STUDENTS' EMERGENT MODELLING OF STATISTICAL MEASURES – A CASE STUDY

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

SUSANNE SCHNELL
University of Cologne, Germany
susanne.schnell@uni-koeln.de

ABSTRACT

The present study investigates the processes of how German middle school students (age 14) interpret, contrast and evaluate different (informal) statistical measures in order to summarise and compare frequency distributions. To trace the developing insights into the properties of these measures, this paper uses the 'emergent modelling' perspective: measures are understood as models, which can either be used to make sense of a given situation or to reason about the statistical measures themselves, e.g. in terms of when they can be applied adequately. The emergent modelling approach is used (1) as a theoretical framework for describing students' conceptual development, and (2) as a design heuristic for developing a teaching-learning arrangement aiming at developing insights about (frequency) distributions and statistical measures. In the qualitative analysis of a design experiment, two students' emerging contextual and statistical knowledge is identified, revealing the intertwined nature of both types of knowledge. Overall, this paper illustrates the important role the emergent modelling perspective can play for designing as well as describing students' learning pathways in statistics education.

Keywords: *Statistics education research; Design research; Informal statistical measures; Contextual knowledge*

1. INTRODUCTION: THE ROLE OF STATISTICAL MEASURES

While distributions can be understood as being at the heart of statistical activities and as the 'fundamental given of statistical reasoning' (Wild, 2006, p. 10), comparing them and drawing inferences is known to be challenging for students (Bakker & Gravemeijer, 2004; Ben-Zvi, 2004; Ben-Zvi & Arcavi, 2001). This is due to the fact that distribution as a 'conceptual entity' (Bakker & Gravemeijer, 2004) – or as we will argue an overarching statistical model – consists of a variety of different, intertwined concepts such as centre, spread, density and skewness (Bakker & Gravemeijer, 2004; Ben-Zvi, 2004). These concepts can be addressed by statistical measures (Bakker & Gravemeijer, 2004). Regular measures in descriptive statistics (e.g. mean, mode, median, or standard deviation) describe and summarise certain aspects of a given data set. Regular measures from inferential statistics like hypothesis tests or interval estimations are intended to draw inferences from the given data to the underlying population. Crucial to these considerations is that measures emphasise only certain aspects of distributions and the corresponding data while neglecting

Statistics Education Research Journal, 16(2), 144-162, http://iase-web.org/Publications.php?p=SERJ

others. Depending on the given problem, one measure can be more useful than another or a combination of different statistical measures may be necessary. Therefore, the properties of the measures themselves such as robustness, but also scope of applicability or validity in certain situations, have to be taken into account.

Whereas the aforementioned measures are regularly used by statisticians and are highly formalised, research shows how learners come up with their own ways of describing data sets, which we will call ‘informal measures’. For instance, Pfannkuch, Regan, Wild, and Horton (2010) elaborated on how students draw on visual aspects of frequency distributions presented as dot plots and summarise them by ‘gaps’ or ‘clusters’. These informal measures can be understood as students’ ways of making sense of situations and as starting points for productive learning pathways through which a profound understanding of regular measures and the concept of a distribution emerge (Bakker & Gravemeijer, 2004).

In this contribution, we will argue that formal and informal measures can be understood as *models of* certain situations or as *models for* more abstract reasoning. Therefore, we will first briefly elaborate on the classical perspective on modelling in mathematics education, which we will call the perspective of modelling as translating, before introducing the emergent modelling approach by Gravemeijer (1999, 2007). This presents an alternative perspective which is fruitful for describing and designing learning pathways in statistics education. After the description of the research design and tasks, we will present empirical excerpts from design experiments with students from a German middle school (age 14). The aim is to illustrate processes of emergent modelling with regard to informal statistical measures when students are asked to compare and reason with frequency distributions. The analysis will use emergent modelling as an analytic framework to show ways in which context can play an important part in informing the development of statistical knowledge, and vice versa.

2. THEORETICAL BACKGROUND

2.1. MODELLING AS TRANSLATING

In mathematics education research, a common perspective on the process of modelling (c.f. Blum & Niss, 1991; Kaiser, 2005; Lesh & Lehrer, 2003) portrays mathematics and the real world as distinct counterparts which can be connected by translating between them. In this view, a mathematical model is a conceptual system which constructs, describes or explains the structural characteristics of a specific situation (Lesh & Doerr, 2003). The activity of modelling is understood as a cyclic activity of moving between reality and models: an initial open situation in the real world is *idealised* into a real model by identifying and specifying relevant structures. Next, this real world model is *mathematised* by quantifying, systematising, and algebraising relevant data, concepts, relationships, patterns etc. (Blum & Niss, 1991; Lesh & Lehrer, 2003). The developed mathematical model can be *processed* within mathematics and the results are *interpreted* against the backdrop of the initial real world situation. If the validation (i.e. the adequacy of the results and solutions) is not satisfying, the cycle can restart with a new approach to the idealization of the real situation. Modelling problems in reality-based contexts are perceived as challenging but important for students in learning to apply mathematics (Kaiser, 2005).

The modelling cycle is prominently used in many areas of mathematics education, and also in statistics education (Eichler & Vogel, 2013). For statistics, this cycle acknowledges ‘the recognition of the need for data’, which according to Wild and Pfannkuch (1999) is one of the foundations of statistical thinking. The data collected in order to solve a problem can then be understood as the real model (Eichler & Vogel, 2013). The activity of *mathematising* means to fit a statistical model (which can be as simple as an arithmetic mean but also more complex as the normal distribution) to the existing data. As Eichler and Vogel (2013) put it, the purpose of this activity is to identify and describe structures in the fog of the omnipresent variability of statistical data. The fit of the model can be evaluated by analysing the deviations of the data from the model (Garfield & Ben-Zvi, 2008). Thus, the process of statistical modelling can be perceived as a cycle which connects reality and data analysis with familiar measures and concepts (Eichler & Vogel, 2013, p. 132).

For statistical modelling, translating the real model (data) into a statistical model is the key activity as it includes selecting the aspects that are modelled (Eichler & Vogel, 2013). The relations within the statistical domain should mirror those within the problem situation as adequately as possible. Thus, the statistical concepts act as stand-ins for the real world, with the relation between the two domains being purposefully constructed to describe, explain or predict a specific real situation (cf. Lesh & Lehrer, 2003). While the variety of suitable mathematical models can be narrow (e.g. for word problems, Greer, 1993) or broad (e.g. open problems, Kaiser, 2005), it is in any case necessary that some mathematical or statistical concepts are already available to the students, which must be used for the mathematising process. Thus, according to the translation perspective on modelling, while existing concepts can be deepened and elaborated, new concepts cannot be acquired by modelling tasks only.

