Common patterns of thought and statistics: accessing variability through the typical

Christian Büscher

TU Dortmund University, Germany; christian.buescher@math.tu-dortmund.de

Allowing students to construct meanings of statistical concepts like variability requires building on their individual experiences. The notion of patterns of thought is utilized to conceptualize the difference between formal statistics and learners' initial thinking and to describe a pathway to bridge this gap between individual and mathematical thinking. Students' patterns of thought and the processes of their conventionalization are reconstructed in a qualitative study with n = 10 students of grade 7. An outlook is given on the possibility of connecting students' patterns of thought with formal statistics.

Keywords: Informal inferential reasoning, design research, concept development.

Introduction

One of the aims of statistics education is to foster informal inferential reasoning (IIR), the ability and disposition to use data in order to reason about some wider universe (Makar, Bakker, & Ben-Zvi, 2011). In IIR, statistical concepts are combined with contextual knowledge under certain statistical norms and habits, such as a "critical stance towards data" (see Makar et al., 2011 for a more thorough overview). While the framework of IIR can explain the *role* of statistical concepts in producing informal statistical inferences (ISI, see Makar & Rubin, 2009), it does not account for the *development* of statistical concepts. Accordingly, there exists a lack of research concerning statistical concept development of students with little experience in statistics. Tracking of the development of concepts, Bakker and Derry (2011) draw on the background theory of inferentialism, linking students' emerging concepts in inferential practices with IIR. Looking at the micro level of students' reasoning processes, Bakker and Derry argue that students can show complex ideas regarding statistical concepts such as center, variation, distribution, and sampling. These ideas are not formally articulated, but can rather be seen as "precursor notions" of regular statistical concepts (Bakker & Derry, 2011, p. 20).

These precursor notions could provide a promising basis for the development of more formal statistical reasoning. The challenge remains however to design tasks that draw on these precursor notions in order to foster students with little prior experience in formal statistics. These tasks need to draw on students' individual ways of reasoning as a resource in order to develop statistical concepts meaningful to the students. This requires paying careful attention to students' learning processes, and to uncover the links between formal statistical concepts and students' everyday thinking. This study aims at reconstructing students' individual ways of thinking on a micro level in order to find ways to connect students' everyday thinking to regular statistical concepts.

Patterns of thought in everyday thinking

The mathematician and philosopher Wille (1995) argues that in order for mathematics to become learnable and possibly meaningful to non-experts (i.e. the general public), the discipline of mathematics itself has to be restructured in a program he coined *Generalistic Mathematics*

(“Allgemeine Mathematik”). Wille calls for mathematicians to reveal the reasons, the aims, the common patterns of thought, and the boundaries and dangers of mathematical concepts and of whole mathematical theories. This however must not be done in the language of mathematical theory. Opening mathematics to the general public necessitates the use of common language, rather than specialized vocabulary, to describe those reasons, aims, patterns, and boundaries.

This approach to make mathematics open and meaningful resonates well with Lengnink and Peschek (2001), who see the challenge and goal of mathematics education in explicating the connection between *everyday thinking* and *mathematical thinking*. For them, mathematical thinking manifests as conventionalized everyday thinking. In this way, confidence intervals can be seen as a conventionalized form of the generalizations taking place in daily life: being reasonably sure that some observation of daily life can be taken as true, with a certain intuitive degree of variation.

This places a firm emphasis on the importance of common thought patterns and practices, so that a task of mathematics education becomes the identification of these patterns of thought. While Wille (1995) sees this task in the hands of (the philosophy of) mathematics, Prediger (2008) points out that empirical insights into mathematical learning processes form a valuable and even necessary basis for identifying connections between everyday thinking and mathematical thinking. This also means that thought patterns cannot be approached in a general way, but rather are tied to the specific mathematical content of the learning process under investigation. This study adopts the approach outlined by Prediger (2008), in order to reveal patterns of thought and their connection to the statistical concepts of center and spread when comparing distributions.

The typical as a common pattern of thought

One way students’ thinking differs from formal statistics is in the use of measures. In statistics, measures such as median or standard deviation function as highly specialized tools for talking about statistical concepts like center or variability. In their intuitive approaches to statistics, students use strategies such utilizing “modal clumps” instead (Konold et al., 2002) to point out ‘the majority’ of the data. Since to the students the location and the width of the clump both are important, these modal clumps can simultaneously address the center of a distribution as well as a form of spread. Thus, in their everyday language, learners integrate many different statistical concepts that would formally be strictly distinguished through use of different specialized measures (Makar & Confrey, 2005).

