Chapter 4
Students' Aggregate Reasoning with Covariation

Keren Aridor and Dani Ben-Zvi

Abstract Helping students interpret and evaluate the relations between two variables is challenging. This chapter examines how students' aggregate reasoning with covariation (ARwC) emerged while they modeled a real phenomenon and drew informal statistical inferences in an inquiry-based learning environment using TinkerPlots™. We focus in this illustrative case study on the emergent ARwC of two fifth-graders (aged 11) involved in statistical data analysis and modelling activities and in growing samples investigations. We elucidate four aspects of the students' articulations of ARwC as they explored the relations between two variables in a small real sample and constructed and improved a model of the predicted relations in the population. We finally discuss implications and limitations of the results. This article contributes to the study of young students' aggregate reasoning and the role of models in developing such reasoning.

Keywords Aggregate reasoning · Exploratory data analysis · Growing samples
Informal statistical inference · Reasoning with covariation · Statistical modelling

4.1 Introduction

The purpose of this chapter is to provide an initial scheme for understanding young students' emergent articulations of aggregate reasoning with covariation (ARwC) in the context of informal statistical inference from growing data samples. Handling data from an aggregate point of view is a core aspect of statistical reasoning (Hancock et al. 1992). Predicting properties of the aggregate is the essential aspect of data analysis and statistical inference. To achieve this goal, one should develop a notion of data as an organizing structure that enables seeing the data as a whole (Bakker and Hoffmann

K. Aridor (✉) · D. Ben-Zvi
Faculty of Education, The University of Haifa, Haifa, Israel
e-mail: kerenaridor@gmail.com

D. Ben-Zvi
e-mail: dbenzvi@univ.haifa.ac.il

G. Burrill and D. Ben-Zvi (eds.), *Topics and Trends in Current Statistics Education Research*, ICME-13 Monographs, https://doi.org/10.1007/978-3-030-03472-6_4

2005). The development of such aggregate view of data is a key challenge in statistics education (Bakker et al. 2004). Previous research about students' reasoning with covariation specified aspects of ARwC that are essential to judging and interpreting relations between two variables, such as viewing data as a whole (Moritz 2004). Nevertheless, young students tend to see data as individual cases (local view) rather than a global entity (Ben-Zvi and Arcavi 2001), and often focus on a single variable and not on the bivariate relationship (Zieffler and Garfield 2009). Statistical modelling contexts can help address these challenges by supporting students' search for patterns in data and accounting for variability in these patterns (Pfannkuch and Wild 2004).

To examine how young students' ARwC can emerge while they make informal statistical inferences and model an authentic phenomenon using hands-on tools and TinkerPlots™ (Konold and Miller 2011), we first elaborate on the notions of informal statistical inference, covariational and aggregate reasoning. We then highlight reasoning with modelling and the "growing samples" pedagogy. In the Method section, we describe the Connections project and the fifth grade-learning trajectory to put the tasks of the project in context. Next, we present the main results of this research by specifying the four aspects that structure the fifth grade students' ARwC. We conclude with theoretical and pedagogical implications and limitations of the research.

4.2 Literature Review

4.2.1 Informal Statistical Inference (ISI)

Statistical inference moves beyond the data in hand to draw conclusions about some wider universe, taking into account uncertainty in these conclusions and the omnipresence of variability (Moore 2004). Informal Statistical Inference (ISI) is a theoretical and pedagogical approach for developing statistical reasoning, connecting between key statistical ideas and informal aspects of learning statistical inference (Garfield and Ben-Zvi 2008). ISI is based on generalizing beyond the given data, expressing uncertainty with a probabilistic language, and using data as evidence for these generalizations (Makar and Rubin 2009, 2017). The reasoning process leading to making ISIs is termed Informal Inferential Reasoning (IIR) . IIR refers to the cognitive activities involved in informally formulating generalizations (e.g., conclusions, predictions) about "some wider universe" from random samples, using various statistical tools, while considering and articulating evidence and uncertainty (Makar et al. 2011). IIR includes reasoning with several key statistical ideas such as: sample size, sampling variability, controlling for bias, uncertainty and properties of data aggregates (Rubin et al. 2006).

4.2.2 Reasoning with Covariation

This chapter is focused on statistical reasoning with covariation. Statistical covariation relates to the correspondence of variation of two variables that vary along numerical scales (Moritz 2004). Bivariate relations are characterized by the variability of each of the variables; the pattern of a relation, the shape of the relationship in terms of linearity, clusters and outliers; and the existence, direction and strength of a trend (Watkins et al. 2004). Reasoning with covariation is defined as the cognitive activities involved in coordinating, explaining and generalizing two varying quantities while attending to the ways in which they change in relation to each other (Carlson et al. 2002).

Reasoning with covariation plays an important role in scientific reasoning and is applied depending on usage, goals and discipline (Schauble 1996). For example, covariation can serve as an alternative to the concept of function. A covariation approach in this context entails being able to move between values of one variable and coordinating this shift with movement between corresponding values of another variable. Such an approach plays an important role in students' understanding, representing and interpreting of the rate of change, and its properties in graphs (Carlson et al. 2002). The approach can also lead to reasoning about the algebraic representation of a function (Confrey and Smith 1994).

Moritz (2004) identified four levels of verbal and numerical graph interpretations while analysing bivariate associations: Nonstatistical, single aspect, inadequate covariation and appropriate covariation. Nonstatistical responses relate to the context or to a few data points, such as outliers or extreme values, without addressing covariation. Single aspects responses refer to a single data point or to one of the variables (usually the dependent), with no interpolating. Inadequate Covariation responses address both variables but either relate to correspondence by comparing two or more points without generalizing to the whole data or to the population; or, variables are described without relating to the correspondence or by mentioning it incorrectly. Appropriate covariation responses refer to both variables and their correspondence correctly.

Moritz's hierarchy as well as other studies reflect students' challenges while reasoning with covariation. Students tend to focus on isolated data points rather than on the global data set and trend; focus on a single variable rather than the bivariate data; expect a perfect correspondence between variables, without exception in data (a deterministic approach); consider a relation between variables only if it is positive (the unidirectional misconception) ; reject negative covariations when they are contradictory to their prior beliefs; have a hard time distinguishing between arbitrary and structural covariation (Batanero et al. 1997; Ben-Zvi and Arcavi 2001; Moritz 2004).

Several studies suggest that a meaningful context for reasoning with aggregate aspects of distribution, such as shape and variability, can support developing reasoning with covariation (e.g., Cobb et al. 2003; Konold 2002; Moritz 2004; Zieffler and Garfield 2009).