2.2. MODELLING AS ORGANISING: EMERGENT MODELLING

In contrast to the translation perspective, Gravemeijer (1999, 2007) described the acquisition of new mathematical concepts as the main purpose of modelling activities. Building on Freudenthal (1991), he criticised the division into a real world and a mathematical domain which is the basis of the 'modelling as translating' perspective. For Freudenthal (1991), this distinction was dissolved, and in fact, mathematics can become as real as any 'traditional' real world. He portrayed mathematics as the result of reorganising a situation in terms of mathematical relations. Developing new mathematical concepts can be understood as expanding one's mathematical reality. Thus, in the emergent perspective, modelling is understood as 'a process of mathematisation by which the situation is being structured in terms of mathematical relationships' (Gravemeijer, 2002, p. 2). Situation and model evolve at the same time and are mutually constituted (Gravemeijer, 2002). In this way, the learners develop a 'common sense of higher order' (Freudenthal, 1991, p. 8) for reasoning about new phenomena within their reality, and formal mathematics emerges organically from students' activity (with guidance of a teacher, described as 'guided reinvention' by Freudenthal). In this perspective, the term 'model' has to be understood in a broader sense as an overarching idea (e.g. as for Gravemeijer, 2002, a graphical representation of the shape of a distribution). The term 'emergent' refers to two different but interrelated processes: the emergence of a model for mathematical reasoning at the same time as the emergence of more formal mathematical knowledge (Gravemeijer, 1999).

In learning trajectories perceived from the emergent modelling perspective, the role of a model changes during learning processes. In the beginning, emerging models are context-specific *models of* a situation, and derive their meaning from the activity in the task setting. For example, in arithmetic education in primary school, an empty number line without any marks for specific numbers can be used to solve a specific subtraction problem. It is a model of a specific process (Gravemeijer, 1999). Through extensive activity in the task setting, learners eventually start to generalise from the differences between two numbers so that a conceptual shift occurs: numbers are no longer only characterised by counting a set of objects, but by their relations to each other, which can be shown on the empty number line. The model does not change, but its meaning does. No longer signifying a specific situation, it has become a *model for* more formal ways of reasoning. The learning process leading to the shift from model-of to model-for manifests itself in a sequence of intermediate sub-models, which should be anticipated when designing specific learning trajectories (cf. Gravemeijer, 2002; Gravemeijer, 2007; Gravemeijer & Bakker, 2006).

2.3. EMERGENT MODELLING IN STATISTICS EDUCATION

In statistics education, the metaphor of interrogating data and producing evidence has a long tradition. Tukey (1977) called for investigators to act as 'data detectives' who should try to uncover what the data 'seem to say' instead of only routinely conducting standardised procedures. Wild and Pfannkuch (1999) continued and expanded this thought by introducing the interrogative cycle of data analysis. Makar and Rubin (2009) emphasised the importance of evidence-based inferences about phenomena beyond the data in statistical investigation.

Makar and Rubin (2009) also used the metaphor of 'tool use' to explain the role of concepts:

> Averages, distributions, variation, samples, modal clumps – these can be studied as objects in themselves, or as tools for understanding processes or group characteristics. It is vital that the focus in using statistical tools is embedded in the reason that we do statistics – to understand underlying phenomena (p. 84).

This view on concepts as tools can also be understood under a modelling perspective. Statistical concepts are the tools used in modelling; they create the conceptual system that describes a specific phenomenon. In statistics education research, this results in conceptualising many difficulties in learning statistics as difficulties in connecting 'context and data worlds' (Ben-Zvi & Aridor-Berger, 2016) or 'contextual and statistical spheres' (Dierdorp, Bakker, van Maanen, & Eijkelhof, 2016).

In this way, these approaches implicitly follow the view of modelling as translating (see above), resulting in the same epistemological problem: in order to engage in modelling activities, students have to apply statistical measures, and thus statistical concepts need to be already available to them. Students can 'learn to assess whether a model is more of less adequate against the backdrop of contextual factors' (Gravemeijer, 2002, p. 2), which is undoubtedly an important and highly valuable competence.

However, to put more emphasis on the emergence of concepts needed for modelling, Gravemeijer (2002, 2007) proposed an alternative view in which 'a model is a result of an organizing activity. It is in the process of structuring a problem situation that the model emerges' (Gravemeijer, 2002, p. 2). From this perspective, statistical concepts serve to create *models of* a situation in order to gain insights into the phenomenon and make informal generalizations beyond the data. At the same time, the concepts need not remain situationally specific but can also serve in the creation of *models for* reasoning about many different situations, thus allowing for deeper insights into the statistical concepts themselves.

Gravemeijer (2007) called for careful consideration of the teaching-learning arrangement in order to support processes of emergent modelling in statistics education. Two questions arise for such a task design: what is the new concept the students should construct and what is the overarching model? For data analysis, Gravemeijer (2007) proposed the concept of 'distribution-as-an-entity'. For him, idea of shape was central to this concept and could be acquired through a series of changing sub-models consisting of different types of diagrams.

Summing up, we argue that acquiring new statistical concepts can be conceptualised from the emergent modelling perspective as switching between the roles of statistical concepts:

Models-of a situation are context-specific and act as a stand-in used for talking about one specific phenomenon. The purpose of the model is to make sense of the given situation and to enable informal inferences beyond the given data. Statistical measures are used to structure the situation and highlight certain aspects in the model while neglecting others.

Models-for statistical reasoning are generalised over several situations and allow organising a variety of similar and new contexts accordingly. They involve insights into the properties, scope of applicability, and (dis-) advantages of these models.

2.4. RESEARCH QUESTIONS

The intent of this paper is to provide insight into the processes of students' emergent modelling, i.e., the conceptual shift from model of a situation to model for deeper reasoning. The present study focuses on the process of acquiring the concept of distribution, specifically a frequency distribution and its properties such as shape, centre and spread (cf. Biehler, 2007), which can be described by different measures. As stated above, using the emergent modelling perspective as a design heuristic puts great importance on the elements of the teaching-learning arrangement which elicit conceptual shifts. Therefore, the interplay of design elements and students' learning processes needs to be carefully evaluated.

Gravemeijer (2007) called for development of the concept of distribution-as-entity, facilitated by the idea of shape. While we follow Gravemeijer in this new concept to be developed, we place the idea of measures at the heart of our learning trajectory, since the ability to see measures as numerical descriptions of aggregate features is an important aspect of the concept of distribution (Bakker & Gravemeijer, 2004). In order to study processes of emergent modelling, first design elements have to be identified that can initiate those processes. Thus, this paper addresses the following questions:

1. How do design elements of a teaching-learning arrangement elicit processes of emergent modelling?
2. How do switches between model-of and model-for manifest themselves in learning processes as students use, contrast and evaluate different measures?

