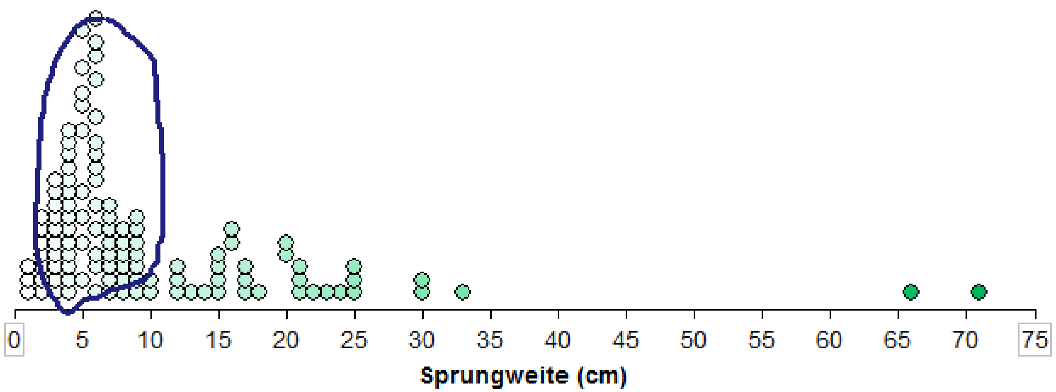
Data (local view)

(see Bakker & Gravemeijer, 2004)

Universität Paderborn
Die Universität der Informationsgesellschaft

Modal clumps

“Main interval” in the data

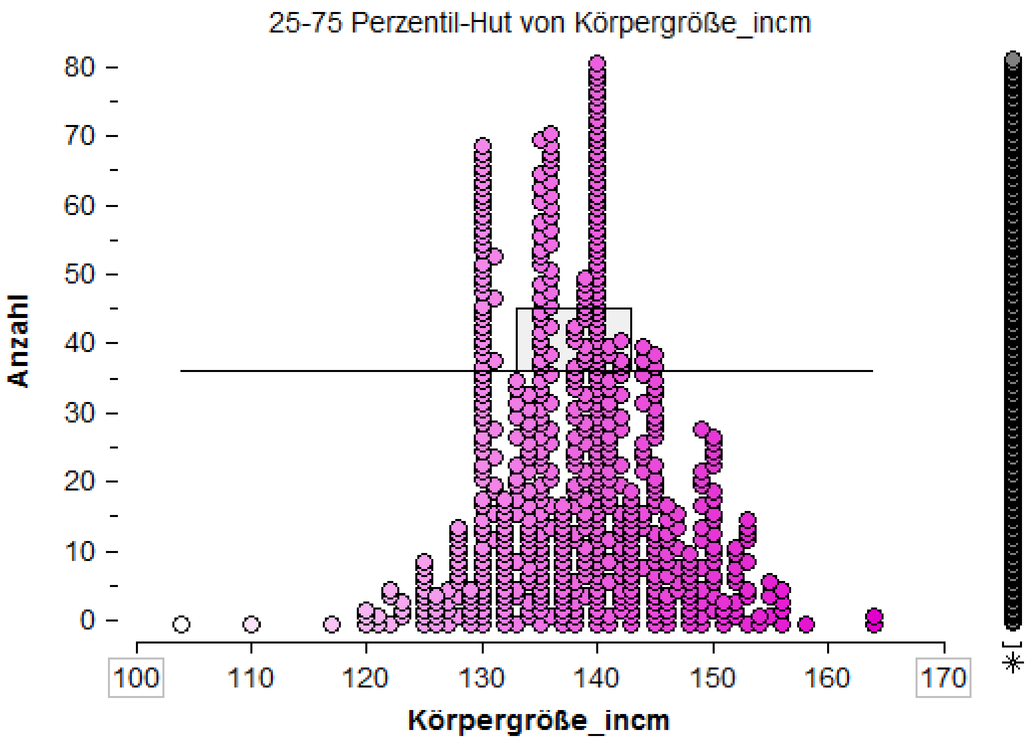


- “range of data in the heart of a distribution of values”
- “appear to allow students to express simultaneously what is average and how variable the data are”
- “may provide useful beginning points for explorations of more formal statistical ideas of center“ (Konold et al. 2002, p. 1).

Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A., Wing, R., & Mayr, S. (2002). *Students' use of modal clumps to summarize data.* Paper presented at the Sixth International Conference on Teaching Statistics, Cape Town, South Africa.

A hat for the data

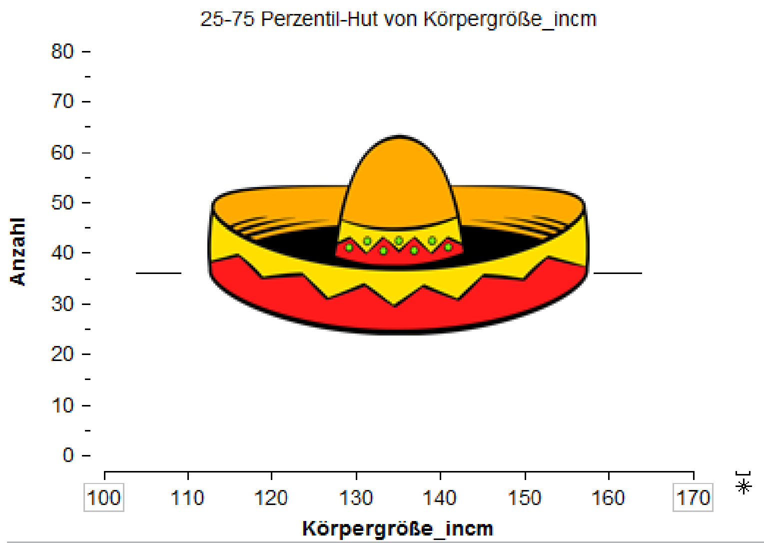
Hatplot visualizes the middle 50% of the data

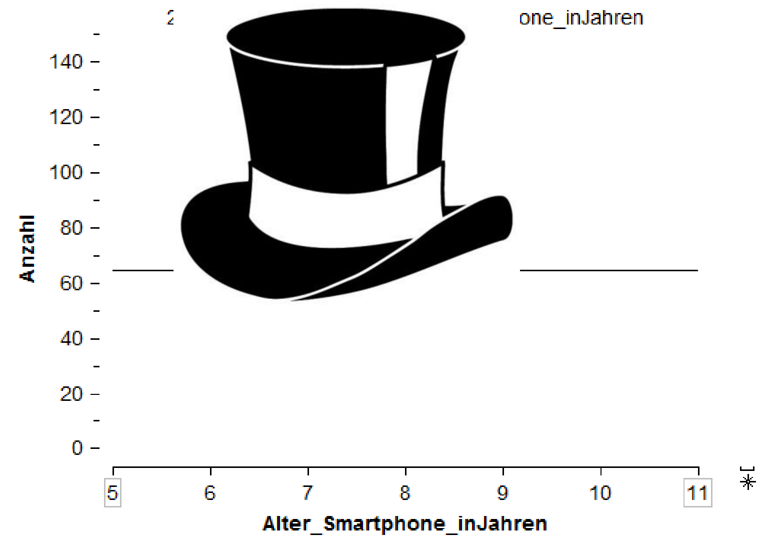


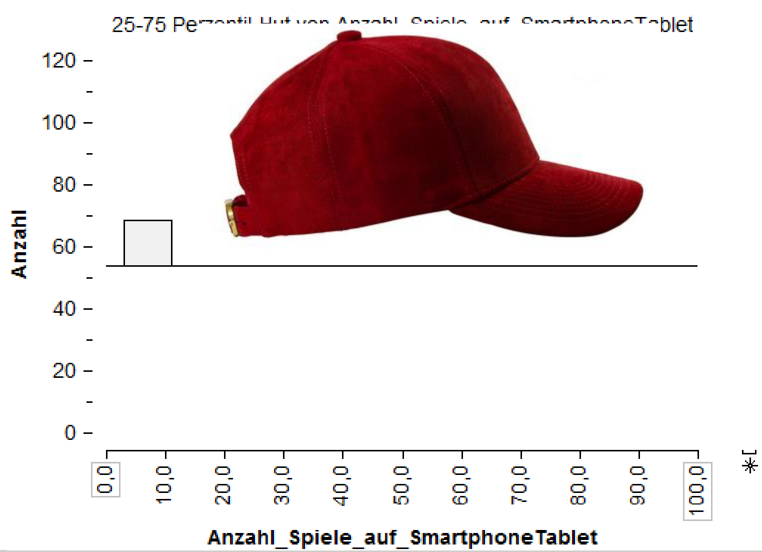
- can offer an easier approach to identifying spread in distributions since they are composed of three major parts (two brims and a crown):
- the brim is a line that extends to the range for each group; the crown is a rectangle that, […] shows the location of the middle 50% of the data – the Interquartile Range (IQR) (Konold 2002, p. 1).

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

Methaphorical hats: Describing and comparing distributions







Design and realization of a teaching unit to enhance statistical reasoning (especially with regard to group comparisons) in primary school

Outline of our research project

Goals

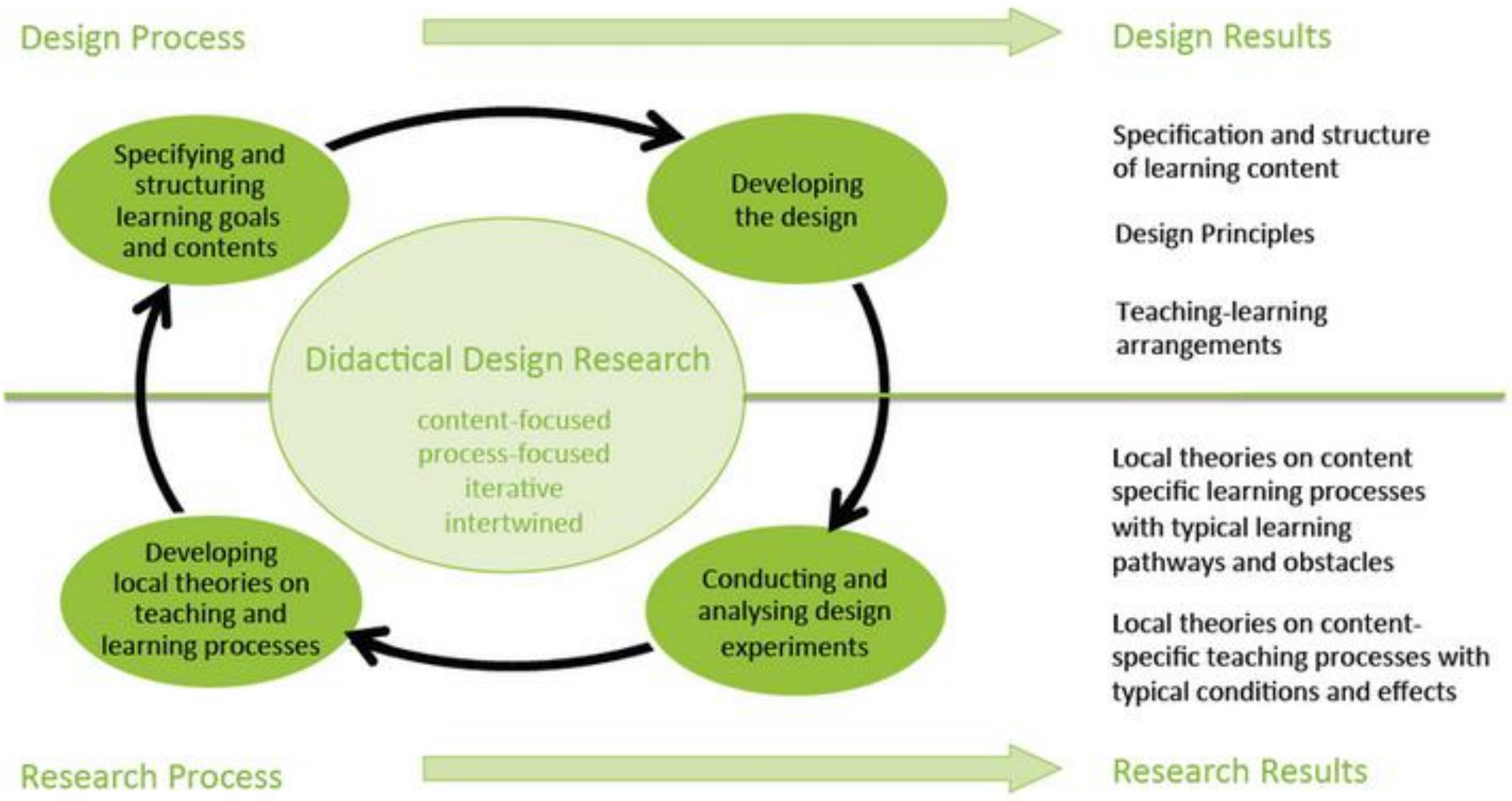
- introduce primary school students to modal clump and hatplots and lead them towards a global view on distribution
- to provide them with a first concept for comparing groups.