4.2.3 Aggregate Reasoning

Developing statistical reasoning involves flexibly shifting between a local view of data and a global view of data according to the need and the purpose of the investigation (Ben-Zvi and Arcavi 2001; Konold et al. 2015). Aggregate reasoning is a global view of data that attends to aggregate features of data sets and their propensities (Ben-Zvi and Arcavi 2001; Shaughnessy 2007). When viewing data as an aggregate, a data set is considered as an entity with emergent properties, which are different from the properties of the individual cases themselves (Friel 2007). Two important aggregate properties are the distinction between signal and noise and the recognition and diagnosis of various types and sources of variability (Rubin et al. 2006).

Aggregate reasoning is discussed in the literature mostly in the context of data and distribution. The notion of distribution as an organizing conceptual structure is conceived by aggregate aspects of distribution, such as the general shape, how spread out the cases are, and where the cases tend to be concentrated within the distribution (Bakker and Gravemeijer 2004; Konold et al. 2015). Reasoning with bivariate data is mostly discussed without using the terminology of aggregate reasoning. For example, Ben-Zvi and Arcavi (2001) describe the way students' previous knowledge and different types of local observations supported and hindered the development of their global view of data. In the beginning, they reasoned with the investigated association as an algebraic pattern, with relation to local data cases and adjacent difference. However, this focus on pointwise observation eventually supported the development of the students' reasoning with the notion of trend while relating to the data as a whole.

Konold (2002) recognizes the gap between people's ability to make reasonable judgments about relations in the real world and their struggle to make judgments about covariation from representations such as scatterplots and two-by-two contingency tables. Konold suggests that this struggle stems from a difficulty to decode the ways in which these relationships were displayed (Cobb et al. 2003; Konold 2002). One goal of the current study is to extend the understanding of aggregate reasoning to the context of statistical modelling and covariation, which we term ARwC. The analysis of ARwC will consider various aspects of students' aggregate reasoning including reasoning with variability.

4.2.4 Reasoning with Variability

Variability is the aptness or tendency of something to vary or change (Reading and Shaughnessy 2004). Variability is omnipresent in data, samples and distributions (Moore 2004). While reasoning with data, students should search for signals in the variability, as well as for potential sources of such variability (Shaughnessy 2007). A signal can be considered as the patterns which have not been discounted as ephemeral. Such patterns can become evident only in the aggregate. Noise can be considered as the unexplained variability around these patterns, if identified (Wild and Pfannkuch 1999).

Reasoning with variability has both informal and formal aspects, from understanding that data vary, to understanding and interpreting formal measures of variability. Students seem to reason intuitively with informal aspects of variability, such as the representativeness of variability by spread and the idea that data vary. However, students tend to focus primarily on outliers and have difficulties measuring variability in a way that depicts thinking of variability as representing spread around the center (Garfield et al. 2007).

A conceptual understanding of variability includes: (a) developing intuitive ideas about variability (e.g., repeated measurement on the same characteristic are variable); (b) the ability to describe and represent variability (e.g., the role of different representation of a data set in revealing different aspects of variability, the representativeness of spread measurements); (c) using variability to make comparisons; (d) recognizing variability in special types of distributions (e.g., the role of the variability of both variables' distributions to a bivariate data distribution) ; (e) identifying patterns of variability in fitting models; (f) using variability to predict random samples or outcomes; and (g) considering variability as part of statistical thinking (Garfield and Ben-Zvi 2005). Modelling a phenomenon entails the search for differences and similarities in the population, which is an initial step toward reasoning with variability (Lehrer and Schauble 2012).

4.2.5 Statistical Models and Modelling

Freudenthal (1991) viewed mathematics as a human activity. As such, students should learn mathematics by "mathematising": they should find their own levels of mathematics and explore the paths leading there with as much guidance as they need. The process of "mathematising" lasts as long as reality is changing and extending under various influences, including the influence of mathematics, which is absorbed by this changing reality. One component of "mathematising" is modelling. As such, modelling is defined as simplifying or grasping the essentials of a static or dynamic situation within a rich and dynamic context (Freudenthal 1991). Modelling can be perceived as interrelating processes in which the role of the model is changing as thinking progresses. At the first process, a model emerges as a "model of" informal reasoning and develops into a "model for" more formal reasoning. At the second process, a new view of a concept emerges along the transition from "model of" to "model for". Such view can be perceived as formal in relation to the initial disposition toward this concept. These two processes are accompanied by a third one—the shaping of a model as a series of signs that specifies the previous reasoning process (Gravemeijer 1999).

Models and modelling are essential components of statistical reasoning and thinking (Wild and Pfannkuch 1999). The practice of statistics can be considered as a form of modelling, as the development of models of data, variability and chance are paving the way for a statistical investigation (Lehrer and English 2017). The modelling process entails an evaluation and improvement of models to include new theoretical ideas or data based findings (Dvir and Ben-Zvi 2018; Lesh et al. 2002). A statistical model is an analogy that simplifies a real phenomenon, describes some of the connections and relations among its components, and attends to uncertainty (Wild and Pfannkuch 1999). Based on the recognition that aggregate reasoning requires summarizing and representing data in multiple ways depending on the nature of the data, various pedagogical approaches have been developed. With this in mind, a modelling pedagogical approach can support the emergence of aggregate views of data (Lehrer and Schauble 2004; Pfannkuch and Wild 2004). In this chapter, we focus on students' emergent ARwC in relation to statistical models that were developed by them to describe a real phenomenon and predict outcomes for an unknown population. These models were constructed as part of the Connections learning environment, which was built on the growing samples and the purpose and utility ideas.

4.2.6 Task Design

The *growing samples* educational approach is an instructional idea mentioned by Konold and Pollatsek (2002), worked out by Bakker (2004) and elaborated by others (e.g., Ben-Zvi et al. 2012). In this approach, students are introduced to increasing sample sizes that are taken from the same population. For each sample, they pose a research question, organize and interpret the data, and draw ISIs. Later, they face "what if" questions that encourage them to make conjectures about same sized samples, or about a larger sample. In this approach, students are required to search for and reason with aggregate features of distributions and to identify signals out of noise. They need to compare their conjectures about the larger samples with insights from the data, to account for the limitations of their inferences and to confront uncertainty with regard to their inferences. The growing samples approach can be a useful pedagogical tool to support coherent reasoning with key statistical ideas (Bakker 2004; Ben-Zvi et al. 2012).

Another task design approach used in the Connections project was *purpose and utility*. Although everyday contexts can support learning statistics or mathematics, the strength of meaningful learning environments is in the design for purpose and utility (Ben-Zvi et al. 2018). The term purpose refers to students' perceptions. A purposeful task is a task that has a meaningful outcome (a product or a solution) for students. Such a purpose might be different from the teacher's intentions. The utility of ideas means that the learning process involves construction of meaning for the ways in which these ideas are useful. Purpose and utility are strongly connected. Purposeful tasks, encompass opportunities for students to learn to use an idea in ways that allow them to reason with its utility, by applying it in that purposeful context (Ainley et al. 2006). With this literature review in mind, we now formulate the research question of this study.