As the research questions are of overarching interest in the underlying larger research project, they will partly be addressed by four design principles which have been identified in previous cycles of design research and applied in the presented study. Their effects on the students' learning processes will be illustrated in this paper.

3. RESEARCH DESIGN

3.1. TOPIC-SPECIFIC DIDACTICAL DESIGN RESEARCH AS A FRAMEWORK

Design Research as methodological frame The present study is situated in the methodological framework of topic-specific didactical design research (Prediger et al., 2012; for an example from statistics education cf. Prediger & Schnell, 2014). This approach aims at providing two different, but strongly connected, types of results (Figure 1, right side): empirically grounded local theories on the nature of topic-specific learning processes and learning goals (i.e. what and how to learn), and design principles and concrete teaching-learning arrangements for learning this topic. This framework is an example of design research with a focus on learning processes (Prediger, Gravemeijer, & Confrey, 2015), with a special focus on the careful (re-)specification and (re-)structuring of learning goals and content as well as developing content-specific local theories of teaching and learning (Hußmann & Prediger, 2016).

Research is conducted in iterative cycles of four interconnected working areas (Figure 1). A starting point is (Figure 1, (1)) the identification of relevant aspects and necessary insights of a topic (e.g. for distribution: centre, spread, density, skewness etc.) as well as documenting the inherent relations of these aspects, which are used to structure a hypothetical learning trajectory. In addition to content analyses, empirical investigations of the learners' perspectives are crucial components of this phase. Next, a specific teaching-learning arrangement is (re-)developed (2), which serves as the base for design experiments (3) (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). The emphasis here is not only on the viability of the design, but also on investigating the initiated learning processes in order to contribute to the development of a local theory of teaching and learning (4). These contributions to the local theory inform further cycles of topic-specific didactical design research, enabling a further refinement of the theory and the design with each cycle.

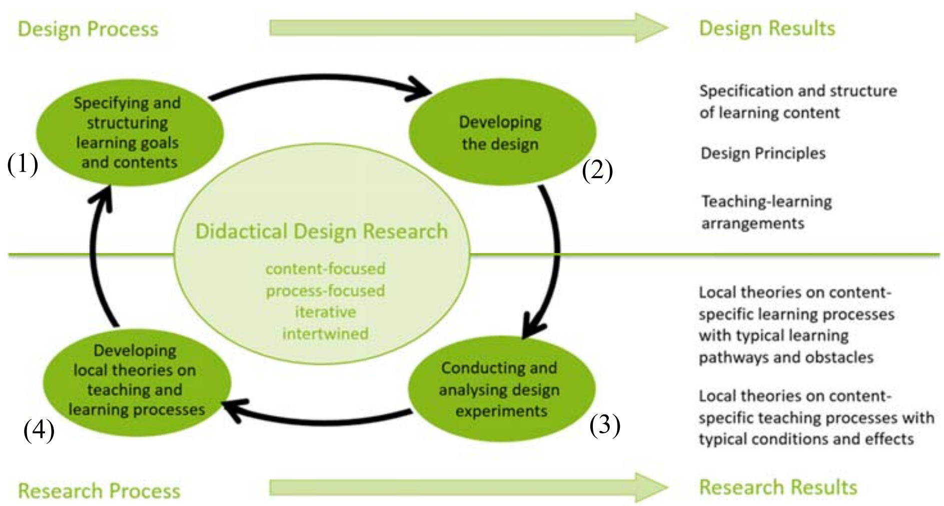

Figure 1. The cycle of topic-specific didactical design research (Prediger et al., 2012; as cited in Prediger & Zwetzschler, 2013, p. 411)

Participants and data collection This study reports on results from the fifth cycle of an ongoing project on the acquisition of statistical concepts (for some results from the first cycle see Schnell & Büscher, 2015). The design experiment series in this cycle took place in a laboratory setting with three pairs of students in a German middle school (aged 13 to 15). Each pair took part in a series of three consecutive design experiment sessions of 45 minutes each. The participating students were evaluated by their teacher as performing well or average in mathematics, which includes statistics education in German curricula. At the time of the experiments, the students had very little experience with statistics, but had learned about simple measures such as the arithmetic mean and median in Grade 6. They were familiar with frequency distributions, but only on a rather superficial level (e.g.

reading information out of frequency distributions without computing measures of centre) and were not used to tasks concerning the comparison of distributions.

All experiments were completely videotaped (altogether 405 minutes of video in the fifth cycle). The students presented here, Kaan and Nesrin, were chosen due to their communicativeness and the richness of their discussion and interpretation of different measures. The related design experiment sessions were fully transcribed.

Data analysis: Concepts-in-action as micro-level access to emergent modelling The video data and transcripts were qualitatively analysed in an category generating approach based on a turn-by-turn analysis. We adapted the theoretical constructs of concepts-in-action and theorems-in-action from Vergnaud's (1996) theory of conceptual fields in order to access students' reasoning on a micro level (see Glade & Prediger, 2017, for a similar approach). 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). Thus, they organise what students focus on and in this case how they interpret certain features of given frequency distributions. Carefully tracing students' concepts-in-action can then allow insights into why students focus on certain aspects (Schnell & Büscher, 2015). Theorems-in-action are defined as 'propositions that [are] held to be true by the individual subject for a certain range of situation variables' (Vergnaud, 1996, p. 225). They are intricately connected to the learners' concepts-in-action: theorems-in-action give meaning to concepts-in-action, which in turn give content to the theorems-in-action. Therefore, the aspects of the situation modelled in the processes of emergent modelling are reflected in the students' concepts-in-action; the statistical relations created or discovered by the students are reflected in their theorems-in-action.

As the goal was to identify the learners' individual processes of emergent modelling, it was necessary to not only compare the learners' process to an intended learning trajectory, but to opt for an open data-led approach to adequately reflect the heterogeneity in these individual processes. We thus identified concepts- and theorems-in-action by systematically comparing and contrasting different cases of learning processes from the fifth cycle of the larger research project (n = 6). Consequently, the identified concepts- and theorems-in-action reflect the individual and informal nature of learners' thinking and therefore do not necessarily correspond to formal statistical concepts. In the analysis, the identified concepts-in-action are symbolised by ||…||, theorems-in-action by <…>.

3.2. DESIGN PRINCIPLES

The overarching concept students were supposed to acquire was that of a distribution with its specific characteristics and attributes (similar to Bakker, 2001; Gravemeijer, 2002). Over the course of the preceding design research cycles, four design principles were identified to support this learning goal:

Focusing on measures To identify the processes of emergent modelling and students' acquisition of the concept of (frequency) distribution, we focus on their use of statistical measures. By this term, we refer to all conceptual tools which can be used to describe observations (e.g. mean, mode, standard deviation in formal descriptive statistics) or to draw inferences from observations (e.g., interval estimations or hypothesis testing in formal inferential statistics). This includes regular formal measures used by experts as the ones mentioned above in parentheses. Learners, on the other hand, invent their own ways of describing and interpreting data, which we call informal measures (Büscher, 2016). Our teaching-learning arrangement thus aims at eliciting these informal measures and encourages students to compare and gradually develop them into regular measures. In line with the emergent modelling approach, we refer to these measures as ‘models’.