This raises the question on what would be promising patterns of thought to build on. One such candidate would be the practice of identifying ‘typical’ values or ranges of values within datasets. Some research indicates that ‘typical’ might be a good way for students to think about the average (Makar, 2014), while other research finds ‘typical’ or ‘normal’ to be a term for talking about ideas combining center and spread (Büscher, 2016a; Büscher, 2016b). Thus, the pattern of thought of identifying the ‘typical’ of a distribution seems a promising candidate as a resource for concept development, although it remains unclear for which concepts exactly.

Patterns of thought and concept development

To address the question of how patterns of thought can support concept development, this study follows a conceptual change approach (Duit & Treagust, 2003). Learning is understood as a restructuring of prior conceptions, occurring when these conceptions no longer satisfactorily explain phenomena. The goal of statistics teaching is then to initiate the development of conceptions into statistical concepts. Since statistical concepts show a high degree of connection to each other, learning trajectories in statistics however should not address concepts in an isolated way, but rather in a holistic way (Bakker & Derry, 2011). This calls for organizing structures that (a) connect to learners' prior conceptions, (b) holistically address statistical concepts, and (c) lead to regular statistical concepts. Patterns of thought can provide just such a structure, as they encompass different prior conceptions and thus enable the connection of everyday thinking to mathematical thinking.

Research questions

The theoretical background of this study suggests that designing learning trajectories towards statistical concepts should start from fruitful patterns of thought. It is thus an issue for empirical investigations to find those patterns of thought in students' thinking which can indeed serve this function as starting points in learning trajectories. Looking at what is 'typical' was identified as one potential pattern of thought that could result in such concept development. This study therefore aims to answer the following research question: *What concepts do students address and develop when conventionalizing the vague pattern of thought of 'identifying the typical'?*

Research design

This study adopts the methodological framework of topic-specific didactical design research (Prediger & Zwetzschler, 2013) with a focus on learning processes (Prediger, Gravemeijer, & Confrey, 2015). This approach aims at providing empirically grounded local theories on topic-specific learning processes as well as design principles and concrete teaching-learning arrangements for the topic. While the framework utilizes iterative cycles of design experiments, this study reports on the third cycle of design experiments.

Data gathering

Design experiments were conducted with five pairs of grade 7 students (aged 12 – 14) who had only little experience with statistics within the mathematics classroom. Experiments consisted of two sessions of 45 minutes each, with each session having its own arrangements of data and tasks. They were fully videotaped and partially transcribed. The process-focused analysis of the video data from the first session of the design experiments allows to reconstruct the development of students' patterns of thought, in their relation to task design.

Task Design

For designing the teaching-learning arrangements, various design principles have been implemented throughout the different cycles of design research, two of which play an important part in this study.

Drawing on 'typical' as pattern of thought. As outlined above, connecting to the pattern of thought of characterizing what is 'typical' of a certain distribution can potentially provide a starting point for processes of conventionalization that lead to meaningful use of formal statistics. Therefore, 'typical' has to be explicitly addressed in the task design, and the setting of the task has to provide a context in which this pattern of thought can naturally occur.

Criticizing conventionalizations. As Lengnink and Peschek (2001) point out, mathematical learning has to explicitly address the relation between everyday thinking and mathematical thinking. Following this, it is not enough for a task to just utilize students' patterns of thought and to encourage conventionalization of those patterns of thought. It is rather these conventionalizations that have to become the object of investigation. Under a Generalistic Mathematics perspective this could mean addressing reasons, aims, and boundaries of these conventionalizations.

These design principles were realized in the design of the *Antarctic weather task*. The initiated activity puts the students into the role of consultants to researchers at the Norwegian Antarctic research station *Troll forskingsstasjon*. In the first phase of the task, the students investigate dot plots of daily temperatures for the month of July 2004 and predict the weather for next July by giving a distribution of ten days. In Phase 2, additional data for July 2002 and 2003 were included (Figure 1). Predicting the weather was chosen as activity because it is rooted in everyday thinking, combining experiences of short-term variability (one can never be too sure about the weather…) with long-term signals (… but there are typical temperatures after all).