4.3 Research Question

In this case study, we focus on two fifth grade students (age 11) who were involved in modelling activities of bivariate data and drawing ISIs in growing samples investigations. In this context, we ask: *What can be the characteristics of the students' emergent ARwC?*

4.4 Method

4.4.1 The Setting

To address this question we draw on data from the 2015 Connections Project in a fifth grade Israeli classroom. In this project, a group of researchers and teachers designed and studied an inquiry-based learning environment to develop statistical reasoning. The focus of the 2015 Project was aggregate reasoning using modelling activities in the context of making ISIs. The design of the learning trajectory was guided by three main approaches: growing samples, statistical modelling, and purpose and utility. Students investigated samples of increasing size that were drawn from the same population. The goal of each investigation was to model an authentic phenomenon within the target population—all fifth graders in Israel. To do that, the students reasoned with the meaning and utility of statistical concepts, such as data, center, variability and distributions, using hands-on tools (pen and paper) or TinkerPlots™ (Konold and Miller 2011). TinkerPlots™ is an innovative data analysis, visualization and modelling tool designed to support students' (grades four to nine) reasoning with data.[1] TinkerPlots™ provides a dynamic graph construction tool that allows students to invent their own elementary graphs and evaluate them (Biehler et al. 2013). Models were constructed using the TinkerPlots™ Sampler, which can be used to model probabilistic processes and to generate random data from a model.

The students participated in ten activities (28 lessons, 45 min each) organized in two main cycles of data investigations of samples (see Table 4.1): (1) their whole class and grade (2–3 attributes, samples of 25 and 73 cases); and (2) their grade (18 attributes, samples of 10, 24 and 62 cases). For each sample, students posed a research question, organized and interpreted the data, and drew ISIs. They made conjectures about a larger sample to confront uncertainty. They modeled their conjecture about the investigated phenomenon in the target population, first as hands-on and after a while as TinkerPlots™ representations. Handouts included questions, such as, "would your inference apply also to a larger group of students such as the whole class?" Each lesson of the first nine activities included a whole class introductory discussion of the investigated topic, data investigation in small groups, and a whole class synthesis discussion of students' findings. In the tenth activity students summarized

[1] www.tinkerplots.com

their investigations to present them in a student-parents event and write a final report about their findings. To build a collaborative culture of inquiry in the classroom, we encouraged the students to share ideas, products and actions, reflect about the learning processes and share their insights. The intervention began with a statistics pre-test and ended with an identical post-test that focused on reasoning with data, distribution, covariation, and informal inference. This learning trajectory preceded and significantly expanded the instruction according to the national fifth grade statistics curriculum, which focuses only on mean and median in a procedural manner.

4.4.2 The Participants

We fully documented the learning processes of 12 pairs of students. In this study, we focus on the development of ARwC of a pair of boys—Orr and Guy. Orr is an academically successful student who has high achievements in mathematics and science. He has a learning disability, which limits somewhat his ability to express himself verbally or in writing. Guy's academic achievements are usually low. The students were selected due to high motivation, creativity and interest in the investigation.

In the first six activities of the learning trajectory (Table 4.1), Orr and Guy investigated univariate distributions and associations between categorical and numerical attributes. During whole class meetings, they discussed statistical ideas, such as center and representativeness, variability and outliers, comparing groups and covariation, and ways to represent and articulate these ideas. During the seventh to the ninth activities, the data investigations dealt with the relations between the amount of push-ups one can make in a row ("Push-ups"), and the 900-m running time in seconds ("Running").

4.4.3 Data Collection and Analysis

The students' investigations were fully videotaped using Camtasia™ to capture simultaneously their computer screen, discussions and actions. Data were observed, transcribed and annotated for further analysis of the students' ARwC. Data that were significant to this article were translated from Hebrew to English. Differences of meanings between Hebrew and English connotations of words were discussed extensively, to make sure the original intention of the speaker is clear.

The analysis process focused on the students' ARwC, using interpretative microgenetic method (Siegler 2006). We examined the entire cohort of data and narrowed it down to reasoning aspects that assemble a narrative of the students' emerging ARwC. Each reasoning aspect is composed of one or more statements of the participants. This process involved many rounds of data analysis sessions and meeting with expert and novice statistics education peers, in which interpretations were suggested, discussed, refined or refuted. This process involved searching forward and

Table 4.1 The Connections 2015 actual learning trajectory

Activity synopsis	Statistical ideas and concepts	No. of lessons
The first cycle: several contexts (favorite movie, usage of technology, favorite snack etc.)		
Activity 1: Introduction—modelling and data analysis of the class (n = 25 cases)		4
Ask a question, plan, collect data from the whole class, model a conjecture about the target population, analyze and organize data, present investigations (hands-on)	The investigative cycle, Uncovering children's initial concepts of reasoning with data, modelling and IIR	
Activity 2: Data analysis of the class (n = 25 cases)		2
Investigate a data sample from activity 1 (25 cases 2 variables, TinkerPlots™) , present investigations	How to ask a statistical question, The investigative cycle, initial reasoning with data, modelling and IIR	
Activity 3: The favorite snack activity—modelling and data analysis of the class (n = 25 cases)		2
Investigate data sample (25 cases, 3 variables), model a conjecture about the target population (TinkerPlots™), present investigations	Initial reasoning with data, distribution, comparing groups, modelling and IIR	
Activity 4: The favorite snack activity—modelling and data analysis of the grade (n = 73 cases)		2
Investigate the same question (activity 3) on a larger sample (73 cases, 3 variables, TinkerPlots™), model a conjecture about the target population (hands-on), present investigations	Initial reasoning with data, distribution, comparing groups, modelling and IIR	
The second cycle: the amazing race activities		
Activity 5: Data collecting of the entire fifth grade		2
Introduction to the activity context, the electronic questionnaire, measurements errors, data collection	How to answer a statistical question, Awareness to the importance of precise data	
Activity 6: Modelling and data analysis (n = 10 cases)		3
Model a conjecture about the target population (hands-on), investigate data sample (10 cases, 18 variables, TinkerPlots™) , refine model, present investigations	Correlations, Noise in the data as a result of measurement errors, The meaning of the median, reasoning with modelling	

(continued)

Table 4.1 (continued)

Activity synopsis	Statistical ideas and concepts	No. of lessons
Activity 7: Modelling and data analysis of the class (n = 24 cases)		2
Investigate the same question (activity 6) on a larger sample (24 cases, 18 variables, TinkerPlots™), refine the previous model (hands-on), present investigations	Growing samples, reasoning with data, distribution, comparing groups, modelling and IIR	
Activity 8: Modelling and data analysis of the grade (n = 62 cases)		2
Investigate the same question (activity 7) on a larger sample (62 cases, 18 variables, TinkerPlots™) , refine the previous model (hands-on)	Growing samples, reasoning with data, distribution, comparing groups, modelling and IIR	
Activity 9: Modelling of the target population		4
Model the conjecture about the target population (TinkerPlots™ Sampler)	Reasoning with data, modelling and IIR	
Activity 10: Summarizing		5
Prepare a presentation of the investigation and the target population final model, present investigations and models in a student-parents event, write a final report with parents	Reasoning with data, modelling and IIR	
Total number of 45-min lessons		**28**

backward over the entire data to find acceptable evidences for the researchers' local interpretations and hypotheses (e.g., Ben-Zvi and Arcavi 2001). To strive for "trustworthiness" (Creswell 2002), inferences about students' reasoning were called only after all data sources (interviews, TinkerPlots™ files and students' notes) provided sufficient evidence, and interpretations from different theoretical perspectives and by a number of researchers were examined (Triangulation, Schoenfeld 2007).