→ **Design and the realization of lessons leading to group comparison activities in grade 4 (age: 10-11 years) in primary school.**

→ **Accompanying study which assesses the performance of students before and after attending to a teaching unit about comparing groups.**

Research Methodology

Didactical Design Research (Prediger, & Zwetzschler, 2013).



Design of the teaching unit

The lessons were designed and taught by a preservice teacher for primary school within the scope of her Bachelor Thesis

Major design ideas:

- **Elements of the „Statistical Reasoning Learning Environment“:** Cobb & McClain (2004) and Garfield & Ben-Zvi (2008)
- **PPDAC-cycle** (Wild & Pfannkuch 1999)
- **Working with real data** (Engel 2007, Garfield & Ben-Zvi 2008)
- **Using educational software TinkerPlots** (Konold 2007, Garfield & Ben-Zvi 2008, Ben-Zvi & Pfannkuch 2011)

Realization of the course (Cycle 2)

- Primary school in rural area in Germany, 4th grade students (ages 10-11)
- 10 sessions à 45 minutes
- Participants: 12 grade 4 students (ages 10-11)

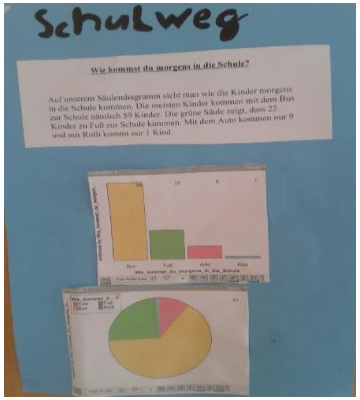
UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Overview of the lessons of the teaching unit

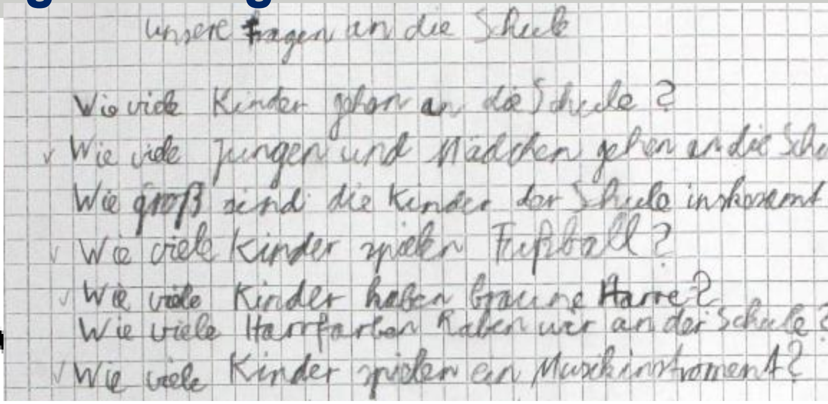
Lesson	Content of the lesson
	Introduction empirical work
1	Getting to know about variables, values, etc.
2	Data analysis cycle (PPDAC), learning to generate statistical questions
3	Collecting data
	Analysing data about the class
4	Drawing and reading bar graphs and pie charts
5	Drawing and reading stacked 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

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

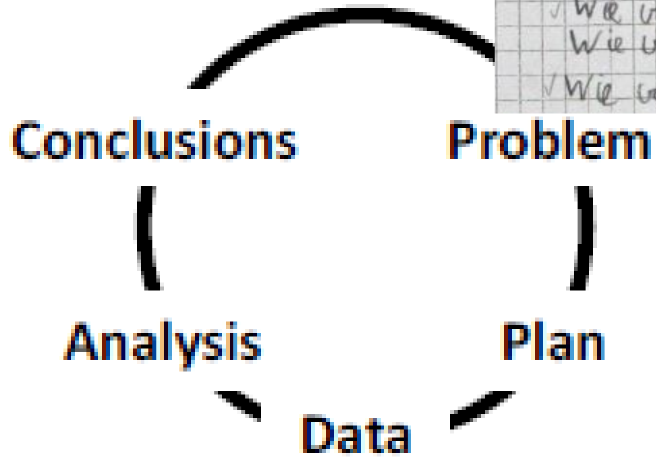
Applying PPDAC cycle in the teaching learning environment

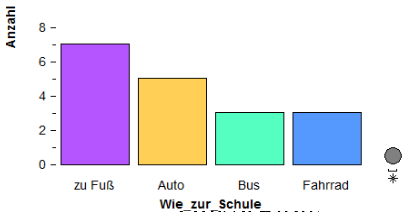


Conclusions, Data posters

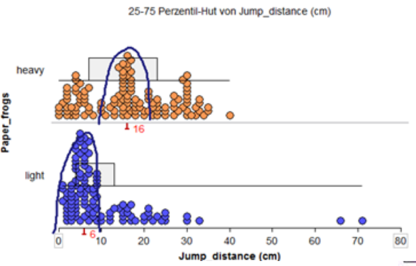


Posing statistical questions





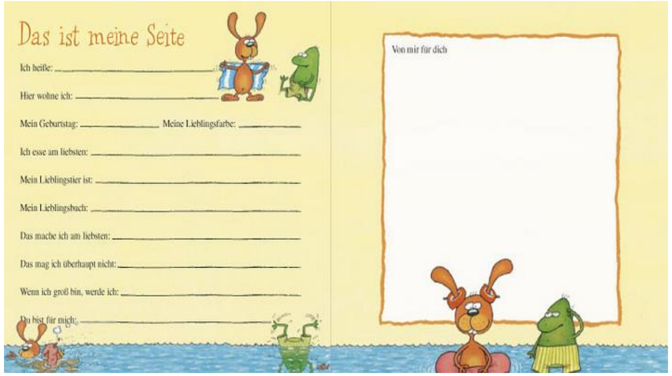
TinkerPlots



kmal	Wert	Ein...	Formel
	Willich		
ßstadt_Stadt_D...	Stadt		
tasiename	Stropi		
chlecht	Junge		
r	9		
se	4		
Körpergröße_incm	138		
Schuhgröße	35		
Anzahl_Kinder_Fa...	mehr als 4		
Haustier	ja		
FallsHaustier_Wel...	Hund		
FallsHaustier_Wie...	mehr als 2		
FallsHaustier_Wie...	6,00		
Minuten_zur_Schule	2		

Data cleaning, Import of the data in TinkerPlots

Creating a questionaire for grade 3 students, collecting data in class



Contents of the lessons on comparing groups

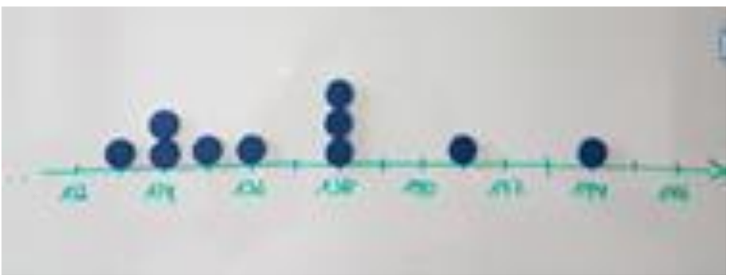
1. Learning stepwise to compare groups

- Reading the data and reading between the data on stacked dot plots, Pupils get to know modal clumps to identify characteristics of the distribution
- Pupils learn to use modal clumps, medians and hatplots to compare groups

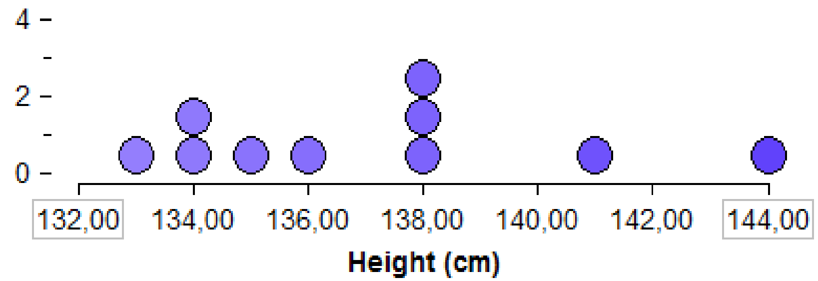
2. Pupils working on group comparison tasks

Learning to compare groups

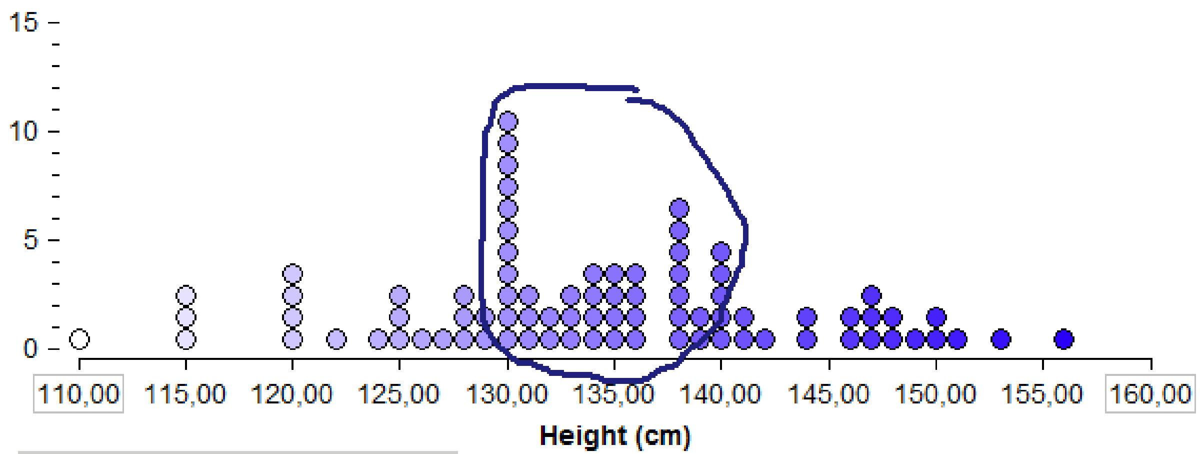
First: Learning to reason about one distribution of a numerical variable



Distribution of height in small dataset on board with magnetic dots



Distribution of height in small dataset in TinkerPlots

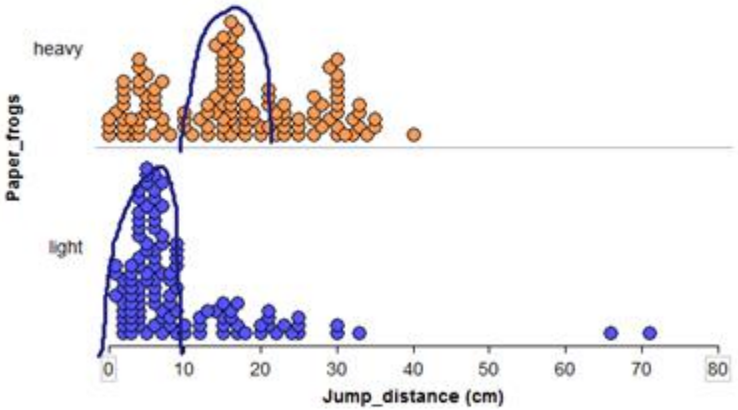


Distribution of height in larger dataset with modal clump in TinkerPlots

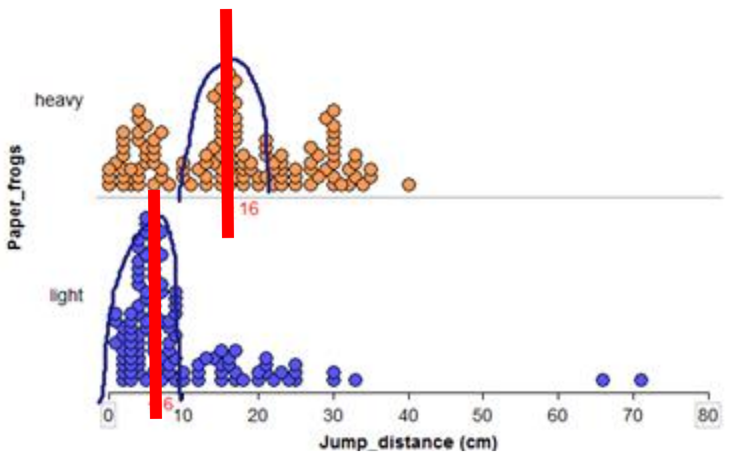
UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Comparing distributions