Scaffolding the use of measures in argumentation Previous cycles of our project showed how students did use informal and occasionally even regular measures when comparing distributions. While this situationally showed promising starting points for concept development, their appropriate use was elusive. Students sometimes lacked the language to specify what was measured by their informal measures and how they inferred conclusions from them. On other occasions, they simply had forgotten their train of thought when prompted by the researcher or other students. This indicated the need to scaffold their use of measures. To facilitate concept development, we specifically decided to not only ask students to invent their own measures, but also presented them given informal measures meant to contrast with their own modelling of the situation. This design heuristic was implemented by the experimental use of so-called report sheets (see below).

Contrasting models The insight that different measures of the same distribution can result in different views of the situation by emphasising different aspects was central to the measure-focused approach of this study. Thus, contrasting and evaluating different models is one of the main activities, which can be done focussing on different aspects (cf. Lesh & Lehrer, 2003) including (a) their usefulness regarding specific investigations, (b) their correspondence to learners’ experienced reality, (c) their applicability in different situations, or (d) their advantages or disadvantages in argumentation.

Modelling and communicating accessible real-life phenomena In line with the emergent modelling perspective, the context plays a crucial role in developing students’ shifts between models-of and models-for. Thus, tasks should be situated in the context of real-life phenomena. However, simply being situated in a real-life situation is not enough for students to be able to meaningfully engage with the context: they also need to be accessible to students, with the possibility of connecting to an existing pre-understanding of the context. Specifically, this means offering a situation in which the emergence of a statistical concept can be seen as an answer to learners’ questions generated through activity in the task setting. In this case, we decided to emphasise the role of adequately communicating and summarising the phenomena, which can be done by applying statistical measures accordingly.

3.3. TASK DESIGN

Whereas the design experiments consisted of a series of three sessions of 45 minutes each, this paper focuses on the second session. Throughout the design experiments we used the context of weather and climate in order to connect to students’ everyday experiences as well as to their possibly existing knowledge of the phenomenon of global warming. We briefly outline the first session before detailing the design of the second session.

The design of the first session revolved around the analysis and prediction of Antarctic weather. The context was to explore for which range of temperatures a research team should prepare. The task drew on students’ experiences with the natural variability of weather regarding uncertainty in the short term (one never really knows how a single day will be…), yet stable patterns or tendencies in the long term (… yet there is a range of ‘typical’ temperatures every month).

The second session centred on the ‘arctic sea ice’-task and was structured in five phases.

Phase 1 The aim of the first phase was to ensure a shared understanding of the context, data, and diagrams (Figure 2).

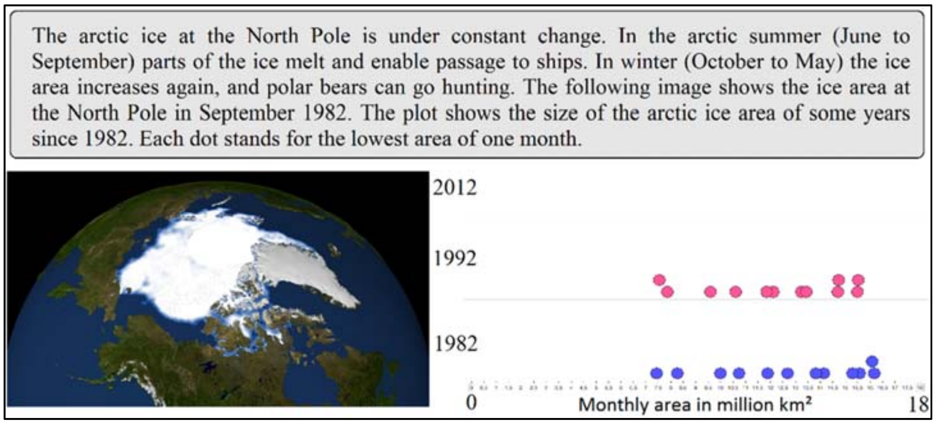

Figure 2. Phase 1 of the arctic sea ice task

The data underlying Figure 2 were taken from the National Snow and Ice Data Center (Fetterer, Knowles, Meier, & Savoie, 2002), with minor modifications for educational purposes (e.g. a small number of data points were added when missing to prevent confusion arising from differences in distributions that were not the focus of this design experiment). In this representation, each dot stands for the minimal area of ice coverage for one month, but no information is given about the exact month represented. This type of diagram was chosen to provide a means for investigating overlaps and differences in the distributions of monthly ice coverage presented in Figure 2. Students were first asked to discuss the meaning of the dots in the given context ('what does each dot mean?', 'why does the area change at all during a year?' etc.) and to verbally summarise the situation. The dot plots were presented to students as static images on a tablet computer with screen overlay software to allow students to freely draw on the screen as additional means to visualise their reasoning (e.g. highlighting certain dots or intervals, cf. Schnell & Büscher, 2015).

Phase 2 In this phase, the students were given filled-in *report sheets* (Figure 3), a design element the students were already familiar with from the first session.

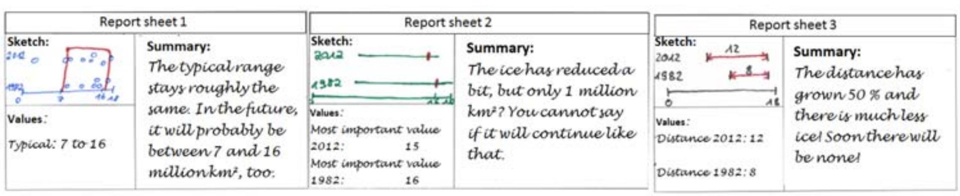

Figure 3. Report sheets in phase 2

To realise the aforementioned design principles, report sheets were the central element in this cycle of design experiments. Their primary intention was to engage students in discussing different (more or less appropriate) models and contrasting them with their own descriptions of the situation. The report sheets (Figure 3) focus on different properties of the distributions by using a variety of informal measures. Each report sheet consisted of a sketch, some (mostly informal) measures, and a statement about future developments of the arctic sea ice area. Wordings and drawings were mostly inspired by written or verbal products of students from previous design cycles (for the informal measure 'typical' see

Büscher, 2016). While the informal measures used vocabulary from students' everyday life (such as 'distance' in report sheet 3), an exact definition was deliberately not given to them. Thus, students were encouraged to find their own interpretations and develop a need for using more reliable and precise descriptions and inferences.