4.5 Results

In this, we present and explain Orr and Guy's emergent ARwC while they made informal inferences and modeled the population. We identify and characterize four reasoning aspects of the students' ARwC in their learning progression. The reasoning aspects varied according to: (a) the analysis unit the students used to examine covariation (for example, a single case, a small group of cases, etc.); (b) the way the students reasoned with signal and noise; and (c) accounted for variability within and between attributes. These aspects represent the key stages of the students' reasoning and helped us to follow the complex process of students' flow of ideas, hesitations,

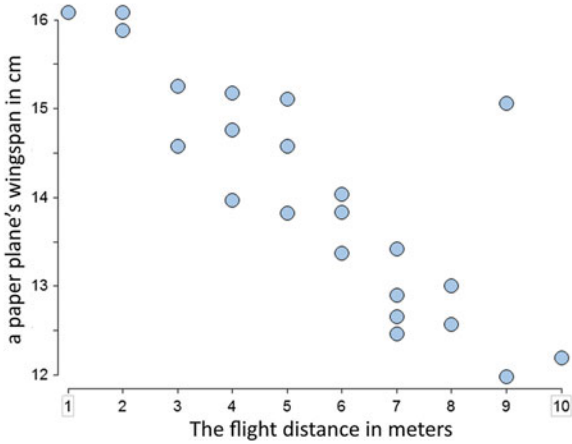

Fig. 4.1 A relationship between paper planes' wingspan (cm) and their flight distance (m)

mistakes and inventions. The generalizability of this suggested categorization needs to be further studied. Before the description of the four reasoning aspects, we provide the results of the pre-test analysis, and at the end we provide the results of the post-test analysis (the pre- and the post-tests were identical).

4.5.1 Initial Aggregate Reasoning with Covariation (Pre-test)

We analyzed relevant questions from Orr and Guy's pre-tests to reveal the students' initial perceptions of data and variability. In the first question, the students were asked to describe the height distribution of all fifth grade students in Israel, Orr and Guy wrote a single value (145 cm). Later at the same question they were asked to describe the distribution of students' heights in a typical fifth grade class (about 30 students). Orr suggested a range (135–160 cm) while Guy suggested specific height values attached to names of students in his school's fifth grade. In the third question, a scatter plot of a relation between a paper plane's wingspan and its flight distance was presented (Fig. 4.1). The students were asked what the graph described and what can they learn about the relation between paper plane's wingspan and flight distance. Then, they were asked to find where a paper plane of 14 cm wingspan would arrive. Guy's answers were related to only one attribute of the relation. He did not suggest any additional data case. Orr described the relation using the language: "the more… the more". He speculated that a 14 cm wingspan plane would land at the same place as a plane with the exact same wingspan that already appeared in the graph.

Thus, we learn that the students' reasoning with data and variability at the beginning was not aggregative. They held a local view of distributions and although Orr articulated aspects of data aggregation in covariation ("the more… the more"), it seems that his perception of variability was partially interpolating a value deter-

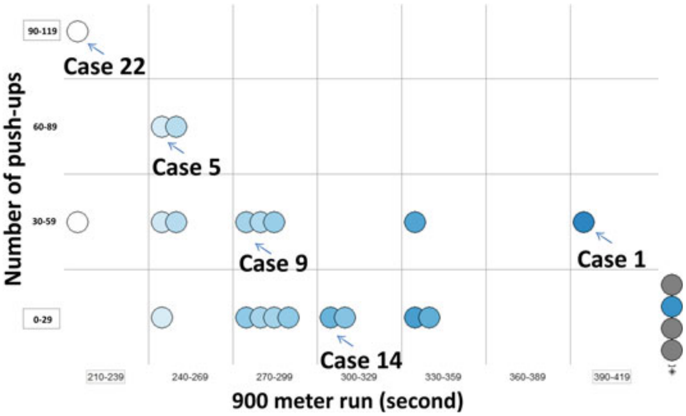

Fig. 4.2 Noticing an initial trend line in the class sample (n = 24)

ministically. Thus, the students' responses in the pre-test can be considered "single aspects" responses (Moritz 2004).

4.5.2 Aspect 1: A Pointwise-Based Covariation Model

In Activity 7, Orr and Guy extended their previous investigation about the running distribution to study the relations between running and push-ups (a sample of 24). They struggled with formulating an aggregate research question in the first task that reflected their contextual knowledge as well as the essence of the relation. Orr suggested to ask whether push-ups influenced the 900-m run results, while Guy rejected the idea of dependency and suggested asking: "Do other sports relate to running?" The students then drew a scatterplot in TinkerPlots™ and discussed the relations in the data (Fig. 4.2).

132 Guy We saw that 93 [push-ups, Case 22] also made a good running [time].
133 Int. That a person who did many push-ups…
134 Guy Many push-ups, which is 93, so the [running] result is also good. …
143 Int. Do you see, that the higher the push-ups is …
144 Orr It gives less… The [running] result is lower.
148 Guy Here [Case 5], it [push-ups] went down a little, and the [running] result is lower; also here [Case 9] the [running] result is lower, and it [push-ups] went down further more [Case 14] …
149 Orr [Continues] and the [running] result is lower. So let's draw a line. [They eventually did not draw this trend line.]
152 Guy And here [Case 1], It [push-ups] is also quite low, and the lowest [worst running] result.