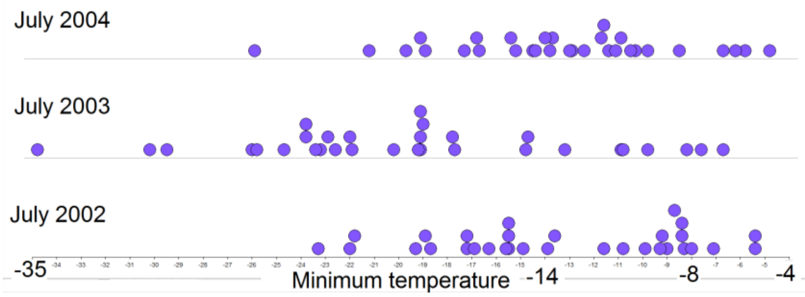


Figure 1: Distributions of the Antarctic weather task (translated from German)

In order to support conventionalization of patterns of thought, the third phase introduces a design element called *report sheets* (cf. Figure 2 and 3). The report sheets are introduced to serve as a brief summary of the Antarctic weather in July. They combine graphical representations and the use of measures with a brief inference. First, the students are asked to fill out their own report sheet. After that, in the fourth phase, the students receive three different filled-in report sheets by fictitious students (Figure 2). These filled-in report sheets differ in their interpretation of typical. This serves as a basis for discussing and criticizing the different conventionalizations of typical, as the students are asked to evaluate the correct use of typical.

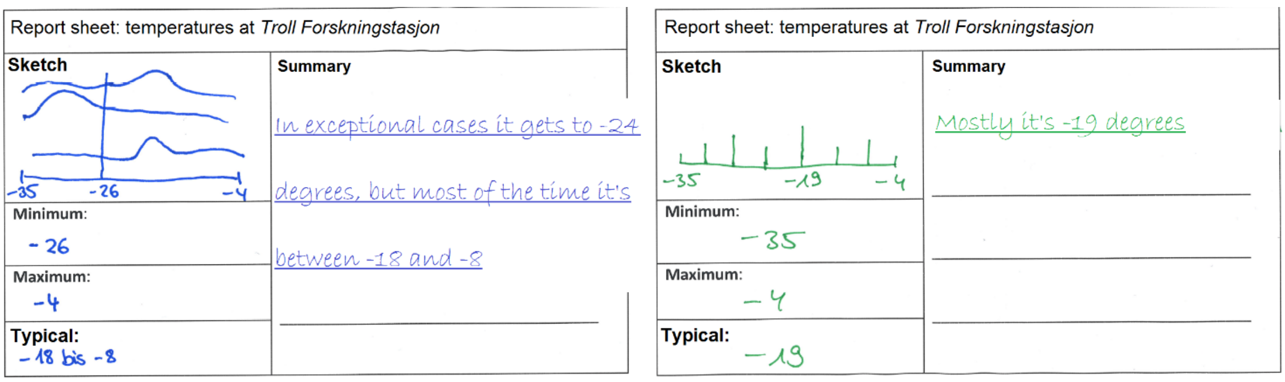


Figure 2: Fictitious students' filled-in report sheets (translated from German)

Data Analysis

The students' patterns of thought were reconstructed in an interpretative analysis using concepts-in-action and theorems-in-action from Vergnaud's (1996) theory of conceptual fields. Concepts-in-action are "categories (objects, properties, relationships, transformations, processes etc.) that enable the subject to cut the real world into distinct elements and aspects, and pick up the most adequate selection of information according to the situation and scheme involved" (Vergnaud, 1996, p. 225). Theorems-in-action are statements held to be true by the learner.

Which concepts-in-action and theorems-in-action are activated depends on the pattern of thought utilized by the learners. Thus, patterns of thought are conceptualized as groups of concepts-in-action concurrently occurring in the learners' activity. Concepts-in-action and theorems-in-action are reconstructed from the students' point of view, and do not necessarily correspond to regular statistical concepts. In the analysis, the reconstructed concepts-in-action are symbolized by ||…||, while theorems-in-action are denoted by <…>.

Empirical Insights: The case of Maria and Natalie

The first snapshot starts with Phase 2 of the Antarctic weather task. After getting the additional data of the years 2002 and 2003, Maria and Natalie, Grade 7, try to explicate their view on the data.

1	Maria:	We are pondering what the relationship, like, how to…
2	Natalie:	Yes, because we want to know what changes in each year. And we said that there *[2003]* it came apart.
[…]		
8	Maria	Yes, I think it *[2004] i*s somehow similar to that *[2002],* but that one *[2003]* is different.
9	Natalie:	Like here *[points to 2004, around -12 °C]* are, like, like the most dots, and here *[2002, -12 °C]* are almost none. And there *[2002, -8 °C]* are the most and here *[2004, -8 °C]* are almost none.