The filled-in report sheets were introduced to the students as reports from other students, giving a report on the change in arctic sea ice between 1982 and 2012. The students were asked to hypothesise about the state of arctic sea ice in 2012. Since they did not yet have access to the data of 2012, they had to base their predictions on their interpretations of the filled-in report sheets and the data of 1982 and 1992. These report sheets however were carefully constructed to give conflicting views on the situation by using different measures and thus highlighting different aspects. The task progressed to the next phase when the students had detected these differing views on the situation and had hypothesised about reasons for these differences.

Phase 3 In this phase, the students received the missing data with an additional illustration of ice area coverage in the North Pole in September 2012 (Figure 4). Looking at the data, one can now see how the fictitious students came to their different conclusions: Focusing on intervals containing most of the data dots would not yield much difference between the distributions, while the range increased dramatically between 1992 and 2012. By this specific construction of the data, the students were faced with the relation between the use of a measure and the possible conclusions to be drawn when accepting this measure as a valid description of the phenomenon being explored.

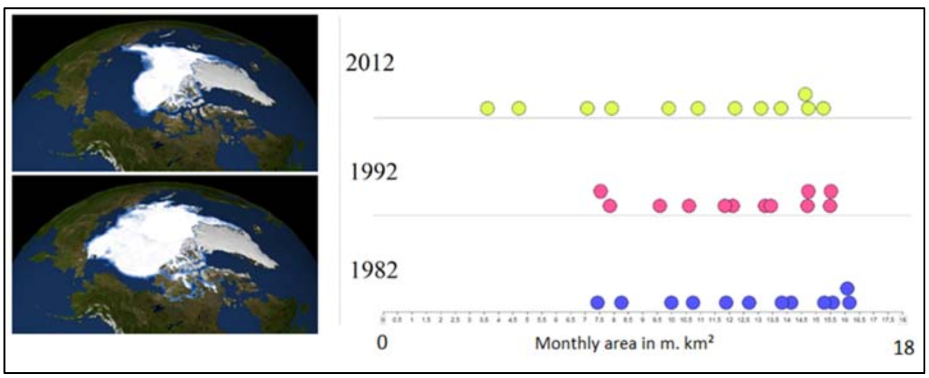

Figure 4: Third phase of the arctic sea ice task

The discussion then revolved around judging whether the filled-in report sheets really captured the phenomenon of arctic sea ice decline, and what they captured better or worse than others.

Phase 4 After having judged these report sheets, the students received an empty report sheet and were instructed to create their own report. Since this activity followed the investigation of already filled-in report sheets, the students were influenced by these report sheets in the creation of their 'own' report. Whereas this does not give insights into the students' initial summary of the situation, it allows deeper insights into how they interpreted, contrasted and evaluated these given (and possibly their own) measures. As this was a scaffolding element, we expected the students to use elements of the filled-in report sheets that captured their ideas expressed in phase 1 and adapted in phase 2 and 3. Afterwards, the students were asked to give reasons why they adapted certain elements and why their report sheet worked better than the others.

Phase 5 In the last phase, the students were asked what should generally be used and considered when creating report sheets.

Summing up, the design aimed at a learning trajectory of acquiring the concept of distribution by focussing on informal statistical measures as central ideas for emergent modelling. The learning processes were initiated by contrasting different models given as filled-in report sheets, which act as scaffolds for measure-based argumentation. The students then created their own report sheets, providing a basis for further critique of the measures' uses and scope of applicability in a more general way.

4. EMPIRICAL RESULTS: THE CASE OF KAAN AND NESRIN

The following excerpts originate from the second session of the design experiment with the students Kaan and Nesrin (both 14 years old), and the first author as researcher (R). Overall, the scene lasts about 30 minutes, during which the students work on the 'arctic sea ice'-task.

In this case, the students gradually differentiated their view on the phenomenon of arctic sea ice decline. At the same time, their statistical knowledge grew as they gradually showed a more sophisticated use of measures. We describe this development by identifying their concepts-in-action, which expand into addressing more specialised aspects of the situation. The informal measures presented on the report sheets (Figure 3) played a key role, since they encouraged students to discuss their meaning and negotiate the representativeness and adequateness of statistical measures. As we will illustrate, this change in the roles of the informal measures can be interpreted as a shift from the use of measures as models-of to models-for.

The first exchange began right after the transition to phase 2. Here, given the data of the area of arctic sea ice per month for 1982 and 1992, students were presented with the filled-in report sheets, which included summaries of the data from 2012. After some time to make sense of the sheets, the students were asked to make an inference about the development of the arctic sea ice area in 2012 (Figure 3).

1	R	What do you think was 2012 like? You have these three report sheets lying there.
2	Kaan	Well, probably in 2012 the ice declined more, and then- That is written in these two *[reports 2 and 3]*, here *(points to report sheet 1)* they reckon that the typical area stays the same, but for example in the one written in black *[report sheet 3]* they reckon the ice area reduced by 50% and maybe there will soon be no more ice.
3	R	Mhm
4	Nesrin	And here *[report sheet 3]* you can tell that the distance, let's say from 18 on, is closer to 2012 than from 8 on, because it's like global warming and because of that, the ice melts more, and you see that here somehow. I think.

Asked to evaluate the statements concerning 2012, Kaan gave preference to the interpretation of declining sea ice by referring to report sheets 2 and 3 (line 2). Nesrin (line 4) introduced her context knowledge about global warming: She referred to the given term 'distance' in report sheet 3 (intended as an informal way for measuring the range between minimum and maximum). While her statement on the distance being 'closer' was unclear, it became apparent that she used the sketch to argue that the change in distance was a result of global warming. Analysing this scene in terms of the students' identifiable concepts-in-action, it seems that the students drew on their context knowledge ||*arctic sea ice decline*|| as the central aspect, which guided them in selecting and interpreting the given representations accordingly. Nesrin specifically also used the (more statistically grounded) concept-in-action ||*informal measure: distance*|| as a means to make sense of the situation on this contextual background.

It is important to note, however, that Nesrin treated the phenomenon of global warming as a given. Rather than using the distance as a justification for the claim of declining ice, she seemed to use her context knowledge of the underlying phenomenon to give meaning to this informal measure. This can be identified as the theorem-in-action <*because arctic sea ice is declining, so does the distance*>, which links both mentioned concepts-in-action.

A few minutes later, the students were again asked to make an estimate for 2012, which led them to compare the record sheets in more detail.

21	Kaan	Well here *[report sheet 3]* they already say 50%, so it's almost half of-
22	Nesrin	So, I don't understand how they say that it declined, but here *[report sheet 1]* they say that it roughly stayed the same.
23	Kaan	Maybe they didn't measure it correctly or something.
24	Nesrin	Yeah. I mean they *[report sheet 3]* say that it increased by half. But they *[report sheet 1]* say that it stayed the same, roughly the same.