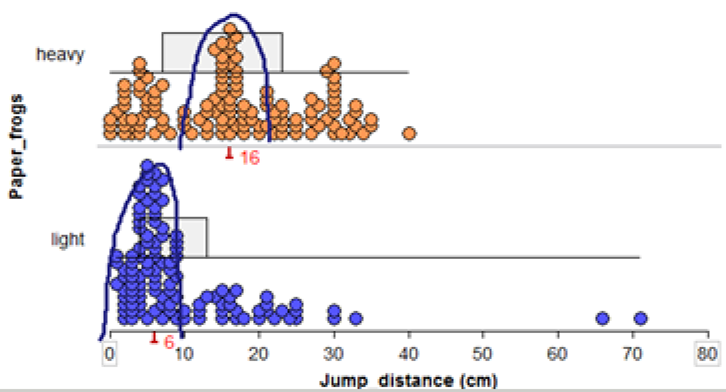
Identifying modal clumps



Identifying median in modal clumps



Using hatplots



Contents of the lessons on comparing groups

1. Learning stepwise to compare groups

- Reading the data and reading between the data on stacked dot plots, Pupils get to know modal clumps to identify characteristics of the distribution
- Pupils learn to use modal clumps, medians and hatplots to compare groups

2. Pupils working on group comparison tasks

- Pupils generate statistical questions leading to group comparisons and conduct group comparisons with given data
- Pupils conduct group comparisons with given data and document their findings on posters.
- Pupils present their findings to their classmates

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

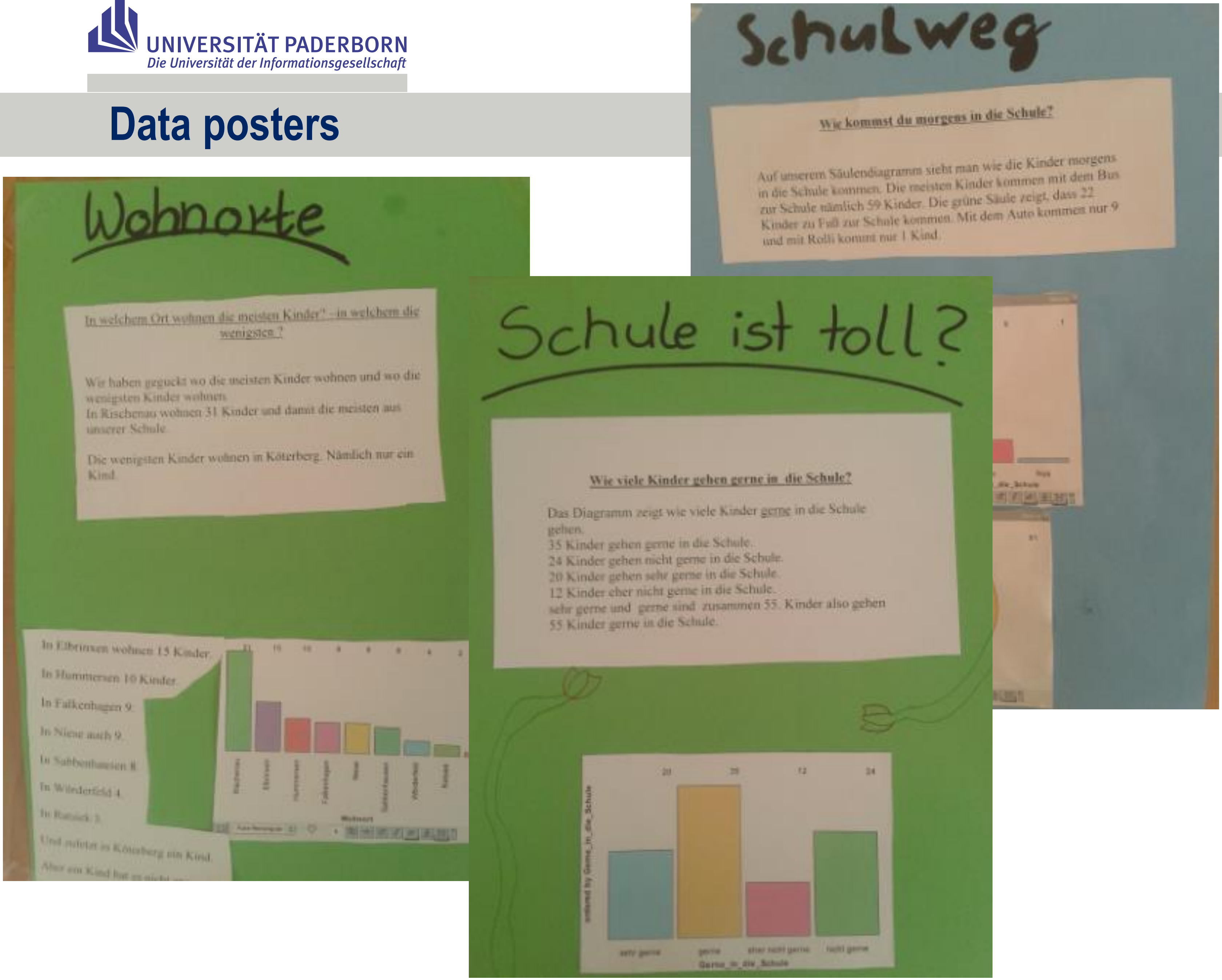
Impressions from the teaching unit

Exploration of data with TinkerPlots in pairs

Presentation of data posters

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Data posters



Presenting all findings in school

Design-based Research: Collected data

- Notebooks of students
- TinkerPlots files of students
- Data posters
- Pre- / post test
- Feedback survey
- Interviews with selected students

Design-based Research: Collected data

- Notebooks of students
- TinkerPlots files of students
- Data posters
- **Pre- / post test**
- Feedback survey
- Interviews with selected students

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Research Questions

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

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

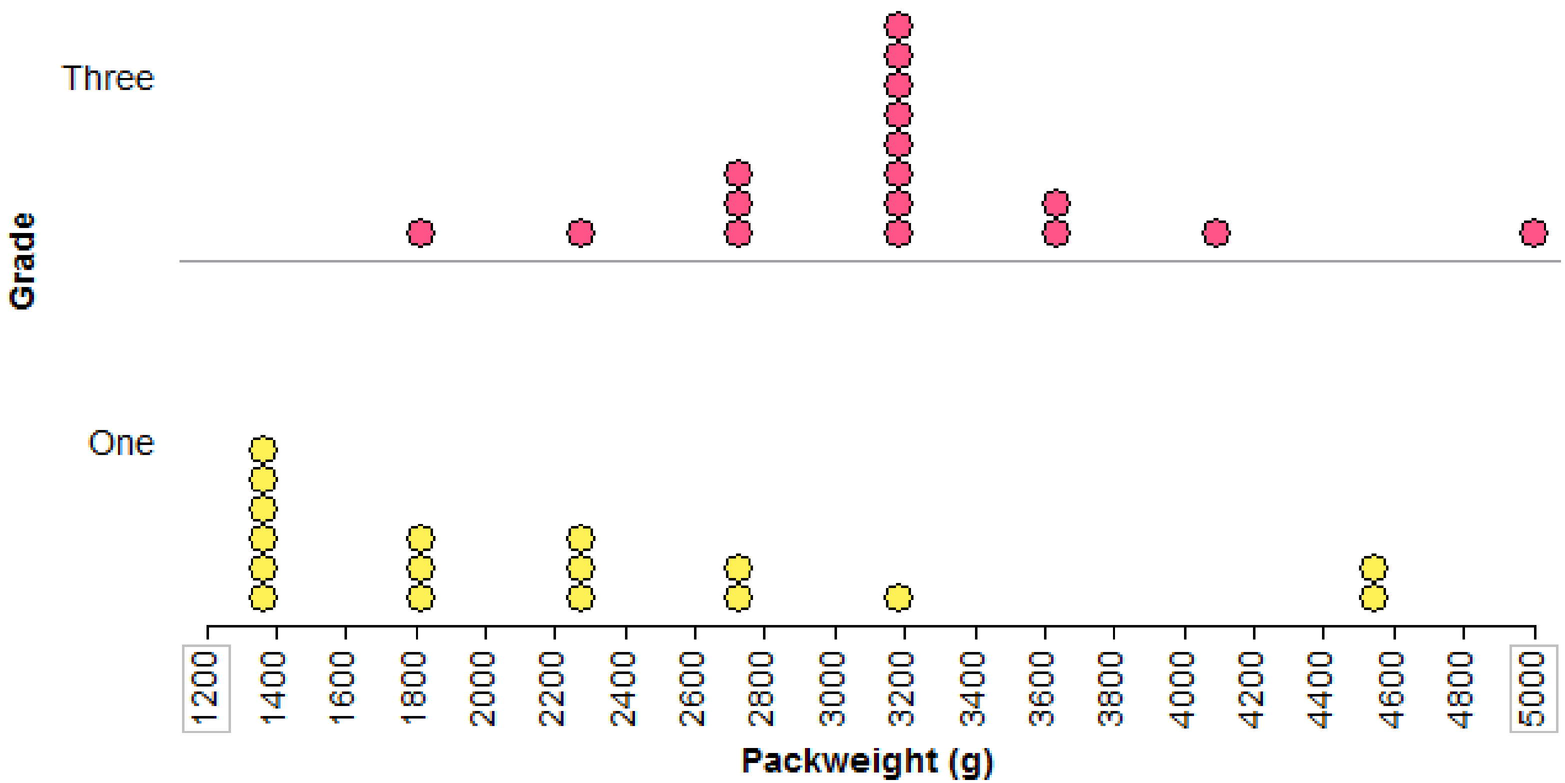
Design of study

Examine group comparison strategies of primary school students before and after they were taught about group comparisons in our teaching unit

For the evaluation we gave the students a group comparison task before they have attended our teaching unit and we gave them the same task after they have attended our teaching unit two weeks later.

The task

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



Participants, Data collection and Data analysis

Participants

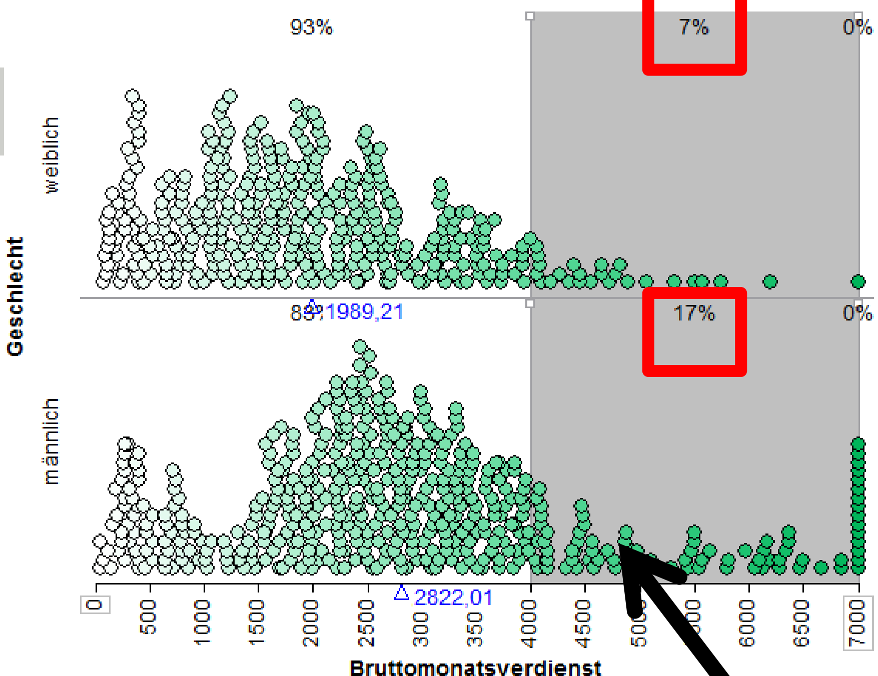
- 12 grade four students (age 10-11)
- All of them have participated in the teaching unit

Data collection

- written notes on the task from all students (n=12 before; n=12 after)

