DEVELOPING DATA COMPETENCE IN PRIMARY SCHOOL

Daniel Frischemeier (dafr@math.upb.de)

UNIVERSITÄT PADERBORN
Die Unternehmer der Informationsgesellschaft

Introduction

"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)

In Germany the leading idea *data, frequency, and chance* sets requirements for students after finishing primary school (Hasemann & Mirwald, 2012) – fundamental components: *posing statistical questions*, *collecting data*, *getting to know how to represent data* and *interpreting representations of data*

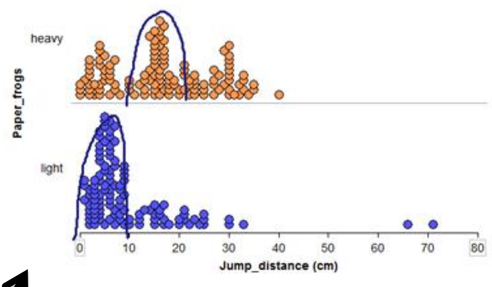
But: German textbooks and teaching materials for primary school often emphasize only singular aspects such as reading a diagram rather than introducing context-rich and complex data analysis projects as sustainable teaching-learning arrangement (Bakker, 2004)

→ **Goal of this research project**: Design of a teaching-learning arrangement to develop data competence (exploring real and meaningful data, using digital tools, working on data projects) in primary school

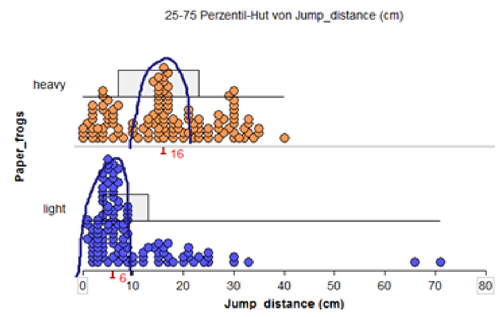
Developing data competence in primary school: Comparing groups as fundamental activities

- Group comparison activities include many of the fundamental statistical ideas like data, variation, representation, etc. (see Burrill & Biehler 2011).
- Proto-concepts like modal clumps and representations like hatplots can provide useful pre-stages for young learners for center, spread and shift (see Konold et al., 2002; Bakker, 2004; Watson et al., 2008).
- Modal clumps (Konold et al., 2002, p.1): "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"
- Watson et al. (2008) introduce the idea of hatplots: "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)

TinkerPlots



Group comparison using modal clumps

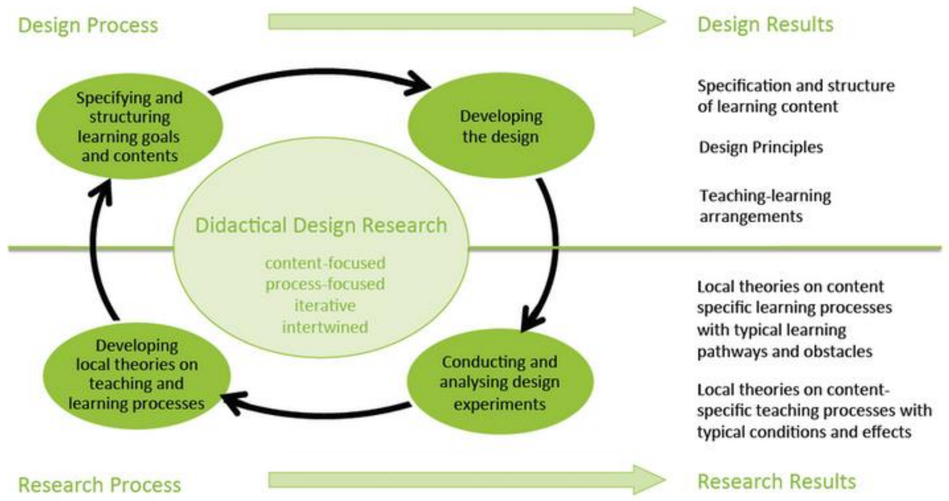


Group comparison using modal clumps, medians and hatplots

Design of a teaching-learning arrangement to lead primary school students to comparing groups

Major design ideas
Elements of the „Statistical Reasoning Learning Environment" (Garfield & Ben-Zvi, 2008):