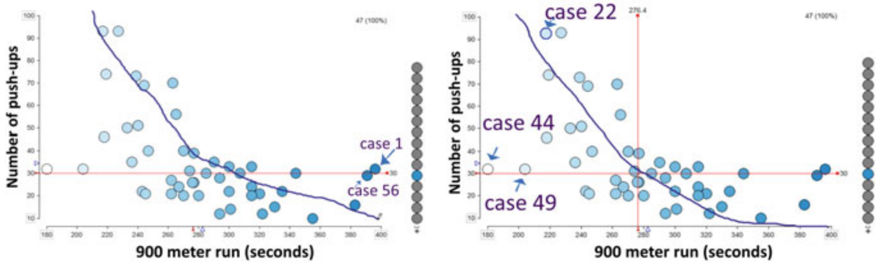

Fig. 4.3 **a** The first trend line and the outliers. **b** The second and "reasonable" trend line

Although the students responded to the researcher's efforts to encourage them to articulate an aggregative expression about the bivariate data [143], they looked mostly locally at the data, starting from an extreme case [Case 22]. They discussed variability by attending to the attributes' values of each of the five cases (Fig. 4.2) [e.g., 134] and to the value change in relation to the previous case ["it went down a little", 148]. To provide evidence for covariation, they identified a collection of four points [Cases 22, 5, 9 and 14]. These points made a pattern—*a pointwise-based covariation line*, which they planned to draw but eventually did not [149]. They were possibly inspired by a previous class discussion about trends. Attention to an outlier [152], which did not fit their line [Case 1, worst running result], led them to relate to the noise in the data. We consider their ARwC reasoning at this juncture as *a pointwise-based covariation model*.

4.5.3 Aspect 2: An Area-Based Covariation Model

In Activity 8, following the growing samples method, the students studied a larger sample of the entire fifth grade (47 cases out of 62 cases). They generated a scatter plot, added the means, medians and a horizontal reference line for the median of the push-ups (Fig. 4.3a). Guy drew a descending trend line using the TinkerPlots™ pen and commented that there were cases not represented by this line [Cases 1 and 56 in Fig. 4.3a]. He drew a new "reasonable" [89] trend line in the middle of the data cloud, mostly without passing through cases, but rather between them, and added a vertical reference line of the median running time, which separated the graph into four quadrants (Fig. 4.3b).

The students explained their findings while describing the relation in each of the quadrants.

122 Guy The more you approach the height here [at the upper-left quadrant in Fig. 4.3b] … in the push-ups, the more you progress here, besides maybe these ones [Cases 44 and 49] … When you are in this area [the upper-left quadrant], it means that he did a lot [a good result] in the running … If

you are here [Case 22], it means that you do a lot of push-ups and also a good running result.

123 Orr But if you are here [the lower-right quadrant], it means that you ran …

124 Guy [interrupting] It means that it is quite little push-ups

125 Orr And the result [of the running] is slow.

Orr noticed then that there is another area type in this graph that was created by the intersection of the two median reference lines. Guy drew a circle around it (Fig. 4.4a) and explained that this area is "the center of all these [cases]. The [trend] line that connects between all these [cases] is about the center of the entire [fifth] grade. These are the average [students] of the entire grade" [130]. Later on, the students expanded their view of the covariation and refined their model without any further prompting by the researcher. They speculated about hypothetical data: "If there was one [kid] here [upper-right quadrant], he would not be part of this [attributes relations], since he does many push-ups and a low [running] result" [Guy, 219]. The students summarized their informal inference by saying that "the faster a person runs, the more push-ups he does" [Orr, 134]. When they compared between their class (n = 24) and their grade (n = 62) samples, they used their new model (Fig. 4.4a) to construct a new representation for their class data (Fig. 4.4b) and were surprised by the similarity between the two samples (Fig. 4.4, b). Although their trend line did not pass through the intersection of median lines, the implementation of their model led the students to discover that the medians' locations are related to the trend line. They concluded that the medians "show you where the line passes" [Orr, 147].

The students thus reasoned at this juncture with, what we term, an *area-based covariation model*. The combination of the trend line, the four quadrants and the area around the medians' intersection constructed an *area-based model* for the main features of covariation. This model defined the presence of covariation as a phenomenon in which data cases are located on either upper-left or lower-right quadrants, spread around the signal—the trend line, and sometimes vary a lot (e.g., Cases 44 and 49 in Fig. 4.3b). Outliers were considered as data that appeared on the edges of quadrants. They reasoned with three aspects of the trend line: (a) the location of the trend line; (b) the trend line representativeness of the covariation; and (c) the trend line features (such as, its relation to center measurements). We view the students' cautious atten-

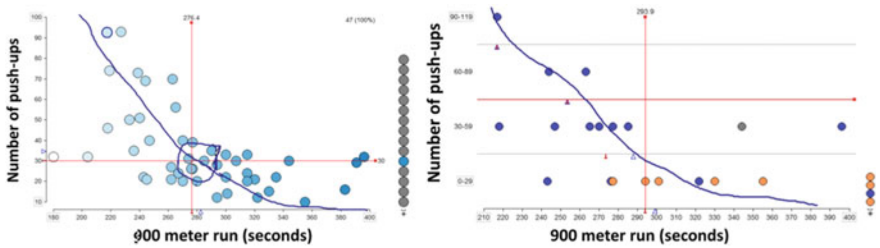

Fig. 4.4 **a** The relation between push-ups and running an area-based covariation model (n = 62). **b** The relation between push-ups and running colored by gender—the class data sample (n = 24)

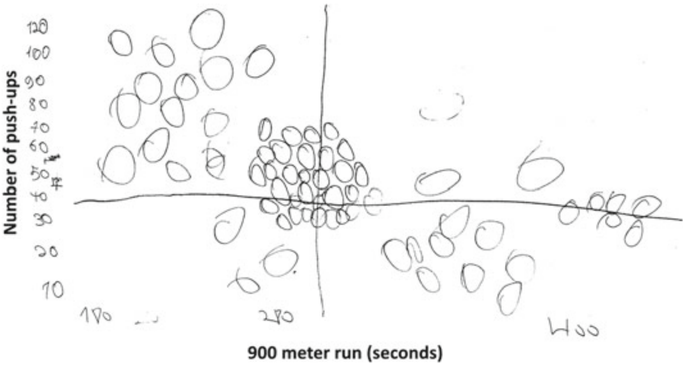

Fig. 4.5 Orr and Guy's conjecture about the target population

tion to these aspects as an attempt to summarize covariation in a way that attended to a large amount of cases and as part of their struggle to view bivariate data aggregately. They reasoned with variability between and within attributes, in relation to a prototype case of a certain area [122]. Using their *area-based covariation model*, they talked about covariation in a rather advanced language: "the more… the more" [122].

4.5.4 *Aspect 3: A Cluster-Based Covariation Model*

The next task in Activity 8 was to draw a conjecture about the population of all fifth graders in Israel ("the target population") on a piece of paper. The students followed the researcher's advice to first describe the push-ups distribution and drew a normal-shaped distribution. This time, they chose not to duplicate the data representation from the TinkerPlots™ real data graph, as they did previously. Rather, they spent time reasoning with aspects of the distribution, considering the data at hand and their beliefs. Later on, they drew the covariation between the attributes and explained their conjectured graph (Fig. 4.5): "[The graph] will be about the same as this one (Fig. 4.4a). There will be many in the middle [the medians intersection]. There will be a many here [the upper-left quadrant] and here [the lower-right quadrant]. About the same amount here and here [upper-left and lower-right quadrants], and here [in the center] also a lot, more than the two of them [upper-left and lower-right quadrants]" [Guy, 323].