Method for data analysis

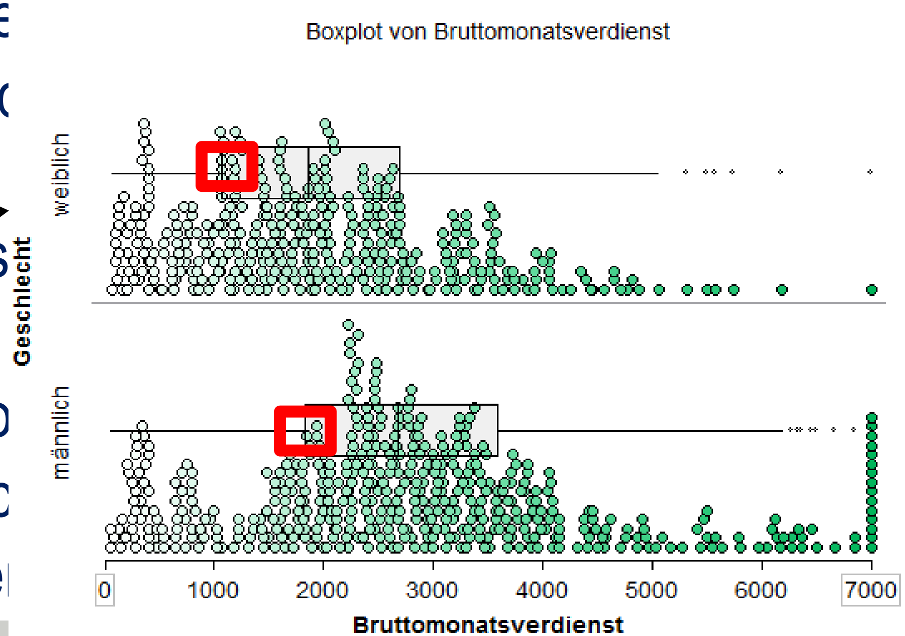
- Structured qualitative content analysis (Mayring 2015) and mixed approach of coding (Kuckartz 2012) to identify the group comparison elements used



ems: coding procedure

)12) → we generated the categories
ed) in a **first step** from a deductive
second step – refined them

rst step) → Biehler (2007),
Pfannkuch (2007), Frischeme
Skewness, Shift, p-based and



- No q-based comparisons
  primary school pupils).
- With regard to compariso
  score" of two groups as c
  sustainable element whe

Deriving group comparison items: coding procedure

- **Use of preconcepts like modal clumps (deductive)**
- **Use of precursor visualisations like hatplots (deductive)**
- **Use of shift of points (inductive)**
  - 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”.

Results

	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
	8/12	4/12	12/12	11/12

Student answers: The case of Maria,

Pretest (shift of points):

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

Posttest (median and hats):

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.

Design-based Research: Collected data

- Notebooks of students
- TinkerPlots files of students
- Data posters
- Pre- / post test
- **Feedback survey**
- Interviews with selected students

Research Questions

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

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

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Survey: 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

N=12 students filled out the survey after the teaching unit

Summary of results

- The study offers interesting insights in fourth graders´ group comparison strategies for the further development of group comparison activities. → it is possible to establish first notations of comparing groups in primary school classroom (after introducing basics of data analysis, stacked dot plots, etc.).
- Even in the pretest, most of the students have shown that they were able to decide whether third graders have heavier backpacks compared to first graders, but the group comparison items of many of the students have been inadequate or have been focusing on local features.
- After attending to the teaching unit the quality of group comparison elements used has improved massively.
- All students show a positive attitude towards statistics and the teaching unit on comparing groups with TinkerPlots

Some Implications…

- TinkerPlots is an adequate digital tool to introduce young stundents in the analysis of larger datasets and into more sophisticated statistical activities (like e.g. group comparisons)
- Use real and meaningful datasets can and should already be implemented in primary school classroom
- Connect data analysis with TinkerPlots with enactive data analysis with data cards, so that both processes “learning data analysis” and “learning the software for analysing data” are connected
- To develop a global view on distributions and to facilitate the group comparison process, the stepwise use of modal clumps, medians and hats is very promising

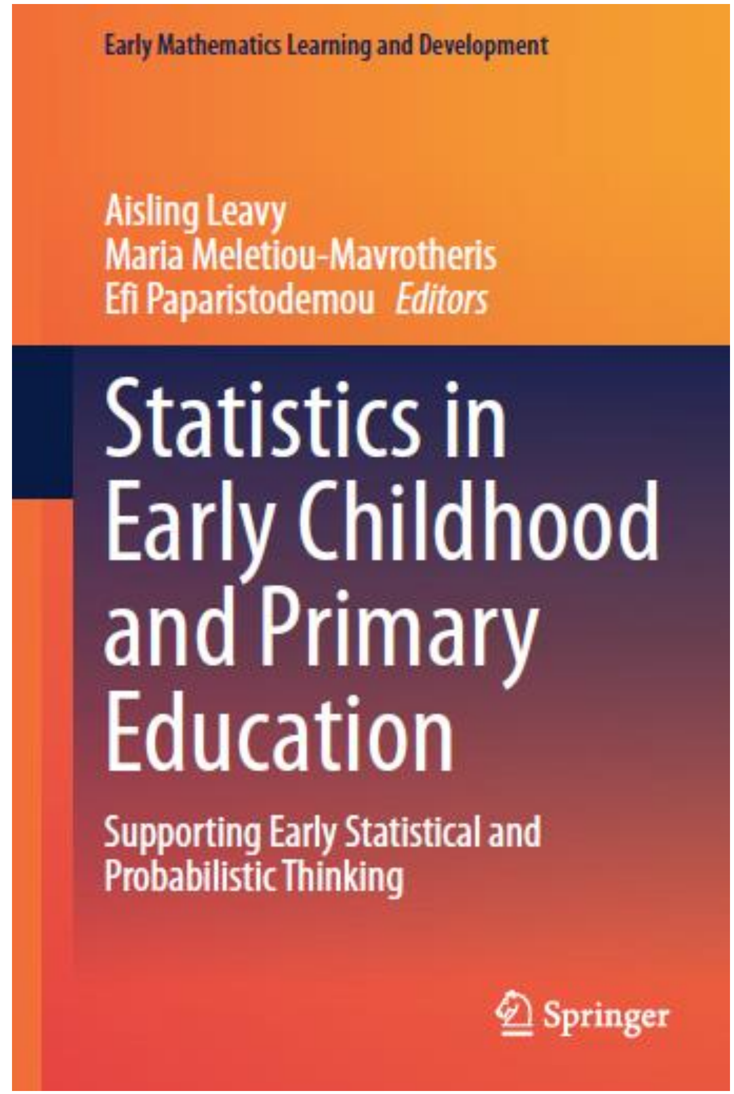
Part 2

Developing reasoning about uncertainty of primary school students using TinkerPlots

The need for Early Statistical Thinking

„Today‘s students need to learn to work and think with data and chance from an early age, so they begin to prepare for the data-driven society in which they live.“ (Ben-Zvi, 2018, vii)

Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2018). *Statistics in Early Childhood and Primary Education: Supporting Early Statistical and Probabilistic Thinking*. Singapore: Springer.



A conceptualization of randomness

Batanero (2015) distinguishes between three conceptualizations of randomness:

(i) Randomness as equiprobability,

(ii) randomness as stability of frequencies,

(iii) subjective view of randomness.

Conceptualizations of randomness

- In “randomness as equiprobability” probability is defined in the sense of Laplace as “the number of favorable cases to a particular event divided by the number of all cases possible in that experiment, provided all the possible cases are equiprobable” (Batanero, 2015, p. 36).

- “Randomness as stability of frequencies” is related to the empirical law of large numbers, probability is defined according to Batanero (2015, p. 37) as “the hypothetical number towards which the relative frequency tends”.

- The subjective view “considered probability as a personal degree of belief that depends on a person´s knowledge or experience” (Batanero, 2015, p. 37).

Conceptualizations of randomness

- In “randomness as equiprobability” probability is defined in the sense of Laplace as “the number of favorable cases to a particular event divided by the number of all cases possible in that experiment, provided all the possible cases are equiprobable” (Batanero, 2015, p. 36).

- “Randomness as stability of frequencies” is related to the empirical law of large numbers, probability is defined according to Batanero (2015, p. 37) as “the hypothetical number towards which the relative frequency tends”.

- The subjective view “considered probability as a personal degree of belief that depends on a person´s knowledge or experience” (Batanero, 2015, p. 37).

Conceptualizations of randomness

- In “randomness as equiprobability” probability is defined in the sense of Laplace as “the number of favorable cases to a particular event divided by the number of all cases possible in that experiment, provided all the possible cases are equiprobable” (Batanero, 2015, p. 36).

- “Randomness as stability of frequencies” is related to the empirical law of large numbers, probability is defined according to Batanero (2015, p. 37) as “the hypothetical number towards which the relative frequency tends”.

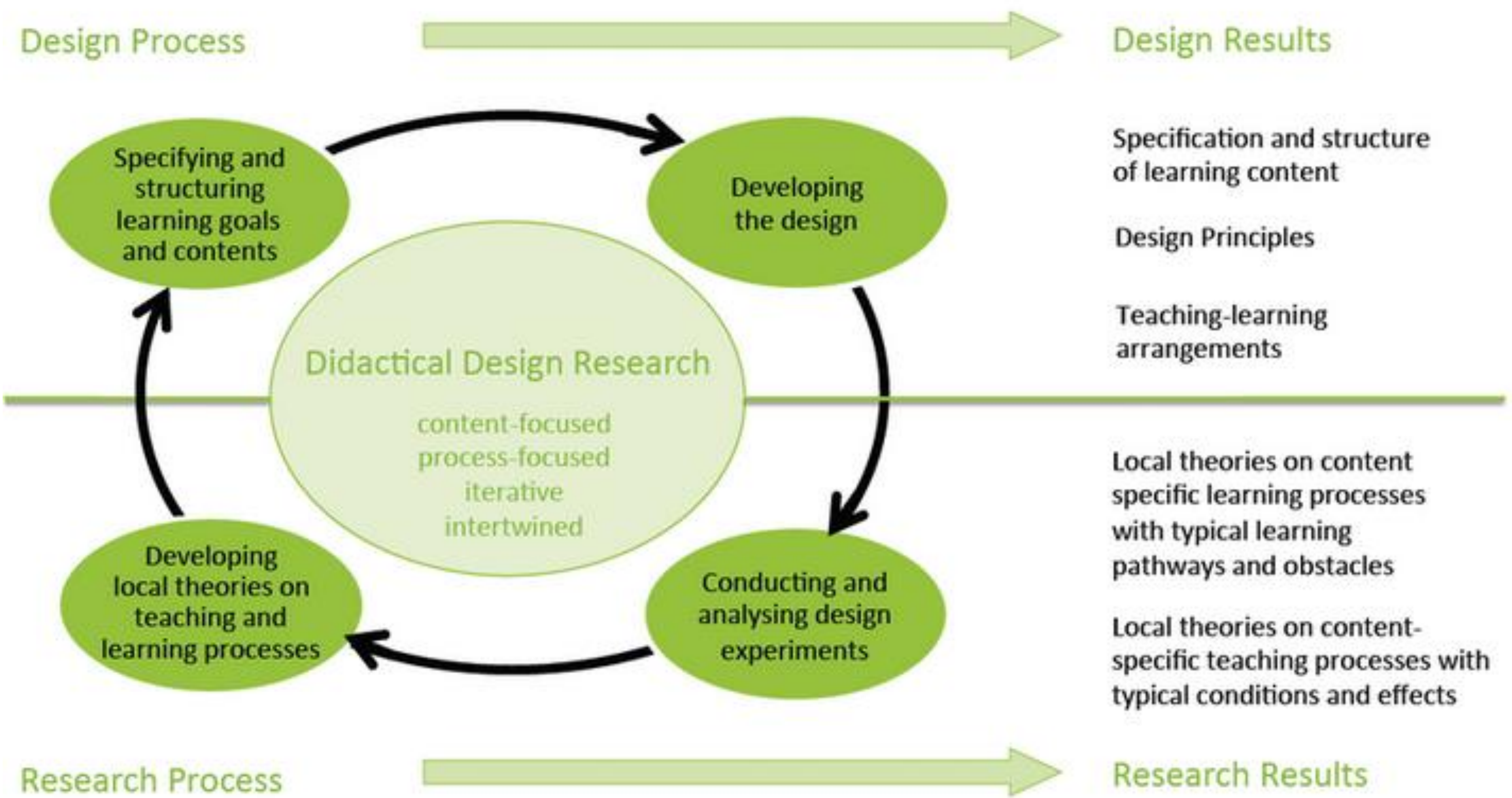
- The subjective view “considered probability as a personal degree of belief that depends on a person´s knowledge or experience” (Batanero, 2015, p. 37).

Goals of this project

Design of a teaching unit which covers each of the three conceptualizations of randomness to develop primary school students´ reasoning about uncertainty

Research Methodology

Didactical Design Research (Prediger, & Zwetzschler, 2013).