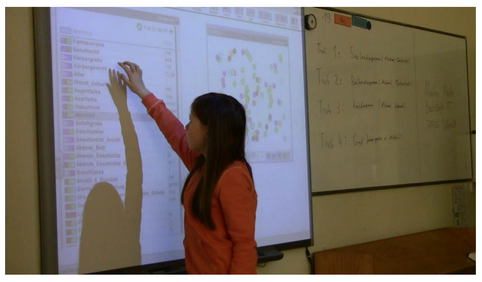
- PPDAC-cycle (Wild & Pfannkuch 1999)
- Working with real data (Engel 2007, Garfield & Ben-Zvi 2008)
- Using educational software TinkerPlots (Konold 2007)
- Implementing collaborative learning settings



Cycle of Didactical Design Research (Prediger & Zwetschler, 2013)

7 Modules of the teaching-learning arrangement
1 First basics in data analysis and getting to know the "Grundschulen NRW" questionnaire, using questionnaire to collect data in class
2 Statistical representations on different representational levels (enactive, iconic, symbolic)
3 Introduction to data analysis with TinkerPlots and creating stacked dot plots in TinkerPlots
4 Getting to know modal clumps, medians and hatplots to compare groups
5 Group comparison projects: Preparing the posters
6 Presentation of the group comparison posters
7 Conclusion of the project

Impressions from the classroom



Student demonstrates her data exploration in TinkerPlots on the interactive whiteboard

Students working in peers on group comparison task in TinkerPlots

Classroom research

Research questions
1. In which way does the group comparison competence of the students improve after the course?
2. How do the students like working on statistical projects and exploring real data with TinkerPlots?

Data collection

- Exercise sheets from the students
- TinkerPlots files from the students
- Data posters (final products at the end of the teaching unit)
- Post evaluation survey
- Pre-/post test

Participants
Twelve Grade 4 students (aged 10-11)
→ only little knowledge on statistics (collecting data and displaying them in tallies, reading of pie graphs).

Results on post evaluation survey

Survey (post evaluation) on students´ attitudes towards specific components of the teaching unit

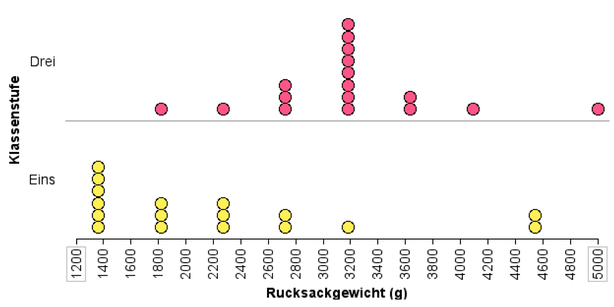
Data analysis with TinkerPlots

Please choose:

	☺	😐	☹	no answer
I liked the project very much.	12 (100%)	0 (0%)	0 (0%)	0 (0%)
I liked the learning about the basics in data analysis very much.	12 (100%)	0 (0%)	0 (0%)	0 (0%)
I liked the data collection very much.	12 (100%)	0 (0%)	0 (0%)	0 (0%)
I liked the creation of stacked dot plots very much.	8 (66.7%)	3 (25.0%)	0 (0%)	1 (8.3%)
I liked TinkerPlots very much.	12 (100%)	0 (0%)	0 (0%)	0 (0%)
I liked comparing groups with TinkerPlots very much.	11 (91.7%)	1 (8.3%)	0 (0%)	0 (0%)
I liked working with the drawing tool and the hats in TinkerPlots very much.	10 (83.3%)	2 (16.7%)	0 (0%)	0 (0%)
I liked the creation of the data posters very much.	11 (91.7%)	1 (8.3%)	0 (0%)	0 (0%)
I liked the presentation of the data posters very much.	10 (83.3%)	2 (16.7%)	0 (0%)	0 (0%)

Results on pre/post test

Group comparison task was handed out before attending to the teaching unit and the same task was handed out after attending to the teaching unit.
The concrete task was: "Do the first graders or the third graders tend to have heavier backpacks? Explain your answer."



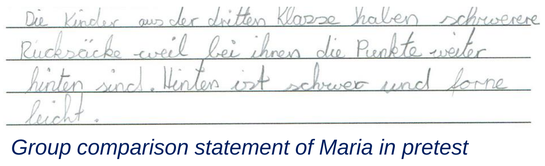
Task in pre-/post test: Distributions of PackWeight (g) in grades one and three

We used qualitative content analysis (Mayring 2015) to categorize and to explore which statistical explanations our students have used to state that there is a difference between both distributions.
→ Categories: Total sum, shift of points, shift of modal clumps, median, hat plots