The students expressed their conjecture about the target population using, what we term, a *cluster-based model of covariation* (Fig. 4.5). They related to the whole data in a way that expressed covariation between the attributes, by presenting the bivariate data in three clusters with common properties of size and density. They accounted for the variability in the data by noticing the signal and the noise. The signal was the pattern of the correlation, its trend and shape, in terms of the three

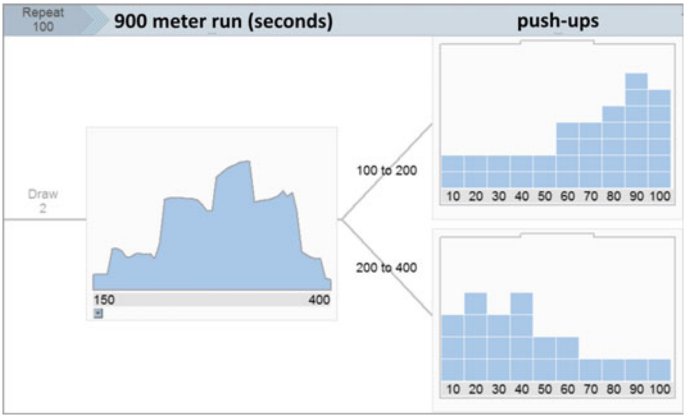

Fig. 4.6 A model of the running–push-ups relations in the population

clusters. They related to outliers and cases at the edges of the quadrants as noise. However, by clustering, they did not relate to the continuity of the aggregate.

4.5.5 Aspect 4: Conditional Distribution Model of Covariation

In Activity 9, the students constructed a model of their target population conjectures (Fig. 4.6) using the TinkerPlots™ Sampler. They first reasoned with the shape and range of the running distribution and defined it as a symmetric "tower with small steps." They then modeled the dependency between the attributes by separating the running range to two (100–200 and 200–400 s) and explained their choice: "A champion runner will run [900 meters] in a minute and a half, which is about a hundred [seconds]. … There is no chance that someone [in our sample] will run in a minute and a half" [Orr 26]. They added: "If you ran fast, then you also made a lot of push-ups. If you ran slowly–you made [less push-ups]" [Guy, 37]. The students thus constructed a model of the relations between the attributes while considering variability within and between them, and both data and context.

We term these actions and reasoning as a *conditional distribution view of covariation*, as they described the dependent attribute as two distinct skewed distributions conditioned on the values of the independent attribute. The students' analysis unit in this aspect was the whole data. Signal was described in relation to the analysis unit and both attributes, while considering continuity in the data. Noise was attended while reasoning with the range, shape, center and tendency of each distribution.

4.5.6 *Articulations of Aggregate Reasoning with Covariation (Post-test)*

We present briefly the results of the post-test, which was identical to the pre-test, to evaluate progress in the students' ARwC. In the first question, both students considered the center, shape and spread of the investigated phenomenon while drawing a height distribution of all fifth grade students in Israel. In the third question, Orr wrote that the graph (Fig. 4.6) described a relationship between planes' wingspans and their flight distance in meters. Both students used an aggregative language to describe this relation as: "the more… the more". When they speculated about a possible flight distance value for a certain plane's wingspan value, Orr interpolated the data case considering variability. Guy considered the center as well, as he explained: "most of the planes are there. Therefore, I thought that [the suggested area] was the average". According to the post-test analysis, the students reasoned with the distribution as an aggregate. Moreover, it seems that there was a progress in their ARwC. When they reasoned with the relation, they described it aggregately and considered aggregate aspects of the relation, such as variability, center and spread.

To sum up, we identified four reasoning aspects that describe the progression of the students' ARwC. In the beginning they described covariation through single cases only. Next, they reasoned separately about areas in the graph, while considering carefully a representative signal within the noise. Their conjecture about the population extended the latter aspect by considering the relations between clusters in the data. Finally, the students related to the whole data, considering continuity and variability in it as well.

4.6 Discussion

This research aims to study the characteristics of the emergent ARwC of two fifth grade students (age 11) who were involved in modelling activities of bivariate data and drawing ISIs in growing samples investigations. We address this goal by carefully analyzing Orr and Guy's emergent processes of ARwC throughout their learning progression. In the following sections, we first describe the students' emergent ARwC and the theoretical implications of our analysis. We then elaborate on the role of the tool and the design approach followed by the research limitations and conclusions. Our main theoretical and pedagogical lessons from this study are:

1. A suggested four-aspect framework of students' emergent ARwC in a learning environment that involves modelling activities and drawing ISIs in growing samples pedagogy.
2. Reasoning with variability and reasoning with modelling play a role in the development of ARwC.

3. The growing samples method, the generation and refinement of models, and the design for purpose and utility are important elements that can support the emergence of ARwC.

4.6.1 Aggregate Reasoning with Covariation

In this case study, we identified four different aspects of students' emergent ARwC. These aspects grew from a "single aspects responses" (Moritz 2004), which we identified in the pre-test stage. The four reasoning aspects depict the students' progress from perceiving covariation as a pointwise-based covariation model, an area-based covariation model, a cluster-based covariation model, to conditional distribution model. These reasoning aspects differ by the ways the students attempted to: (1) define an analysis unit to examine covariation; (2) reason with signal and noise; (3) account for variability; and (4) communicate about the correlations between the attributes (discourse about covariation).

The students initially perceived covariation as *a pointwise-based covariation model* (Fig. 4.2). The analysis unit was a single case, starting from extreme values as the most noticeable signal, and following the descending slope of a pointwise line in selecting additional key cases. Cases that only partly met the defined relation were considered as noise. They reasoned with both variability between and within attributes in relation to single cases. Their discourse related to a single case and the way the attributes behaved with regard to this case. When the students analyzed the bigger sample, a new reasoning aspect had emerged: *an area-based covariation model* (e.g., Fig. 4.4a). The analysis units were four quadrants generated by the median reference lines and the area around their intersection. The trend line was considered to be the signal. Cases that were at the edges of the quadrants were considered as noise and outliers. The students reasoned with variability between attributes and discussed covariation in relation to the way the attributes varied within a prototype case of a certain area, considering all cases in the analysis units. When the students modeled their conjecture about the target population, they extended their previous reasoning aspect to *a cluster-based covariation model*. In this aspect, the students described covariation in three main clusters that have common properties (size and density): (1) the center—the "average students" in the population; (2) the upper-left quadrant—the fast runners who do lots of push-ups; and (3) the lower-right quadrant—the slow runners, who hardly do any push-ups. The analysis unit in this perception was the whole data. The signal was the pattern of the correlation, its shape in terms of clusters and the existence of a trend. Cases at the edge of the quadrants were considered as noise, and variability was discussed in relation to the analysis unit, by attending to both attributes. However, they did not consider continuity in the data. The final reasoning aspect we identified in the students' ARwC was a *conditional distribution model*. In this view, the students described the data as a model of two attributes, where they described one attribute by its conditional distribution given the other. The analysis unit in this perception was the data as a whole. Signal was