Research study

- First cycle of Design research
- Primary school in rural area in Germany
- 7 lessons (45 minutes each)
  – Data: field notes of the teacher, documents of students, videos
- Participants: 20 students, grade 4 (10-11 years old)
  – no specific pre-knowledge in statistics apart from collecting data in tallies and creating and reading bar charts and reading pie charts.

Design ideas for the teaching learning environment

- **Taking into account all three conceptualizations of randomness** of Batanero (2015) to develop the reasoning about uncertainty among grade 4 students.
- **Elements of the „Statistical Reasoning Learning Environment“** (Garfield & Ben-Zvi , 2008), e.g.
  - **Peer learning settings, promoting classroom discourse**
  - **Using educational software TinkerPlots** (Konold 2007, Garfield & Ben-Zvi 2008, Ben-Zvi & Pfannkuch 2011) → TP Sampler

Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in "certain", "possible", "unlikely" and "impossible".
3	Pupils conduct hands-on experiment "throw of a dice" and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment "throw of two dice", collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation "throw of two dice" and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Contents of the second lesson

- Students evaluate events of several chance experiments using the expressions “certain”, “probable”, “unlikely” and “impossible”.
- Teacher offers a strip with a scale from “impossible” to “certain” which the students use to show and to visualize their expectation with regard to specific events.
- Then: Teacher expands the scale and adds frequency descriptions like “never”, “seldom”, “often” and “always” for each terminology

Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Contents of the third lesson

Pupils conduct hands-on experiment "throw of a dice" and collect data to get the insight that the probability for each side of the dice is equal

Students work in pairs (10 pairs) to throw a dice fifty times and document the frequencies of the occurrences of each side (hands-on activities)

→ Teacher uses notations like "In 79 of 500 cases the side 1 has occurred" or "in 78 of 500 cases the side 2 has occurred" to make the frequencies comparable

→ Insight for the students: the probability for each side of the dice is equal.

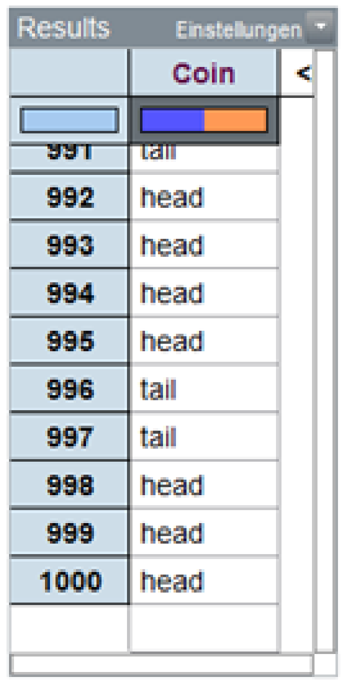
UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

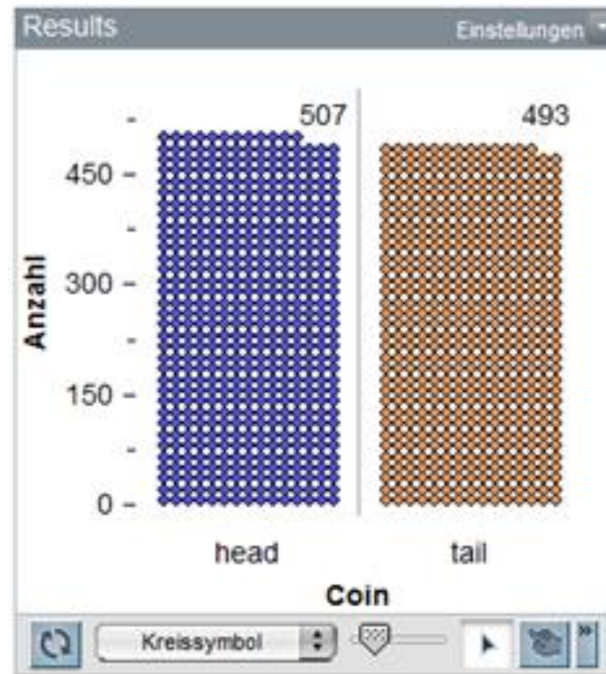
Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Contents of the fourth lesson

Students are introduced in the simulation of chance experiments with the TinkerPlots sampler.
→ Teacher demonstrates how to realize the “toss of a coin” in the TinkerPlots sampler





Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Contents of the fifth lesson

- Teacher chooses a more complex chance experiment (“throw of two dice”) → multi-step chance experiments.
- Task for the students: “Find out which sum is more likely to appear when throwing two dice.”
- Students are paired in groups and each group was given two dice and the task to throw the two dice fifty times and to document their outcomes → teacher collects the results of all ten groups

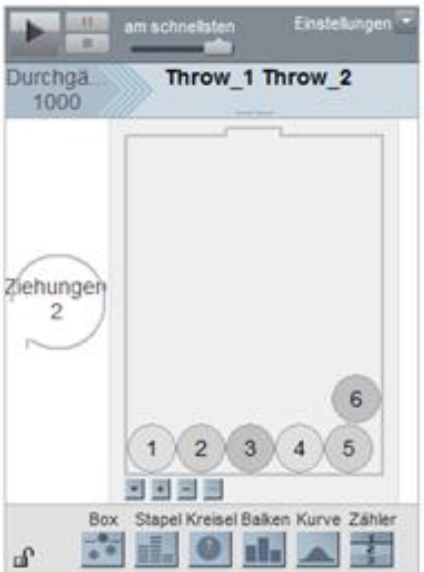
The first insight for the students in lesson 5 is that the different sums of the throw of two dice are not equiprobable

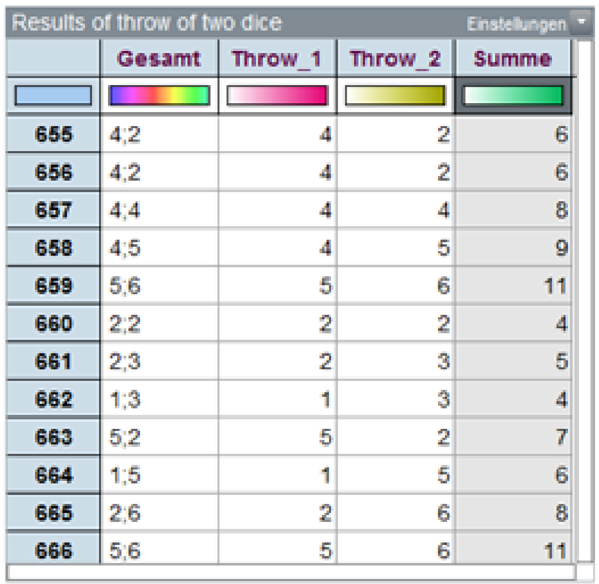
Seven lessons of the teaching learning environment

No	Content of lesson
1	Pupils learn about how to read and interpret statistical bar graphs and pie charts
2	Pupils estimate and compare probabilities of events by classifying them in “certain”, “possible”, “unlikely” and “impossible”.
3	Pupils conduct hands-on experiment “throw of a dice” and collect data to get the insight that the probability for each side of the dice is equal.
4	Pupils are introduced in the sampler of TinkerPlots and get first experiences of the empirical law of large numbers by simulating the throw of a coin with the TinkerPlots sampler
5	Pupils conduct hands-on experiment “throw of two dice”, collect the sum of the two dice for each throw and collect data to use the frequencies of occurrence of the different sums to make statements with regard to the probability of the events of different sums
6-7	Pupils conduct TinkerPlots simulation “throw of two dice” and collect data to compare the probability of events. Pupils try to find explanations why some sums appear more frequent than other sums when throwing two dice.

Contents of the sixth and seventh lesson

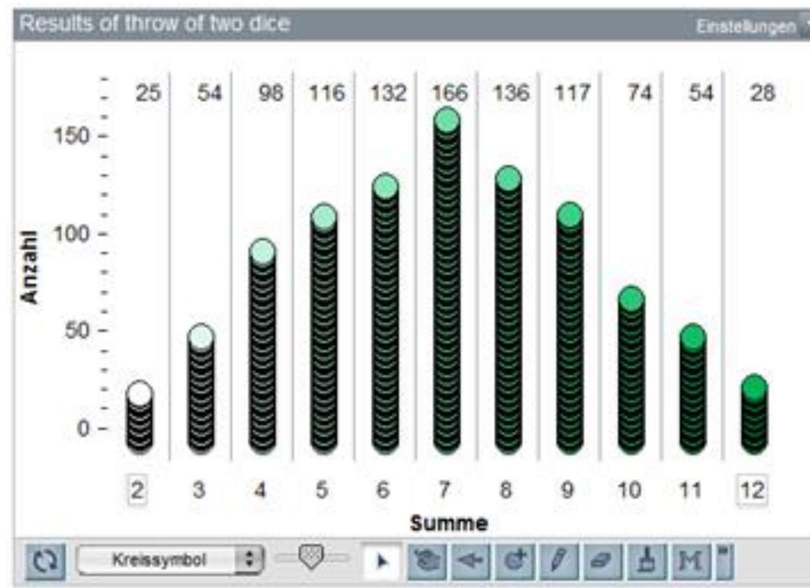
- The teacher reflects the results from the hands-on activity in the fifth lesson on the chance experiment “sum of two dice” and demonstrates how to set up this chance experiment in the TinkerPlots sampler





Results of throw of two dice

	Gesamt	Throw_1	Throw_2	Summe
655	4;2	4	2	6
656	4;2	4	2	6
657	4;4	4	4	8
658	4;5	4	5	9
659	5;6	5	6	11
660	2;2	2	2	4
661	2;3	2	3	5
662	1;3	1	3	4
663	5;2	5	2	7
664	1;5	1	5	6
665	2;6	2	6	8
666	5;6	5	6	11



Contents of the sixth and seventh lesson

- The teacher repeats the simulation in TinkerPlots several times and the students observe that the mode of the distribution is the sum “7”.
- The students have the idea to consider the possible outcomes of the throw of two dice for each of the sums (2-12) → identify the outcomes for each sum and collect their results
- The students explore that the sum “7” has the most outcomes (1+6, 6+1, 2+5, 5+2, 3+4, 4+3) and therefore occurs more often in the experiment than for example the sum “3” (with the two outcomes 1+2 and 2+1) or the sum “2” (with the outcome 1+1).
- The students also find out the outcomes for each sum and that there are 36 outcomes in total (each of them are equiprobable).

Contents of the sixth and seventh lesson

- Comparisons of probabilities of certain events like “the sum 6 is more likely to show up than the sum 3”
  → because there are five favorable out of 36 outcomes for sum “6” (1+5, 5+1, 2+4, 4+2, 3+3) but only two favorable out of 36 outcomes for sum “3” (1+2, 2+1).

First observations from the teaching-learning environment

- Students were very engaged - especially with the simulation of chance experiments (hands-on, TinkerPlots)
- Elements of the three conceptualizations of randomness (see Batanero, 2015) can be already implemented in primary school classroom.
- TinkerPlots sampler seems to be a powerful educational tool to facilitate the modelling and the data production process when simulating chance experiments in primary school
  → Some students needed support (especially when setting up the model)
  → Challenge: interpretation of the produced data

Explorative Study: „Primary school students‘ statistical reasoning when conducting chance experiments with TinkerPlots“

Research question

In which way does the teaching unit develop the competence of pupils with regard to compare probabilities of events of multi-step chance experiments?

Data, participants and data analysis

We collected data on different levels:

(a) written pre/post-tests,

(b) working notes on tasks and activities after each lesson and

(c) interviews after the teaching unit.

Data, participants and data analysis

We collected data on different levels:

(a) **written pre/post-tests,**
(b) working notes on tasks and activities after each lesson and
(c) interviews after the teaching unit.

Written pre/post tests were handed out at the beginning and at the end of the teaching unit

→ Six tasks.
→ Posttest is identical to the pretest.
→ 19 students have participated in the pre- and posttest.
→ We will only concentrate on the tasks which focus on the comparison of probabilities of events of multi-step chance experiments (tasks 5 and 6).

Tasks

Task (5) You and your friend Tim play a game where you toss a coin twice. You win if you get tail both times. Tim wins if he gets tail and head when tossing the coin twice. Do you both have the same chances of winning? Explain!

Task (6) Anna and Tom are playing a game. They throw a dice twice. If the sum of the two dice is "7", Anna wins. If the sum is 3 or 12, Tom wins. Do Anna and Tom have the same chance of winning? Explain!