This was the first moment in which models were contrasted in their correspondence to the learners' experienced reality. Record sheets 1 and 3 presented different conclusions about the change in the area of ice in 2012 by using different informal measures. Especially for Nesrin, this seemed to be surprising. Here, she focussed solely on the statements about the ice area rather than investigating the measures used and presented in the 'sketch' and 'value' boxes. The root of this conflict was possibly in the understanding of the informal measures as *models of* the situation: Since the students took the ||*arctic sea ice decline*|| to be a given, different models of the same situation could not make different claims. A similar problem was noticed when looking at the conclusions of report sheets 2 and 3, even though they were closer in the overall direction of their summary by stating a decline but drawing different inferences from it. This conflict was momentarily pushed aside by Kaan's explanation that the application of the measures must have been erroneous (line 23).

However, it was addressed again a few minutes later when the students received the missing data from 2012 at the beginning of phase 3. Asked to evaluate the report sheets, Nesrin started having a closer look at the measures used.

41	R	(…) well what do you think now, are *[the three report sheets]* correct, did they do well or not?
42	Nesrin	Well it's okay actually. If you – with that *[report sheet 2]* you have to think a lot, and in the beginning you can't understand immediately what they mean. And this *['most important value' in 2012]* 15 *[report sheet 2]*, it's like – at the end, ehm, they could rather say what happens in the other months (…)
43	R	What do you mean?
44	Nesrin	(…) if you look here they *[report sheet 3]* say it declined by half. And here *[report sheet 2]* they don't say anything about declining or rising, they only say 'in the end the ice is this big'.
45	R	Mhm.
46	Nesrin	And they rather could have said what happened in the middle. (…) They could have better said how much the ice declined. Not only 'in the end it's this big'.

Nesrin seemed to be able to resolve the conflict of the contradicting report sheets now: They both depicted the same situation, but with different foci; report sheet 2 used the informal measure 'most important value' (as an informal way of expressing roughly the value around which most dots spread). She seemed to develop the interpretation that <*the* *‘most important value’ represents the final state*> (‘in the end’, line 44). As she used the term ‘end’, she probably referred to measuring a static central tendency rather than a change (‘what happened in the middle’ and ‘how much the ice declined’, line 46). As she seemed to perceive the latter as more important in this task, she deemed report sheet 2 to be ‘okay’ (line 42), but improvable or inadequate.

Overall, this scene can be interpreted as an indication of a shift of the role of the informal measures: They no longer seemed to function as sole models of the situation as a whole, but rather as models of certain aspects of the situation, which are both correct, but useful to different degrees. Specifically, Nesrin might have been beginning to see that the phenomenon of ||*arctic sea ice decline*|| consist of something like a ||*final state*|| and the ||*degree of departure*|| during the ice melt. Here she was able to show a more pronounced view on the situation than before.

Towards the end of the design experiment, the students were asked to reflect on the informal measures applied, and to give more general reasons when one measure should be preferred over the other. For this, they focussed on single elements of the report sheets and discussed them in detail.

81	Kaan	Well here *[report sheet 3, sketch]* the box is also depicted with the line and you show the exact distance, how much it has departed. And here *[report sheet 1]* you show what’s outside of it.
82	R	Mhm
83	Kaan	Well, I think here *[report sheet 3]* you can tell much better how it differs, from 1982 and 2012 and how the difference is.

Kaan in line 81 compared the sketches on the report sheets by focussing on ‘the line’ in report sheet 3 (informal measure ‘distance’ describing the range) versus ‘the box’ in record sheet 1 (describing the interval in which the values of 1992 and 1982 lie). By ‘what’s outside of [the box]’, he probably meant the dots with the lowest value left of the marked box in 2012. Interpreting his statement, he seemed to evaluate the ‘distance’ as a more exact representation of the ||*overall arctic sea ice decline*|| while the ‘box’ put more emphasis on singular values in the sense of outliers from a ‘typical’ interval as in ||*singular deviations as indication for arctic sea ice decline*||. He generally preferred the numerical, more exact representation of the range (line 83). Nesrin differentiated in her preference of when to use which measure:

84	Nesrin	I also think that if you did this with temperatures, that it wouldn’t fit that well, in contrast to *[areas]* here– because temperatures are, like, you would want to look more at what is outside of the red box *[as in report sheet 1]*, which is enough. And with sizes it’s not like this, here, this *[report sheet 3]* fits better. Because with the lines it’s like – that it just depicts it better, how big the distance between them is.

Nesrin drew on her experience from the previous design experiment session in which the students worked on distributions of daily Antarctic temperatures to give an estimation of which equipment to bring to prepare for a research stay. She stated that in the temperature context and its question, she would be more interested in values which were outside of the typical interval represented by the box; the reasoning behind this could be that too many outliers of very cold temperature would have severe consequences for the researchers’ health when being ill prepared (this was discussed to some length during the first experiment). For the sea ice area context, though, she also preferred the range—although her reasoning ‘because it just depicts it better, how big the distance between them is' seemed tautological. She drew on the situation to evaluate the ||*adequateness*|| of the applied measures.

Overall, the students discussed three different uses for measures: quantifying the range within a distribution (how much it 'departs' and how big the 'distance' is, 81/84), comparing the range of different distributions (the difference between 1982 and 2012, 83), and judging single data points as outliers (inside and outside the 'typical' box, 81).

Therefore, Nesrin's and Kaan's view of measures became much more refined: The informal measures were no longer used only as representations of specific situations. Instead they acted as tools for highlighting different aspects of situations, which depended on the question to be answered. Different types of situations called for different measures, which could more or less adequately highlight certain aspects. Thus, measures were put in relation to situation and question and became models not only of a situation but also for reasoning about their properties, applicability, and adequateness.

5. SUMMARY

This case study addressed two research questions in relation to students' process of emergent modelling when dealing with frequency distributions: (1) How do design elements of a teaching-learning arrangement elicit processes of emergent modelling, and (2) how do switches between model-of and model-for manifest themselves as students use, contrast and evaluate different measures?

Regarding research question 1, central to the design of the teaching-learning environment were the report sheets, which provided informal ideas, different views on the phenomenon, and focal points for discussion. Thus, the report sheets reflected the design principles of focusing on (informal) measures, scaffolding argumentation, contrasting models, and modelling and communicating life-world phenomena. All of these design principles played a part in initiating and supporting processes of emergent modelling. Global warming served as a well-known life-world phenomenon that allowed the students to give meaning to the informal measure 'distance'. Consequently, the discussion primarily revolved not around making sense of the phenomenon, but of the informal measures. Conflict was created by report sheets that could not simultaneously serve as the same models of the same situation. This caused the students to differentiate between model and situation, and to use measures as models-for reasoning about situations, where each model had its advantages and disadvantages for certain types of situations.