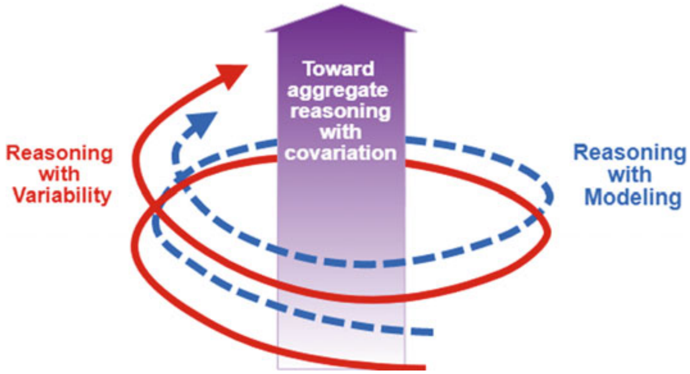

Fig. 4.7 The co-emergence process of ARwC, reasoning with modelling and variability

described in relation to the analysis unit and both attributes, while considering continuity in the data. Noise was attended to while reasoning with the range, shape, center and tendency of each distribution.

It seems, that the emergent process of ARwC involves a progression in the way students view data and covariation. They shift from a local pointwise view of data to an aggregate reasoning with data. Along this journey, they negotiate their understandings of data, variability, center and models to attend to the whole data. This process entails the construction of new understandings of the data at hand, the context of the investigated phenomenon and statistical concepts.

4.6.2 Theoretical Implications

The analysis of this case study distinguished two main processes that seemed to play a central role in the development of ARwC: Reasoning with variability and reasoning with modelling (Fig. 4.7).

4.6.2.1 Reasoning with Variability

We identified in this case study a progression in the students' reasoning in line with the literature (e.g., Garfield and Ben-Zvi 2008; Garfield et al. 2007; Shaughnessy 2007). The students' initial reasoning with variability was first expressed at the pre-test. They hardly attended to variability in data (e.g., represented distribution as single value and did not relate to measurement variability). Later, they related to more informal aspects of variability, such as: identifying that one variable varies more than the other (aspect 2), reasoning with variability with regard to the trend line (aspect 2), reasoning with variability of both variables to reason with the relationship between the two variables (aspects 3 and 4), reasoning with different representations to view different aspects of variability (aspect 3) and considering measures of variability and center as related while reasoning with data (aspects 3 and 4, and post-test).

We view this process as a key component in the emergence of the students' ARwC. The pointwise-based covariation model (the first reasoning aspect) emerged from concentrating on extreme values and an examination of covariation locally. Such a view toward data restrained the students from reasoning aggregately with data. However, it drew their attention to outliers that did not exactly fit their pointwise-based covariation model. This result is in line with Ben-Zvi and Arcavi (2001) concerning pointwise local view of data and the role of an outlier in developing an aggregate view. On the second reasoning aspect, the students' attention to the variability in data led them to refine the covariation model, i.e., the trend line they drew. We noticed their growing sensitivity to the need to attend to a larger amount of data cases while reasoning with covariation. Their attempt to confront this need was the area-based covariation model. In this model, variability raised the need to justify covariation and characterize each area in relating to all data cases within the certain area. In the cluster-based covariation model (aspect 3), the need to represent covariation led the students to confront variability as they compared clusters in their model and characterized the relations between them. This process extended the analysis unit to the whole data. At the final reasoning aspect, the need to consider the variability of one attributes as depending on the other led the students to extend their ARwC. They considered the whole data, as well as possible interpolations and extrapolations of data, as they constructed the conditional distribution model. They also reasoned with the relation between the center of the distribution and its spread and shape.

4.6.2.2 Reasoning with Modelling

We see the developing process of modelling as another important component in the emergence of ARwC. We assume that each step of a statistical investigation entails a process of emergence, development, refinement or verification of a model (Wild and Pfannkuch 1999). In this case study, some of these models were developed to represent and think or make predictions about the investigated phenomenon (the pointwise line, aspect 1; the trend line, aspect 2; representations of the students' conjectures, aspects 3 and 4). These modelling processes involved attempts to simplify the investigated phenomenon and reason aggregately with data (e.g., Pfannkuch and Wild 2004).

The students' modelling process also entailed the development of the students' epistemological model of the ARwC concept, which we term an *ARwC model*. The emergence of the ARwC model is in line with Ainley et al. (2000) epistemological analysis of trend that includes the sub-elements: correlation, linearity, interpolation and extrapolation. In the context of aggregate reasoning, the students' search for meaning of trend included a search for the relationship between two variables and its representation as a trend. We see this search in the students' request to refine the trend to represent more data (aspect 2), and in their discovery of the relationship between distributions' centers and the position of the trend (aspect 2). The area-based covariation model extended the ARwC model to a structure of reference and trend lines and a circle. The students used this structure to express and later to examine and evaluate the existence of covariation in data samples of different sizes. We assume that the development of this model facilitated and even promoted the students' perceptions of ARwC to rely on larger amount of data cases as they reason with data. We see the change of the utility of the ARwC model as a transition from a "model *of* ARwC" to "model *for* ARwC" (Gravemeijer 1999), as it involved the emergence of a new view of the ARwC concept.

To sum up, this case study implies that reasoning with variability and reasoning with modelling can play an important role in the emergence of ARwC. We see this role as supporting the emergence of ARwC, as well as the growing of understandings of the concepts of variability and modelling (Fig. 4.7). We suggest that this process entails reasoning with the data in hand as well as reasoning with the meaning of statistical ideas (e.g., "model for ARwC"). Further research is needed to study the nature of the roles these aspects play in the emergence of students' emergence of ARwC.

4.6.3 Pedagogical Implications

It seems that the students' learning progression was supported by the design of the learning environment (Ben-Zvi et al. 2018): the growing samples method, the generation and refinement of models and the design for purpose and utility. One of the advantages of the growing samples method is the students' focus on predictions, while viewing these predictions as temporary (see Ben-Zvi et al. 2012). In this case study, the growing samples method elicited the need to summarize data in a way that allows the students to examine their inferences within different size data sets. This requirement brought the need to attend to the signal within the noise. When the students reasoned with different sized samples, they needed to adapt their inferences to a larger data sample. This requirement encouraged them to attend to a larger group of data (e.g., the refinement of the trend line, aspect 2) and later to consider the whole data and possible population (aspects 3 and 4). We assume that the need to model the conjecture about the target population provided a reasonable utility to the data analysis. It also encouraged the students to express their ARwC considering signal, noise and uncertainty (aspect 3) and later dependency and continuity in data in relation to the whole data (aspect 4).