Data analysis procedure & Results

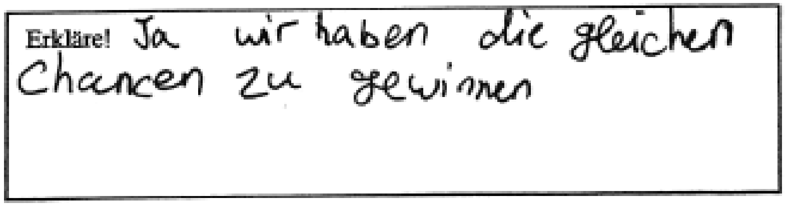
1. Coding the correctness of the statements of each participant in pre-test and post-test
2. Frequency analysis and calculation of the percentages of correct statements of each task in comparison to pre- and posttest.

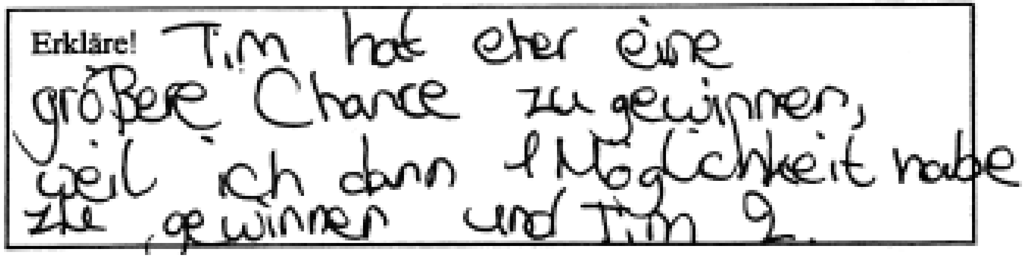
Item	**% pretest** ***n***	**% posttest** ***n***	**Difference** ***n***
5	0.0% *0*	47.4% *9*	+47.4 + *9*
6	9.5% *2*	52.6% *10*	+43.1 + *8*

Exemplary statements of students (task 5)

Task (5): You and your friend Tim play a game where you toss a coin twice. You win if you get tail both times. Tim wins if he gets tail and head when tossing the coin twice. Do you both have the same chances of winning? Explain!

“Yes, we have an equal chance of winning.” (statement in pre-test)


Erklăre! Ja wir haben die gleichen Chancen zu gewinnen


Erkläre! Tim hat eher eine größere Chance zu gewinnen, weil ich dann 1 Möglichkeit habe zu gewinnen und Tim 2.

“Tim has a better chance to win because I have one chance to win and Tim has two.” (statement in post-test)

Exemplary statements of students (task 6)

Task (6): Anna and Tom are playing a game. They throw a dice twice. If the sum of the two dice is "7", Anna wins. If the sum is 3 or 12, Tom wins. Do Anna and Tom have the same chance of winning? Explain!

“Tom will win because he has two numbers and has a better chance to win.” (statement in pre-test)

Jeil er hat zwei Zahlen und hat eine
u gewonnen.

6. Anna wird eher häufiger gewinnen als
Tom, weil die 7 mehr Möglichkeiten
bei 2 Würfeln hat, als die 3 und die 12

“Anna will win more often than Tom, because the 7 has more possibilities with two dice than the 3 and 12.” (statement in post-test)

Summary

- Elements of the three conceptualizations of randomness (see Batanero, 2015) can be already implemented in primary school classroom.
- The performance of the students from pre-test to post-test has increased considerably
- The selected statements in the post-test show that the students´ reasoning about the interpretation of events in multi-step chance experiments has developed in a positive way
  - Some students are now able to use, e.g. elements of the approach of Laplace to explain their statements when comparing the probability of events.

First „Implications“ (although not in a sense of experimental evidence)

- Statistical reasoning should be initiated as early as possible → in primary school
- The TinkerPlots sampler seems to be a powerful educational tool to facilitate the modelling and the data production process when simulating chance experiments in primary school → challenge: interpretation of the produced data (decision making based on data)
- The students should conduct the chance experiments as hand-on activities first and then use software like TinkerPlots for the modelling and data production process.

Outlook

- The teaching unit cannot only be of interest for primary school researchers and teachers but can also be adapted for implementation in secondary school classroom.
- The other data (working notes and interview data) collected in the frame of this research project are still under analysis.
  - Interview: insight into the cognitive processes of primary school students when modelling one/multi-step chance experiment with TinkerPlots

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Part 3

Developing reasoning about uncertainty of preservice teachers for primary school (Randomization tests with TinkerPlots)

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Short review of relevant literature

Reasoning in uncertainty

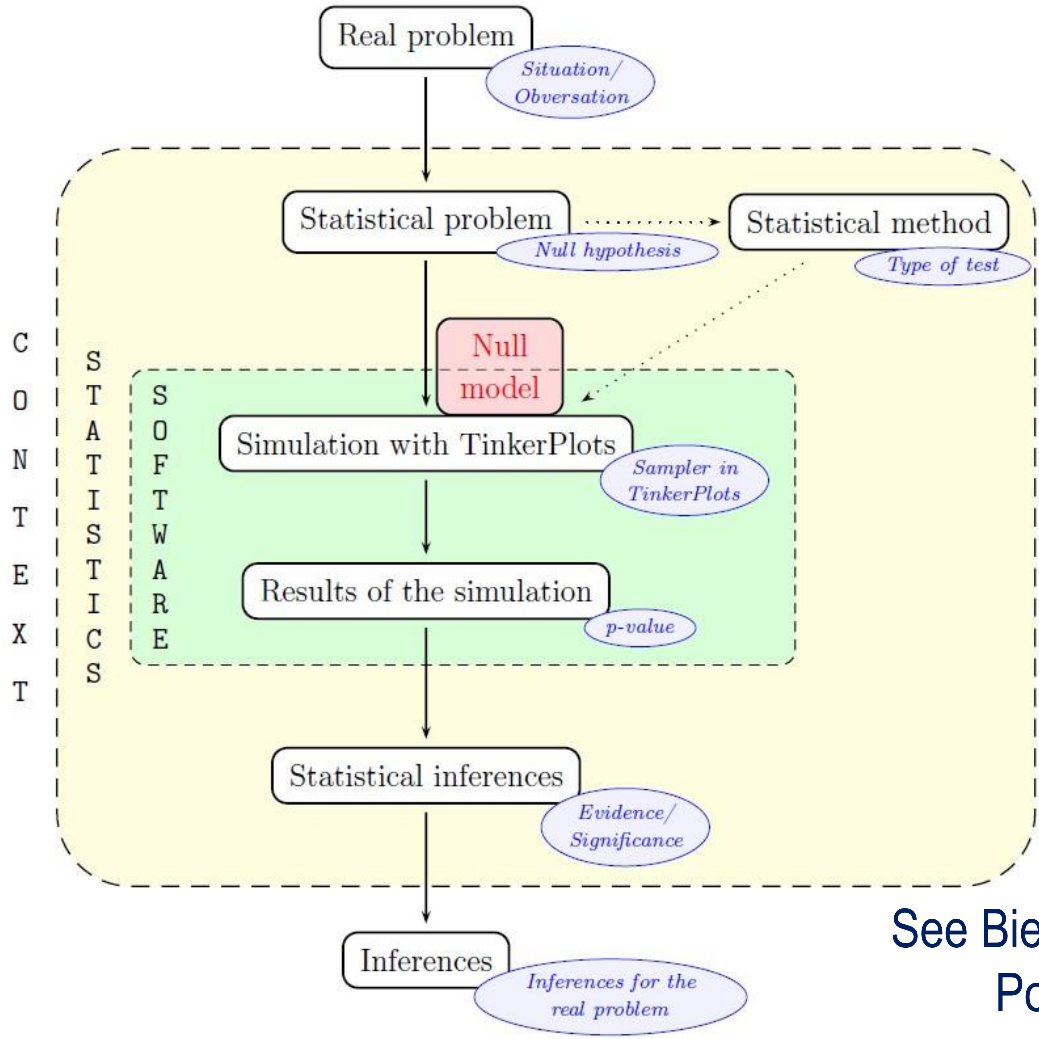
- Problems occur when a result found in a sample has to be generalized –> Randomization tests as an opportunity to make generalizations (Rossman, 2008)
- Cobb (2007): the 3 „Rs“: **R**andomize data production, **R**epeat by simulation and **R**eject any model that puts your data in its tail

Use of randomization tests in German school curriculum

- Randomization tests in textbooks are widely unknown at secondary and tertiary level
- Even hypothesis testing with *p-values* appears seldomly
- Formal hypothesis testing with pre-defined significance levels is a topic in most federal states, but often not obligatory

Research on how German preservice teachers cope with randomization tests see Frischemeier & Biehler (2014); Biehler, Frischemeier & Podworny (2015) and Podworny (2019)

Schematic model for doing a randomization test in a real world context



See Biehler, Frischemeier & Podworny (2015)

Randomization tests with TinkerPlots

Why TinkerPlots?

- Easy introduction to chance experiments via the sampler
- Powerful tool for modelling
- No formulas need
- Making a statistical process visible (drawing out of boxes; random assignment; no black box) (Schnotz, 2002)
- TinkerPlots as a „data factory“ for producing a new sample → Understanding distributions by modeling them (Konold et al., 2008)

An example:

Randomization test with TinkerPlots

Context:

Caffeine vs. non Caffeine (clinical study)

Does caffeine facilitate rapid movement?


Finger Tapping and Caffeine

Many people feel they need a cup of coffee or other source of caffeine to "get going" in the morning. The effects of caffeine on the body have been extensively studied. In one experiment,[14] researchers trained a sample of male college students to tap their fingers at a rapid rate. The sample was then divided at random into two groups of 10 students each. Each student drank the equivalent of about two cups of coffee, which included about 200 mg of caffeine for the students in one group but was decaffeinated coffee for the second group. After a 2-hour period, each student was tested to measure finger tapping rate (taps per minute). The students did not know whether or not their drinks included caffeine and the person measuring the tap rates was also unaware of the groups. This was a double-blind experiment with only the statistician analyzing the data having information linking the group membership to the

Table 4.4 *Finger tap rates for subjects with and without caffeine*

Caffeine	246	248	250	252	248	250	246	248	245	250	Mean ≈ 248.3
No caffeine	242	245	244	248	247	248	242	244	246	242	Mean = 244.8

Research question:

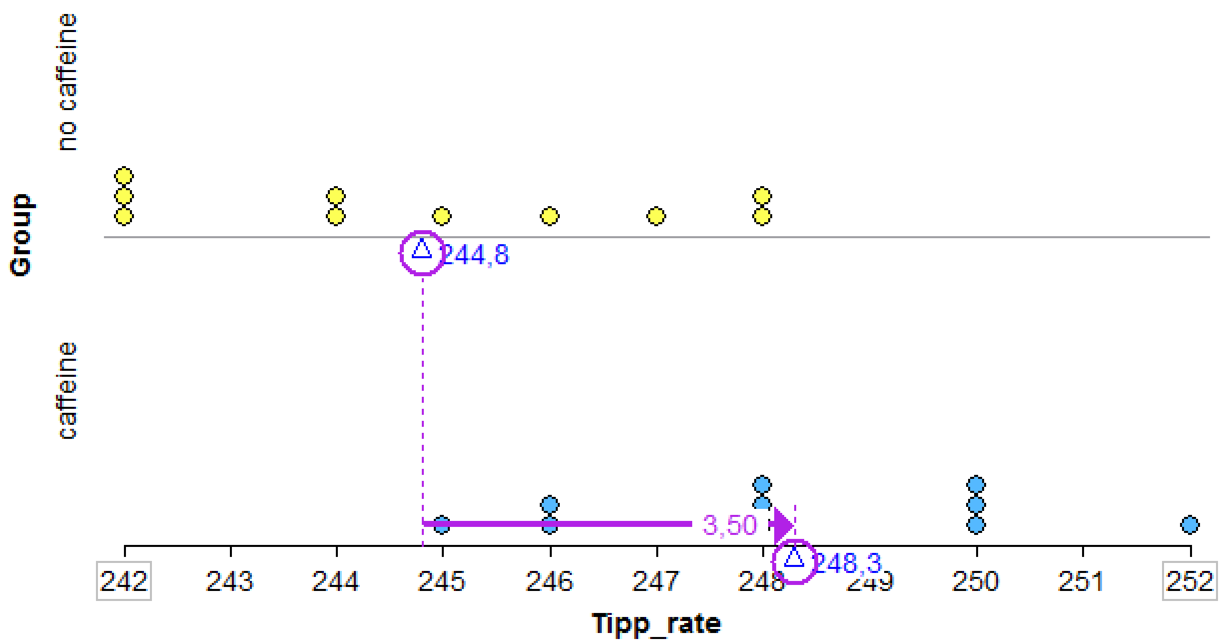
Does caffeine produces an increase in the average tap rate?