Regarding research question 2, during the design experiment, the students firstly interpreted the given measures as representing the situation and evaluated them against their context knowledge of global warming as correctly or incorrectly applied. With the ongoing learning process, they refined their assessment by contrasting the different measures in terms of their validity. Not only did their understanding of the measures change, but so did their view of the phenomenon: the general ||*arctic sea ice decline*|| was conceptually separated into a ||*final state*|| and a ||*degree of departure*||. The measures started to become not only models which gave insights into the situation, but were addressed and evaluated in terms of their properties; at the same time, the interpretation of the context became more refined. While in the beginning measures were thought of in terms of ||*representativeness*||, this later changed to terms of ||*adequateness*||. Finally, by contrasting the use of measures for different phenomena, the measures gained criteria for applicability through the establishment of connections to the given situation and problem.

This sheds some light on the role of context in the development of statistical knowledge. Whereas statistical investigation usually means learning about the context through use of statistics, these roles do not seem as clear cut under an emergent modelling perspective. The students' intuitive knowledge of global warming and its likely impact on arctic sea ice provided the foundation on which their models of the situation could be constructed. From then on, statistical and contextual knowledge co-developed through processes of emergent modelling. In this way, an emergent modelling perspective reveals that not only can statistical knowledge inform contextual knowledge, but the opposite may also be the case.

6. DISCUSSION AND LIMITATIONS

Three major themes have influenced the research carried out in this paper: (1) the use of emergent modelling as a design heuristic, guiding the design of the teaching-learning arrangement; (2) the use of emergent modelling as an analytic framework, used to trace students' developing reasoning; and (3) the use of informal measures as starting points for initiating concept development. In the discussion, we turn towards each of these themes in order to reveal broader lessons to be learned for further research.

6.1. EMERGENT MODELLING AS A DESIGN HEURISTIC

Using the emergent modelling perspective to guide the design of a teaching-learning environment proved useful in terms of eliciting students' processes of reasoning with and about (informal) statistical measures. The design principles of focusing on measures, scaffolding the use of measures in argumentation, contrasting models and modelling and communicating life-world phenomena led students to engage in the task while gradually developing their conceptual understanding of the situation as well as understanding of the (function of) measures.

One goal of design research approaches is to provide design principles that can inform the design of other teaching-learning arrangements (Prediger et al., 2012). Since this study is limited to one case of students, the question remains if these principles could prove useful for other students and other tasks. During the previous research cycle, the principle of contrasting models seemed especially promising, which is in line with the results of Ben-Zvi (2003, 2004) and Bakker and Gravemeijer (2004). They asked students to compare groups or distributions; we additionally showed how comparing measures can also be fruitful for facilitating a deeper understanding of statistical concepts.

One important part of designing a teaching-learning arrangement that is grounded in the emergent modelling design heuristic is the choice of context. Using contexts that are meaningful to students is a common requirement for such a design. An emergent modelling perspective reveals demands to be made of such a context: students should be able to structure the context without relying on the statistical concepts that are to emerge. Yet the context also needs to encourage questions that create the need for more formal concepts. As the students were able to draw on their intuitive knowledge of global warming while engaging in processes of emergent modelling, we believe the context used in the arctic sea ice task to be just such a context. It remains the task of further research to improve on the utilisation of that context and to uncover additional contexts supporting emergent modelling.

6.2. EMERGENT MODELLING AS AN ANALYTICAL FRAMEWORK

Emergent modelling was not only used as a design heuristic, but also as a tool for the analysis of learning pathways. Tracking the shifts in models on a micro level, we tried to illustrate a conceptual difficulty: At the beginning, measures function as *models of* a situation for learners which enable access to make sense of the situation. Understanding these measures as describing only certain aspects of the situation requires a shift to a use of measures as *models for* reasoning about their properties. Both, knowledge of the statistical concepts as well as context-focussed inferences about the underlying phenomenon, are developed as students make this conceptual shift. The importance of tracing statistical and contextual knowledge in students' learning pathways is well-established in statistics education research (e.g. Ben-Zvi & Aridor-Berger, 2016; Dierdorp et al., 2016), as a deeper understanding of the context can be seen as the reason for statistics itself (e.g. Makar & Rubin, 2009). This makes the emergent modelling approach especially fruitful for statistics education by providing the model-of and model-for language as a means to trace these developments.

6.3. INFORMAL MEASURES AS STARTING POINTS

In adopting the learners' perspective in making sense of data and drawing inferences by focusing on the use of informal measures, we have eschewed the question of how to get students to draw inferences based on regular measures. One could argue that this case study provides nothing but entertaining starting points, and the real work of statistics education still has to be carried out. While we would agree that it is still a long way for the students to acquire more specific knowledge of distributions and formalised ways of comparing them, we want to emphasise the importance of acknowledging learners' active roles in the construction of meaning during their learning processes. In our point of view, it is imperative to build on the learners' own reasoning under an epistemological perspective (e.g. following Vergnaud, 1996). The focus on learners' concepts-in-action also reveals the seed of regular statistical concepts embedded in their thinking. Using informal measures as starting points for learning processes thus provides a foundation for building meaningful statistical reasoning; regular measures can be supplied when this foundation proves sustainable (cf. Bakker & Gravemeijer, 2004).

6.4. LIMITATIONS

We presented only one case of two students, who proved to be especially verbal in expressing their thoughts in the design experiments. The intention of this choice was to illustrate the research framework of emergent modelling and highlight certain phenomena regarding the development of informal measures by presenting rich and detailed insights of their learning pathway. While all pairs of students invented, used and discussed different informal measures, the results of these activities differed greatly. The findings presented here are currently being contrasted with the design experiments of other students in our ongoing research to identify more general phenomena (among this specific set of students).

Furthermore, all results must be interpreted against the backdrop of the tasks used. As elaborated above, the task design, sequence, and especially the scaffolding of informal measures is a very specific approach. Changing the report sheets, for example, could produce quite different results. We can make no statements on the students' own invented informal measures with the data presented in this paper. The informal measures used on the report sheets, though, were observed in previous design cycles, in which the students freely invented them without scaffolds (cf. Büscher, 2016; Schnell & Büscher, 2015). However, the focus of this paper was to elaborate on the learning processes under the condition of these given supports to explore students' individual evaluation and comparison of different informal measures. We acknowledge, though, the development and use of individual informal measures as crucial for students' conceptual development as shown for instance by Bakker and Gravemeijer (2004).

REFERENCES

Bakker, A. (2001). Symbolizing data into a 'bump'. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 81–88). Utrecht, the Netherlands: Freudenthal Institute.

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147–168). Dordrecht: Springer Netherlands.

Ben-Zvi, D. (2003). The emergence of reasoning about variability in comparing distributions: A case study of two seventh grade students. In C. Lee (Ed.), *Reasoning about variability: A collection of current research studies* (pp. 42–63). Mount Pleasant, Michigan: Central Michigan University.

Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal, 3*(2), 42–63. http://iase-web.org/documents/SERJ/SERJ3(2)_BenZvi.pdf

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representation. *Educational Studies in Mathematics, 45*(1/3), 35–65. doi:10.1023/A:1013809201228

Ben-Zvi, D., & Aridor-Berger, K. (2016). Children's wonder how to wander between data and context. In D. Ben-Zvi & K. Makar (Eds.), *The teaching and learning of statistics: International perspectives* (1st ed., pp. 25–36). Cham, s.l.: Springer International Publishing. doi:10.1007/978-3-319-23470-0_3

Biehler, R. (2007). Students' strategies of comparing distributions in an exploratory data analysis context. In *Bulletin of the International Statistical Institute* (pp. 783–790). Lisboa: Instituto Nacional de Estatistica (INE).

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects - State, trends and issues in mathematics instruction. *Educational Studies in Mathematics*, *22*(1), 37–68.

Büscher, C. (2016, July). *Students' informal measures between objects and tools*. Paper presented at the 13th International Congress on Mathematical Education, Hamburg.

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

Dierdorp, A., Bakker, A., van Maanen, J., & Eijkelhof, H. (2016). Supporting students to develop concepts underlying sampling and to shuttle between contextual and statistical spheres. In D. Ben-Zvi & K. Makar (Eds.), *The teaching and learning of statistics: International perspectives* (1st ed., pp. 37–48). Cham, s.l.: Springer International Publishing. doi:10.1007/978-3-319-23470-0_4

Eichler, A., & Vogel, M. (2013). *Leitidee Daten und Zufall: Von konkreten Beispielen zur Didaktik der Stochastik*. *Studium*. Wiesbaden: Springer Spektrum.

Fetterer, F., Knowles, K., Meier, W., & Savoie, M. (2002, updated daily). *Sea ice index, version 1: Arctic sea ice extent*. NSIDC: National Snow and Ice Data Center. http://dx.doi.org/10.7265/N5QJ7F7W

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht, Boston: Kluwer Academic Publishers.

Garfield, J. B., & Ben-Zvi, D. (2008). Learning to reason about statistical models and modeling. In J. B. Garfield & D. Ben-Zvi (Eds.), *Developing students' statistical* *reasoning: Connecting research and teaching practice* (pp. 143–163). Dordrecht, the Netherlands: Springer

Glade, M., & Prediger, S. (2017). Students' individual schematization pathways - empirical reconstructions for the case of part-of-part determination for fractions. *Educational Studies in Mathematics*, *94*(2), 185–203. doi:10.1007/s10649-016-9716-5

Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning*, *1*(2), 155–177. doi:10.1207/s15327833mtl0102_4

Gravemeijer, K. (2002, July). *Emergent modeling as the basis for an instructional sequence on data analysis*. Paper presented at the sixth International Conference on Teaching Statistics, Cape Town, South Africa. Retrieved from http://iase-web.org/Conference_Proceedings.php?p=ICOTS_6_2002

Gravemeijer, K. (2007, December). *Emergent modeling and iterative processes of design and improvement in mathematics education*. Plenary lecture at the APEC-TSUKUBA International Conference III, Innovation of Classroom Teaching and Learning through Lesson Study - Focusing on Mathematical Communication, Tokyo and Kanazawa, Japan.

Gravemeijer, K., & Bakker, A. (2006, July). *Design research and design heuristics in statistics education*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil. Retrieved from http://iase-web.org/Conference_Proceedings.php?p=ICOTS_7_2006

Greer, B. (1993). The mathematical modeling perspective on wor(l)d problems. *The Journal of Mathematical Behavior*, *12*(3), 239–250.

Hußmann, S., & Prediger, S. (2016). Specifying and structuring mathematical topics. *Journal für Mathematik-Didaktik, 37*(1), 33–67. doi:10.1007/s13138-016-0102-8

Kaiser, G. (2005). Mathematical modelling in school – examples and experiences. In H.-W. Henn (Ed.), *Mathematikunterricht im Spannungsfeld von Evolution und Evaluation. Festschrift für Werner Blum* (pp. 99–108). Hildesheim: Franzbecker.

Lesh, R., & Lehrer, R. (2003). Models and modeling perspectives on the development of students and teachers. *Mathematical Thinking and Learning*, *5*(2-3), 109–129.

Lesh, R. A., & Doerr, H. M. (2003). Foundations of a models and modelling perspective on mathematics teaching, learning, and problem solving. In R. A. Lesh & H. M. Doerr (Eds.), *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 3–34). Mahwah, N.J.: Lawrence Erlbaum Associates.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, *8*(1), 82–105. http://iase-web.org/documents/SERJ/SERJ8(1)_Makar_Rubin.pdf

Pfannkuch, M., Regan, M., Wild, C. J., & Horton, N. J. (2010). Telling data stories: Essential dialogues for comparative reasoning. *Journal of Statistics Education, 18*(1). http://ww2.amstat.org/publications/jse/v18n1/pfannkuch.pdf

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., Link, M., Hinz, R., Hußmann, S., Thiele, J., & Ralle, B. (2012). Lehr-Lernprozesse initiieren und erforschen–fachdidaktische Entwicklungsforschung im Dortmunder Modell. *Mathematischer und Naturwissenschaftlicher Unterricht*, *65*(8), 452–457.

Prediger, S., & Schnell, S. (2014). Investigating the dynamics of stochastic learning processes: A didactical research perspective, its methodological and theoretical framework, illustrated for the case of the short term–long term distinction. In E. J. Chernoff & B. Sriraman (Eds.), *Advances in mathematics education: Probabilistic thinking. Presenting plural perspectives* (pp. 533–558). Dordrecht, New York: Springer.

Prediger, S., & Zwetzschler, L. (2013). Topic-specific design research with a focus on learning processes: The case of understanding algebraic equivalence in grade 8. In T. Plomp & N. Nieveen (Eds.), *Educational design research - part A: An introduction* (pp. 409–423). Enschede, the Netherlands: SLO.

Schnell, S., & Büscher, C. (2015). Individual Concepts of Students Comparing Distribution. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 754–760). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Tukey, J. W. (1977). *Exploratory data analysis. Addison-Wesley series in behavioral science*. Reading, Mass.: Addison-Wesley Pub. Co.

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.

Wild, C. J. (2006). The concept of distribution. *Statistics Education Research Journal, 5*(2), 10–26. http://iase-web.org/documents/SERJ/SERJ5(2)_Wild.pdf

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–248.

CHRISTIAN BÜSCHER
Institut für Entwicklung und Erforschung des Mathematikunterrichts
Vogelpothsweg 87
44227 Dortmund
Germany