Overview of results of pre- and post-test

No	Name	Correct statement (before)	Explanation (before)	Correct statement (after)	Explanation (after)
1	Tim	Yes	None	Yes	Median, Hat
2	Titus	Yes	None	Yes	Hat
3	Elton	No	None	Yes	Median, Hat, Shift MC
4	Noel	No	None	Yes	Median
5	Maria	Yes	Shift of Points	Yes	Median, Hat
6	Johannes	Yes	None	Yes	Median
7	Finnja	No	None	Yes	None
8	Emilie	No	None	Yes	Hat
9	Oliver	Yes	Shift of Points	Yes	Hat, Median
10	Liv	Yes	Total sum	Yes	Hat
11	Hannah	Yes	Shift of Points	Yes	Hat
12	Loris	Yes	None	Yes	Hat

Exemplary statements of students

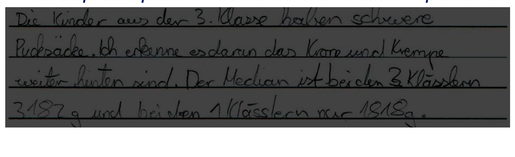


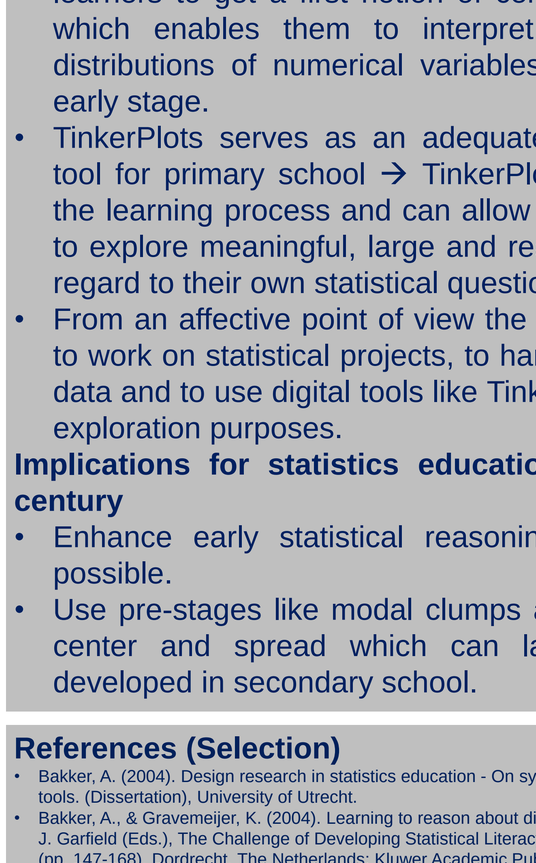
Group comparison statement of Maria in pretest

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

The children in class 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] 3182g and in the group of the grade 1students [is] 1818g.

Group comparison statement of Maria in posttest





Poster "Survey on heights – Big boys, little girls?" [Translation] as a product from the group comparison project of the group blue (Tim, Titus, Elton and Noel)

Conclusions

- The teaching-learning arrangement develops the competence to compare two distributions in larger, real data sets with TinkerPlots.
- Pre-stages like modal clumps can help young learners to get a first notion of center and spread which enables them to interpret and compare distributions of numerical variables already at an early stage.
- TinkerPlots serves as an adequate data analysis tool for primary school → TinkerPlots can support the learning process and can allow young students to explore meaningful, large and real datasets with regard to their own statistical questions.
- From an affective point of view the participants like to work on statistical projects, to handle meaningful data and to use digital tools like TinkerPlots for their exploration purposes.

Implications for statistics education in the 21st century

- Enhance early statistical reasoning as early as possible.
- Use pre-stages like modal clumps and hatplots for center and spread which can later be further developed in secondary school.

References (Selection)

- 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, The Netherlands: Kluwer Academic Publishers.
- Ben-Zvi, D. (2018). Introduction. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), Statistics in Early Childhood and Primary Education: Supporting Early Statistical and Probabilistic Thinking. Singapore: Springer
- 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.
- 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., & Miller, C. (2011). TinkerPlots 2.0. Emeryville, CA: Key Curriculum Press.
- 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.
- Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2018). Statistics in Early Childhood and Primary Education: Supporting Early Statistical and Probabilistic Thinking. Singapore: Springer.
- Prediger, S. & Zwetschler, 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. Enschede: SLO, Netherlands Institute for Curriculum Development, 407-424.
- Watson, J., Fitzallen, N., Wilson, K., & Creed, J. (2008). The representational value of HATS. Mathematics Teaching in Middle School, 14(1), 4-10.