Example from Lock et al. (2013)

Difference of means of the two groups in the real data

Table 4.4 *Finger tap rates for subjects with and without caffeine*

Caffeine	246	248	250	252	248	250	246	248	245	250	Mean = 248.3
No caffeine	242	245	244	248	247	248	242	244	246	242	Mean = 244.8



248,3 - 244,8 = 3.5

Formulating an adequate null hypothesis

The group membership caffeine vs. non-caffeine has no effect on the fingertip rate

The difference between caffeine and non-caffeine users has occured at random

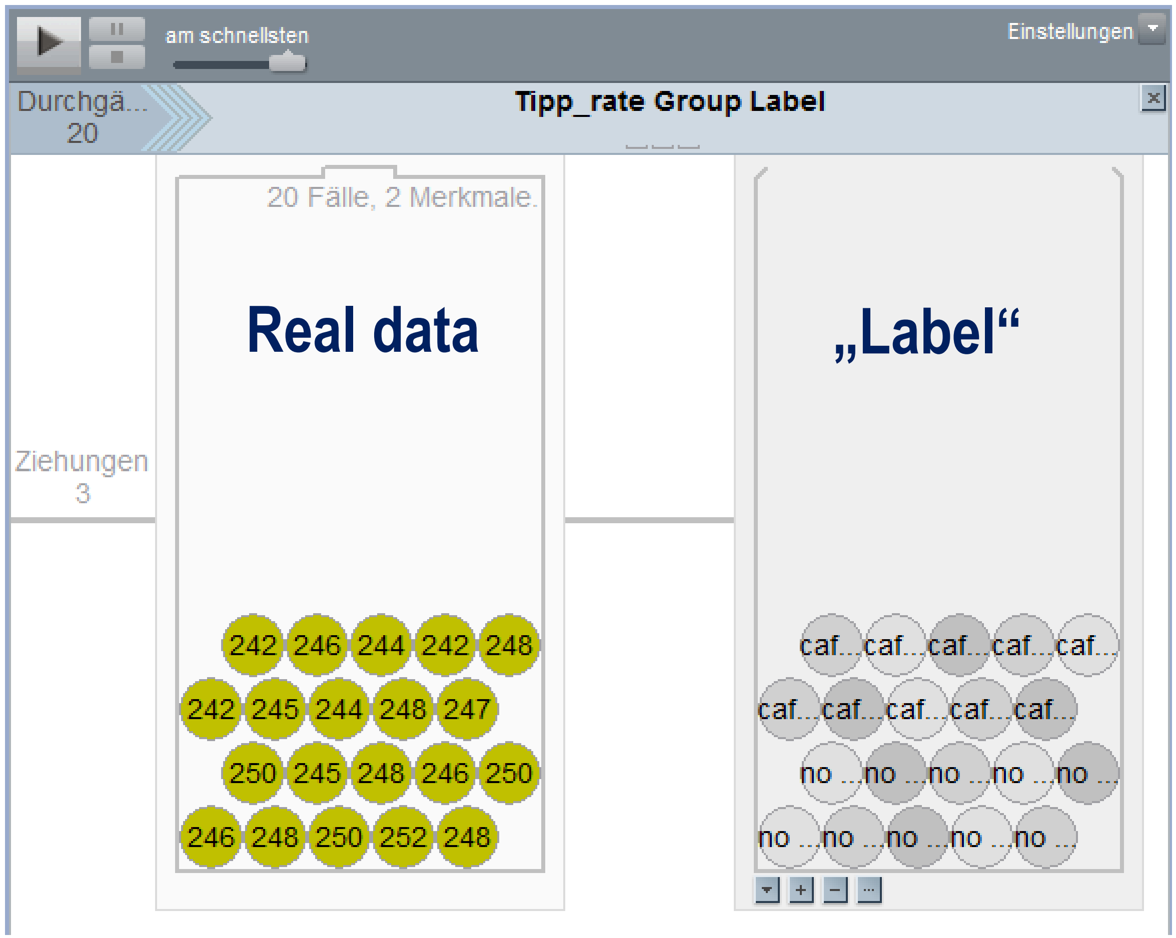
→ Idea: We simulate this Hypothesis in TinkerPlots

We investigate, how likely is it to get a difference of means 3.5 or larger in the case of randomly allocated groups (caffeine vs. non-caffeine)

Setting up the model in TinkerPlots

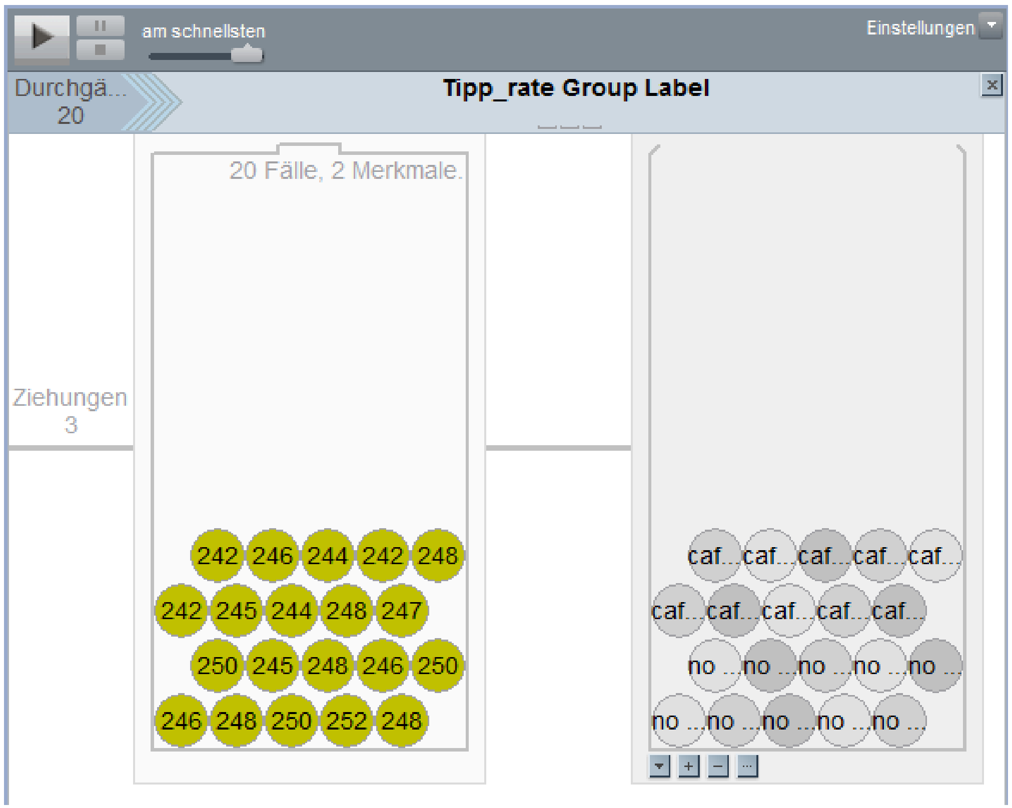
Repeat

Draw



without replacement

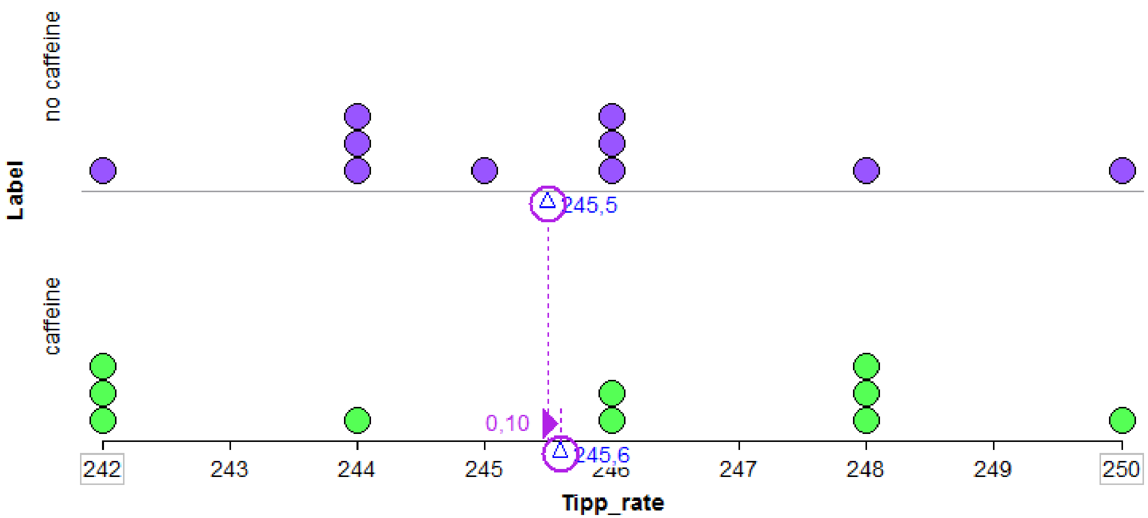
Running the TinkerPlots sampler



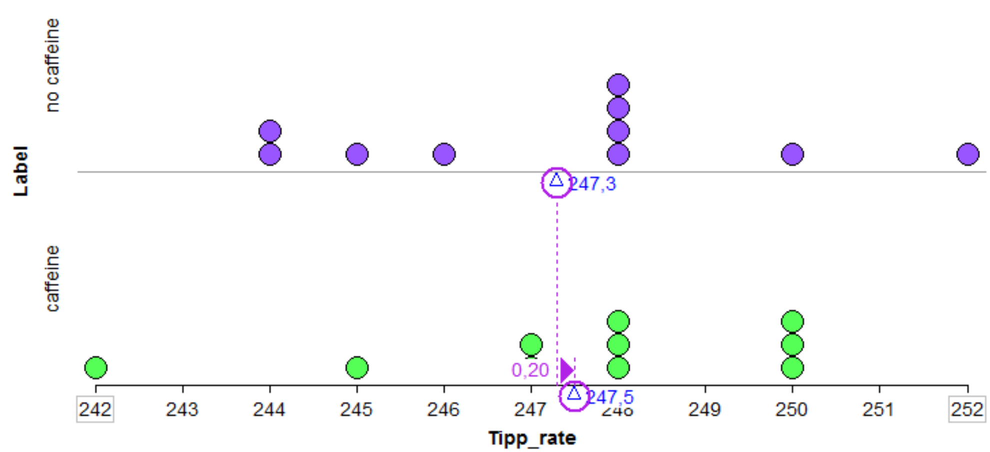
Ergebnisse von Zufallsmaschine 1 — Einstell

	Tipp_rate	Group	Label
1	247	no caffeine	no caffeine
2	248	caffeine	no caffeine
3	244	no caffeine	no caffeine
4	248	caffeine	no caffeine
5	245	caffeine	no caffeine
6	252	caffeine	caffeine
7	248	caffeine	caffeine
8	242	no caffeine	caffeine
9	245	caffeine	caffeine
10	246	caffeine	no caffeine
11	244	no caffeine	caffeine
12	250	caffeine	no caffeine
13	248	caffeine	no caffeine
14	250	caffeine	caffeine
15	250	caffeine	caffeine
16	247	no caffeine	caffeine
17	250	caffeine	no caffeine
18	248	caffeine	no caffeine
19	246	caffeine	caffeine
20	245	no caffeine	caffeine