This excerpt serves as an illustration of the starting point in the students' reasoning. The students are trying to characterize the differences observed in the distributions. At first, the students formulate differences between 2003 and 2002/2004 in terms of ||*spread*||: in 2003, the temperatures "came apart" (#2). Another difference they notice is the difference of the location of the ||*center*|| between 2002 and 2004, indicated through modal clumps ("the majority", #9).

It is important to note that at this stage, the students face difficulties in trying to express their findings. The distributions of 2004 and 2002 are found to be "somewhat similar" (#8) though "different" (#8) to 2003, with further explanation supplied by Natalie through use of gestures and improvised vocabulary ("like, like the most dots", #9). Few minutes later, the students find a way to deal with the complexity.

21	Maria:	Well, we first should look at how many degrees it has risen or fallen. Generally. In two years.
[…]		
27	Natalie	You mean average, like…
28	Maria	The average, and then we look at how the average changed in two years.

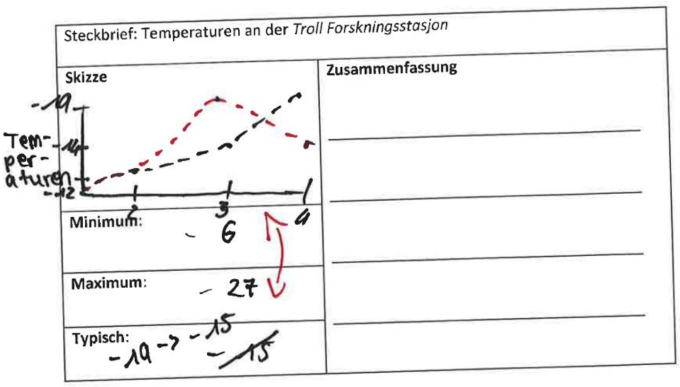
By introducing the notion of reducing the distributions to a ||*general value*|| ("Generally", #21), the students are able to handle the complexity of the differences between the distributions. For this general value, they appear to already know an adequate measure: the ||*average*||. To the students, <*the average represents the general value of a distribution*>. The average acts as a summary to be used in further procedures, as <*the differences between distributions can be described by differences in general values*>.

In the following exchange, after having estimated the averages to be -12 (2002) and -14 (2004), Maria and Natalie try to use their result for a linear extrapolation of the weather in 2015 to be predicted.

41	Natalie:	Wait. If it gets colder by 2 °C in two years, then it gets colder by 1 °C each year, so we have to…
42	Maria:	Nah, eh, yeah okay
43	Natalie:	13 °C colder average temperature. Right?
44	Maria:	Yes.
45	Natalie:	But that's too much, isn't it?

When reflecting on their result however, the students realize that a decline of the average temperature by 13 °C is not a realistic proposition (#45). While their knowledge of the real-world context helps them to identify this contradiction, they are not able to find another solution. In the minute following (not shown here), the students stay insistent in using the average and a linear extrapolation of the trend. It is important to note that, at this stage, their ideas concerning ||*spread*|| as expressed in the first excerpt seem to have disappeared, replaced by the stronger notion of ||*general value*|| expressed through the more conventionalized form of ||*average*||.

The design experiment progresses through the third phase, in which the students create their own report sheet (Figure 3). The analysis picks up at beginning of the fourth phase, with the students comparing the different interpretations of 'typical' in the filled-in report sheets.



Translations:

"Report sheet: temperatures at Troll Forskningsstasjon"

Skizze – Sketch

Typisch – Typical

Zusammenfassung – Summary

Temperaturen - Temperatures

(The black graph was drawn first, labeled a mistake, and immediately replaced by the red graph. Typical was first assigned as -15, then after the fourth phase changed to -19 to -15.)

Figure 3: Maria and Natalie's report sheet

Comparing the different interpretations of 'typical', Maria and Natalie are intrigued by the possibility to use an interval to characterize 'typical'. This leads them to reflect on their use of the average.

61	Natalie:	But the average temperature isn't really typical, is it?
62	Maria:	What, typical? Of course the average temperature is the typical.

[…]

66 Maria: Well, no. Typical is more like where the most… no…

67 Maria: The average temperature isn't the typical after all. Because it's only the general, the whole. The typical would be for example for this *[2004]* here *[points to -14 on the 2004 dot plot]*.

68 Natalie: Typical I think simply is what is the most or the most common.

The students utilize Typical to differentiate between two different ideas: The ||*general value*|| that is expressed through the ||*average*|| ("the general, the whole", #67), and the ||*most common*|| part of the distribution, expressed through the ||*Typical*|| ("the most common", #68) – although at this point it is not yet clear if Typical consists of a number or an interval. With the ||*most common*|| corresponding to the notion of ||*center*|| expressed earlier ("the most dots", #9), Typical seems to help the students to express ideas that got swept aside by the more conventionalized average. Both, average and typical, start to act as conventionalized tools for talking about specific aspects of distributions.