The dynamic TinkerPlots™ affordance to shift easily between representations helped the students to extend their view and role of the trend line as an aggregate representative of data as a whole (aspects 2 and 4). The TinkerPlots™ Sampler allowed generating and evaluating different representations, in the search for the one that best expressed the main properties of the investigated concept.

4.6.4 Limitations

This description is far from being a complete description of students' complex emerging processes of ARwC. The two students chosen for this research were considered by their teacher to be both able. This choice was made to enable the collection and analysis of detailed data about their ARwC during the intervention. Even after validating the data interpretation, the idiosyncrasy of the phenomena observed in this research remain questioned. More analysis of students' ARwC should be done within the Connections 2015 learning environment, as well as, further research in other learning environments, to further study students' ARwC.

4.7 Conclusions

This case study presented a new possible learning progression and reasoning aspects of students' ARwC. Students may initially hold local views of covariation. However, when students face covariation in data in such a multi-faceted learning environment, they start considering aspects of reasoning with covariation and develop a sense of the aggregate. Such a reasoning process is involved with handling and confronting variability in data and creating different types of models to analyze data and give meanings to the concept of variability (Fig. 4.7). It seems that this new line of research can advance our ongoing efforts to understand and improve the learning of statistics.

Acknowledgements This research was supported by the University of Haifa and the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation grant 1716/12. We deeply thank the Cool-Connections research group who participated in the Connections project 2015, and in data analysis sessions of this research.

References

Ainley, J., Nardi, E., & Pratt, D. (2000). The construction of meanings for trend in active graphing. *International Journal of Computers for Mathematical Learning, 5*(2), 85–114.

Ainley, J., Pratt, D., & Hansen, A. (2006). Connecting engagement and focus in pedagogic task design. *British Educational Research Journal, 32*(1), 23–38.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools* (A Ph.D. Thesis). Utrecht, The Netherlands: CD Beta Press.

Bakker, A., Biehler, R., & Konold, C. (2004). Should young students learn about boxplots? In G. Burrill & M. Camden (Eds.), *Curricular development in statistics education, IASE 2004 Roundtable on Curricular Issues in Statistics Education, Lund Sweden*. Voorburg, The Netherlands: International Statistics Institute.

Bakker, A., & Gravemeijer, K. P. E. (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.

Bakker, A., & Hoffmann, M. (2005). Diagrammatic reasoning as the basis for developing concepts: A semiotic analysis of students' learning about statistical distribution. *Educational Studies in Mathematics, 60*, 333–358.

Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environment. In J. B. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 191–205). Voorburg, The Netherlands: International Statistical Institute.

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

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM—The International Journal on Mathematics Education, 44*(7), 913–925.

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. In D. Ben-Zvi., K. Makar & J. Garfield (Eds.), *International handbook of research in statistics education*. *Springer international handbooks of education* (pp. 473–502). Cham: Springer.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643–690). Berlin: Springer.

Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events. *Journal for Research in Mathematics Education, 33*(5), 352–378.

Cobb, P., McClain, K., & Gravemeijer, K. P. E. (2003). Learning about statistical covariation. *Cognition and Instruction, 21*(1), 1–78.

Confrey, J., & Smith, E. (1994). Exponential functions, rates of change, and the multiplicative unit. In P. Cobb (Ed.), *Learning mathematics* (pp. 31–60). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Creswell, J. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Prentice Hall.

Dvir, M., & Ben-Zvi, D. (2018). The role of model comparison in young learners' reasoning with statistical models and modeling. *ZDM—International Journal on Mathematics Education*. https://doi.org/10.1007/s11858-018-0987-4.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht, The Netherlands: Kluwer.

Friel, S. (2007). The research frontier: Where technology interacts with the teaching and learning of data analysis and statistics. In G. W. Blume & M. K. Heid (Eds.), *Research on technology and the teaching and learning of mathematics* (Vol. 2, pp. 279–331). Greenwich, CT: Information Age.

Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal, 4*(1), 92–99.

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

Garfield, J., delMas, R. C., & Chance, B. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 87–116). New York: Lawrence Erlbaum.

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

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic enquiry with data: Critical barriers to classroom implementation. *Educational Psychologist, 27*(3), 337–364.

Konold, C. (2002). Teaching concepts rather than conventions. *New England Journal of Mathematics, 34*(2), 69–81.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2015). Data seen through different lenses. *Educational Studies in Mathematics, 88*(3), 305–325.

Konold, C., & Miller, C. (2011). *TinkerPlots (Version 2.0) [Computer software]*. Key Curriculum Press. Online: http://www.keypress.com/tinkerplots.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259–289.

Lehrer, R., & English, L. (2017). Introducing children to modeling variability. In D. Ben-Zvi, J. Garfield, & K. Makar (Eds.), *International handbook of research in statistics education. Springer international handbooks of Education* (pp. 229–260). Cham: Springer.

Lehrer, R., & Schauble, L. (2004). Modelling natural variation through distribution. *American Educational Research Journal, 41*(3), 635–679.

Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education, 96*(4), 701–724.

Lesh, R., Carmona, G., & Post, T. (2002). Models and modelling. In D. Mewborn, P. Sztajn, D. White, H. Wiegel, R. Bryant, et al. (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 89–98). Columbus, OH: ERIC Clearinghouse.

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning, 13*(1), 152–173.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82–105.

Makar, K., & Rubin, A. (2017). Research on inference. In D. Ben-Zvi., K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education*. *Springer international handbooks of education* (pp. 261–294). Cham: Springer.

Moore, D. S. (2004). *The basic practice of statistics* (3rd ed.). New York: W.H. Freeman.

Moritz, J. B. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–256). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 17–46). Dordrecht, Netherlands: Kluwer Academic Publishers.

Reading, C., & Shaughnessy, C. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201–226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Rubin, A., Hammerman, J. K. L., & Konold, C. (2006). Exploring informal inference with interactive visualization software. In *Proceedings of the Seventh International Conference on Teaching Statistics [CD-ROM], Salvador, Brazi*l. International Association for Statistical Education.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology, 32*(1), 102–119.

Schoenfeld, A. H. (2007). Method. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 69–107). Charlotte, NC: Information Age Publishing.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *The second handbook of research on mathematics* (pp. 957–1010). Charlotte: Information Age Publishing Inc.

Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (6th ed., Vol. 2, pp. 464–510). Hoboken, NJ: Wiley.

Watkins, A. E., Scheaffer, R. L., & Cobb, G. W. (2004). *Statistics in action: Understanding a world of data*. Emeryville, CA: Key Curriculum Press.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion). *International Statistical Review, 67*, 223–265.

Zieffler, S. A., & Garfield, J. (2009). Modelling the growth of students' covariational reasoning during an introductory statistical course. *Statistics Education Research Journal, 8*(1), 7–31.