Comparing two „random“ distributions after the labelling process



245,6 - 245,5 = 0.1



247,5 - 247,3 = 0.2

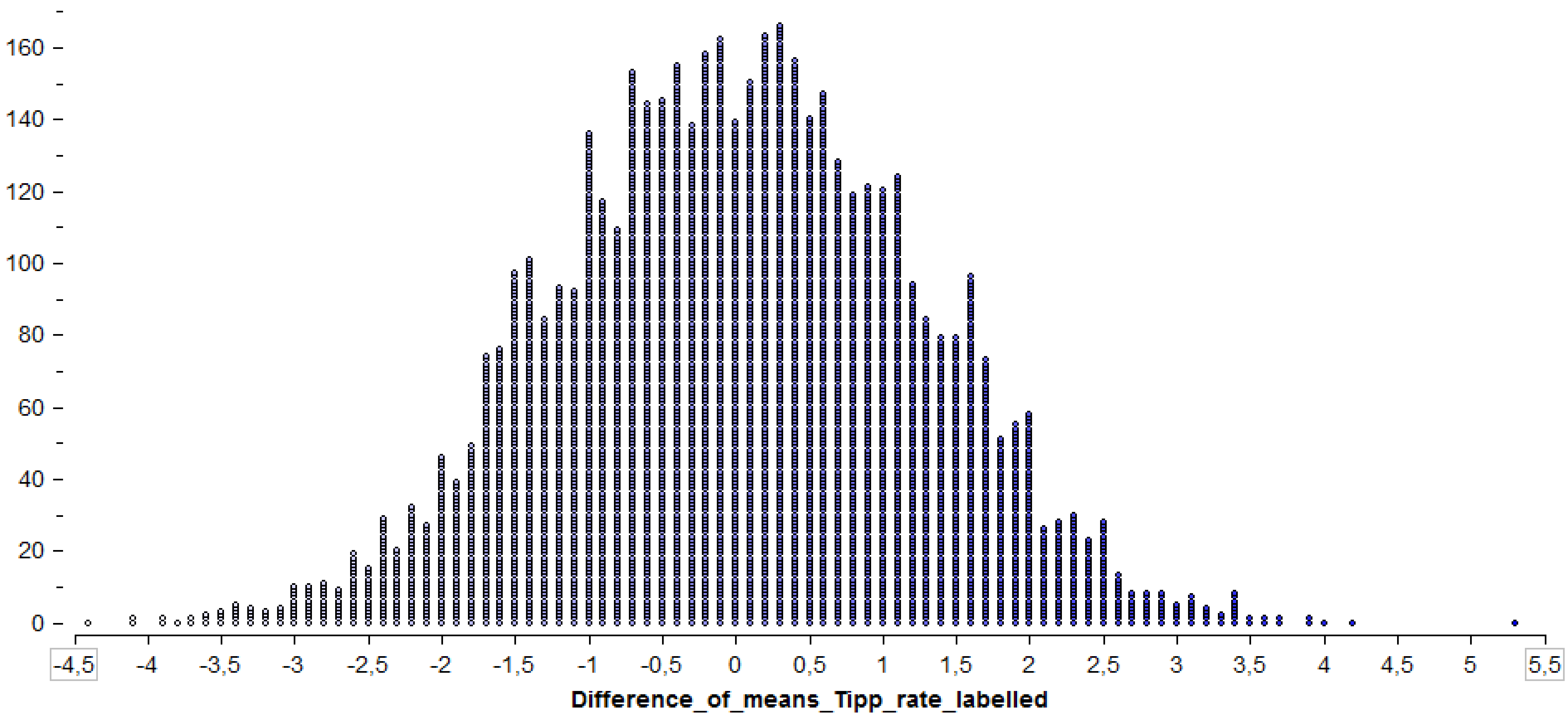
Collecting measures

Collect

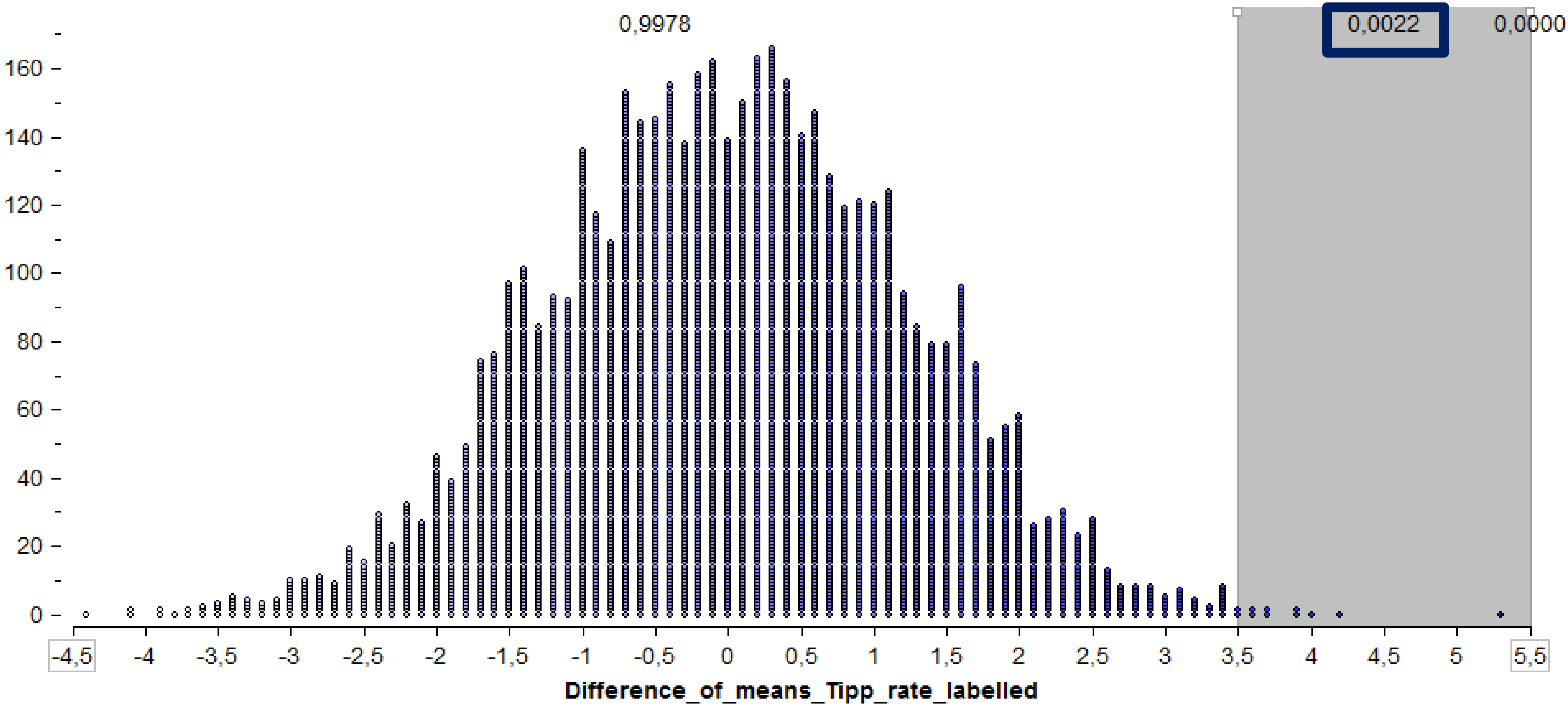
Messgrößen der Ergebnisse von Zufallsmasc... Sammeln 5000 Einstellungen

	Difference_of_means_Tipp_rate_labelled	<neu>
4990	0,5	
4991	0,7	
4992	-0,2	
4993	2,6	
4994	-0,6	
4995	0,1	
4996	-1,5	
4997	0,1	
4998	0,6	
4999	-1	
5000	0,5	

Identifying and interpreting p-value



Identifying and interpreting p-value



This shows clear evidence against our hypothesis that "the group membership caffeine vs. no caffeine has no effect on the fingertip rate".

Statistical steps of the randomization test - Overview

Step 1: Reading off the difference of the means of the groups in the dataset

Step 2: Formulating an adequate null hypothesis

Step 3: Describing the null model

Step 4: Formulating the test statistic

Step 5: Determining the p-value

Step 6: Drawing conclusions from the p-value

Typical problems & Implications

Problems cf. Frischemeier & Biehler (2014); Frischemeier (2017):

- Generation of adequate Null hypothesis
- Identifying p-value correctly
- Interpreting p-value correctly

Implications cf. Frischemeier & Biehler (2014); Frischemeier (2017):

- Discussing different null hypotheses (good vs. bad examples)
- Realizing hands-on allocation process before technical introduction with TinkerPlots
- Providing a scheme with the several steps of a randomization test
- Providing a guide how to interpret p-values

→ TinkerPlots as powerful tool to realize randomization tests

Supporting material: Randomization test scheme

Randomization test scheme

Task:	
1. **Observation** What is the difference in the dataset?	
2. **Null hypothesis** Formulate the null hypothesis, which will be assumed as true.	
3. **Simulation of the null hypothesis** Describe the modeling process, modeled with the sampler of TinkerPlots.	
4. **Test statistic** Define the test statistic.	
5. **P-Value** Read off the *p-value*.	
6. **Conclusions** Draw conclusions concerning your null hypothesis with the *p-value*.	

Supporting material: Handout „evidence“ - language

Hand-out:

*We have weak evidence against the null hypothesis, if $p \leq 10\%$.

*We have medium evidence against the null hypothesis, if $p \leq 5\%$.

*We have strong evidence against the null hypothesis, if $p \leq 1\%$.

*We have very strong evidence against the null hypothesis, if $p \leq 0.1\%$.

*We have no evidence against the null hypothesis, if $p > 10\ \%$.

Final thoughts on my lecture

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Thank you very much for your attention!

dafr@math.upb.de

Literature (Part 1, chosen)

- Bakker, A. (2004). Design research in statistics education - on symbolizing and computer tools. (Dissertation), University of Utrecht.
- Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distributions. In D. Ben-Zvi & J. Garfield (Eds.), The challenge of developing statistical literacy, reasoning and thinking (pp. 147–168). Dordrecht: Kluwer Academic Publishers.
- 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). Dordrecht/Heidelberg/London/New York.
- 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: 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: National Council of Teachers of Mathematics.
- Konold, C., & Miller, C. (2011). TinkerPlots 2.0. Emeryville: Key Curriculum Press.
- Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. Journal for Research in Mathematics Education, 33(4), 259–289.
- Konold, C., Robinson, A., Khalil, K., Pollatsek, A., Well, A. D., Wing, R., et al. (2002). Students' use of modal clumps to summarize data. In B. Phillips (Ed.), Proceedings of the International Conference onTeaching Statistics [CD-ROM], Cape Town. Voorburg: International Statistics Institute.
- Leavy, A.,Meletiou-Mavrotheris,M., & Paparistodemou, E. (2018). Statistics in early childhood and primary education: Supporting early statistical and probabilistic thinking. Singapore: Springer.
- Watson, J., Fitzallen, N., Wilson, K., & Creed, J. (2008). The representational value of HATS. Mathematics Teaching in Middle School, 14(1), 4–10.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223–248.

Literature (Part 2, chosen)

- Batanero, C. (2015). Understanding Randomness: Challenges for Research and Teaching. In K. Krainer & N. Vondrova (Eds.), Proceedings of the 9th Congress of the European Society for Research in Mathematics Education (pp. 34–49). Charles University in Prague: Faculty of Education and ERME.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. Educational Researcher, 32(1), 9-13.
- 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.
- Jackes-Schulte, A. (2016). Design, Durchführung Und Evaluation Einer Unterrichtsreihe Zur Entwicklung Der Kompetenz "Gewinnchancen Beim Doppelten Würfelwurf Einzuschätzen" in Jahrgangsstufe 4. (Bachelor of Education), University of Paderborn.
- Konold, C., & Miller, C. (2011). Tinkerplots 2.0. Emeryville, CA: Key Curriculum Press.
- Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2018). Statistics in Early Childhood and Primary Education: Supporting Early Statistical and Probabilistic Thinking. Singapore: Springer.
- Plückebaum, K. (2018). Simulieren Von Zufallsexperimenten in Der Primarstufe Mit Der Software Tinkerplots. (Master of Education), Universität Paderborn, Paderborn.
- Weustenfeld, W. (2007). Die Augensumme Zweier Würfel Voraussagen: Alles Nur Eine Frage Von Glück Oder Pech. Stochastik in der Schule, 27(3), 2-15.

Literature (Part 3, chosen)

- Biehler, R.; Frischemeier, D. & Podworny, S. (2015). Preservice Teachers´ Reasoning about Uncertainty in the Context of Randomization Tests. In A. Zieffler & E. Fry (Eds.): *Reasoning about Uncertainty: Learning and Teaching Informal Inferential Reasoning (S. 129-162)*. Catalyst Press..
- Cobb, G. (2007). The introductory statistics course: A Ptolemaic curriculum? *Technology Innovations in Statistics Education,* 1(1).
- Frischemeier, D. & Biehler, R. (2014). Design and exploratory evaluation of a learning trajectory leading to do randomization tests facilitated by TinkerPlots. In: B. Ubuz, C. Haser & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 799-809).
- Frischemeier, D. (2017). Statistisch denken und forschen lernen mit der Software TinkerPlots. Wiesbaden: Springer Spektrum.
- Hand, D. J., Daly, F., McConway, K., Lunn, D., & Ostrowski, E. (1993). *A handbook of small data sets*. Chapman and Hall: London.
- Konold, C. & Miller, C. (2011). TinkerPlots TM Version 2 [computer software]. Emeryville, CA: Key Curriculum Press.
- Lock, R. H., Lock, P. F., Lock, K., Lock, E. & Lock, D. (2013). Statistics: *Unlocking the power of data*. Wiley Global Education.
- Rossman, A. (2008). "Reasoning about Informal Statistical Inference: A Statistician´s View." *Statistics Education Research Journal.* 7(2): 5-19.