Some minutes later, Natalie summarizes her view on the relation between 'typical' and average.

81 Natalie: And average is pretty imprecise, because it doesn't say anything about a single day. And with typical, I'd say, that it's a span between two numbers, because that way you can better overlook how it is most of the time.

In the end, typical and average provide two different applications. Whereas the average acts as a summary, 'typical' gives an overview into a distribution. The average can be used in comparing distribution in an efficient way, while 'typical' gives an insight into a range of 'normal' or 'expected' temperatures, to which any single day can be compared. In this way, 'typical' combines aspects of ||*center*|| and ||*spread*||.

Conclusion

The aim of this study was to examine the interaction of concept development and students' patterns of thought. The students showed two different patterns of thought: Characterizing data through a general value and through a range of typical values. These patterns differed in their degree of conventionalization. While the general-pattern was addressed through use of the average, the students lacked a conventionalization corresponding to the typical-pattern. This resulted in the typical-pattern to be suppressed in favor of the general-pattern, as the ideas addressing spread disappeared. Only when the students were supplied with different conventionalizations of Typical, they were able to reconnect to their typical-pattern. This then allowed them to express ideas combining center and spread.

This identification of thought patterns provides a promising starting point. Work still remains however utilizing these thought patterns to develop regular statistical concepts. This could be achieved by reconceptualizing formal statistics in terms of a typical-pattern. One possible connection could be interpreting the 'box' of a box plot as the typical area of a distribution. Additionally, more insight into processes of conventionalization is needed in order to be able to successfully guide students on their way to meaningful statistics (for one example see Büscher, 2016).

References

Bakker, A., & Derry, J. (2011). Lessons from Inferentialism for Statistics Education. *Mathematical Thinking and Learning*, *13*(1-2), 5–26.

Büscher, C. (2016a). Entwicklung von informellen statistischen Maßen zwischen Werkzeugen und Objekten. In *Beiträge zum Mathematikunterricht 2016.* Münster: WTM-Verlag.

Büscher, C. (2016b). *Students' Informal Measures between Objects and Tools*. Paper presented at the 13th International Congress on Mathematical Education, Hamburg.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(6), 671–688. doi:10.1080/09500690305016

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.

Lengnink, K., & Peschek, W. (2001). Das Verhältnis von Alltagsdenken und mathematischem Denken als Inhalt mathematischer Bildung. In K. Lengnink, S. Prediger, & F. Siebel (Eds.), *Darmstädter Schriften zur allgemeinen Wissenschaft: Vol. 2. Mathematik und Mensch. Sichtweisen der allgemeinen Mathematik.* Mühltal: Verlag Allgemeine Wissenschaft, HRW e.K.

Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics*, *86*(1), 61–78.

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The Reasoning Behind Informal Statistical Inference. *Mathematical Thinking and Learning*, *13*(1-2), 152–173.

Makar, K., & Confrey, J. (2005). Variation talk: Articulating meaning in statistics. *Statistics Education Research Journal*, *4*(1), 27–54.

Makar, K., & Rubin, A. (2009). A Framework for Thinking about Informal Statistical Inference. *Statistics Education Research Journal*, *8*(1), 82–105.

Prediger, S. (2008). Do You Want Me to Do It with Probability or with My Normal Thinking? Horizontal and Vertical Views on the Formation of Stochastic Conceptions. *IEJME-Mathematics Education*, *3*(3), 126–154.

Prediger, S., Gravemeijer, K., & Confrey, J. (2015). Design Research with a Focus on Learning Processes: An Overview on Achievements and Challenges. *ZDM*, *47*(6), 877–891. doi:10.1007/s11858-015-0722-3

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 - Part A: An Introduction* (pp. 409–423). Enschede, the Netherlands: SLO.

Vergnaud, G. (1996). The Theory of Conceptual Fields. In L. P. Steffe (Ed.), *Theories of mathematical learning* (pp. 219–239). Mahwah, N.J.: L. Erlbaum Associates.

Wille, R. (1995). „Allgemeine Mathematik“ als Bildungskonzept für die Schule. In R. Biehler (Ed.), *Mathematik allgemeinbildend unterrichten. Impulse für Lehrerbildung und Schule* (2nd ed., Vol. 21, pp. 41–55). Köln: Aulis-Verl. Deubner.