



Statistical modeling processes through the lens of instrumental genesis

Marianne van Dijke-Droogers¹ · Paul Drijvers¹ · Arthur Bakker¹

Accepted: 25 December 2020/Published online: 02 April 2021
© The Author(s) 2021

Abstract

Digital technology is indispensable for doing and learning statistics. When technology is used in mathematics education, the learning of concepts and the development of techniques for using a digital tool are known to intertwine. So far, this intertwining of techniques and conceptual understanding, known as instrumental genesis, has received little attention in research on technology-supported statistics education. This study focuses on instrumental genesis for statistical modeling, investigating students' modeling processes in a digital environment called TinkerPlots. In particular, we analyzed how emerging techniques and conceptual understanding intertwined in the instrumentation schemes that 28 students (aged 14–15) develop. We identified six common instrumentation schemes and observed a two-directional intertwining of emerging techniques and conceptual understanding. Techniques for using TinkerPlots helped students to reveal context-independent patterns that fostered a conceptual shift from a model *of* to a model *for*. Vice versa, students' conceptual understanding led to the exploration of more sophisticated digital techniques. We recommend researchers, educators, designers, and teachers involved in statistics education using digital technology to attentively consider this two-directional intertwined relationship.

Keywords Statistical modeling · Instrumental genesis · Statistical reasoning · TinkerPlots · Simulated sampling distribution

The increasing use of digital technology in our society requires an educational move towards learning from and with digital tools. This is particularly urgent for statistics education where digital technology is indispensable for interpreting statistical information, such as real sample data (Gal, 2002; Thijs, Fisser, & Van der Hoeven, 2014). For such interpretations, understanding underlying statistical models is fundamental (Manor & Ben-Zvi, 2017). Current technological developments offer digital tools—for example, TinkerPlots, Fathom, and

✉ Marianne van Dijke-Droogers
m.j.s.vandijke-droogers@uu.nl

¹ Freudenthal Institute, Utrecht University, PO Box 85170, 3508 AD Utrecht, The Netherlands

Codap—that provide opportunities to deepen understanding of statistical modeling and models. These digital tools enable students to build statistical models and to use these models to simulate sampling data, and therefore offer means for statistical reasoning with data (Biehler, Frischmeier, & Podworny, 2017). As such, modeling with digital tools is promising for today's and tomorrow's statistics education.

Although statistics education is developing as a domain distinct from mathematics, the use of digital tools is a shared problem space and collaboration within shared spaces can strengthen each domain (Groth, 2015). From other domains in school mathematics, for example, algebra, it is well known that as soon as digital tools are used during the learning process, the development of conceptual understanding becomes intertwined with the emergence of techniques to use the digital tool (Artigue, 2002; Drijvers, Godino, Font, & Trouche, 2013). For teachers, researchers, educators, and designers, insight into this intertwined relationship of learning techniques and concepts is a prerequisite for deploying digital tools in such a way that they are productive for the intended conceptual understanding. In the meantime, due to a lack of insight into this intertwining, undesired influence of techniques for using the digital tool on the intended conceptual development can be overlooked. This complex relationship, however, has so far received little attention in research on technology-supported statistical modeling processes.

A useful perspective to grasp the relationship between the learning of digital techniques and conceptual understanding is instrumental genesis (Artigue, 2002). In this theoretical view, learning is seen as the simultaneous development of techniques for using artifacts, such as digital tools, and of domain-specific conceptual understanding, for example, statistical models and modeling. The perspective of instrumental genesis seems promising to gain knowledge about learning from and with digital technology. As such, the aim of this study is to explore the applicability of the instrumental genesis perspective to statistics education and to statistical modeling processes in particular.

1 Theoretical framework

In this section, we elaborate on two main elements of this study: statistical modeling and instrumental genesis.

1.1 Statistical modeling: techniques and concepts

Digital tools for statistical modeling have the potential to deepen students' conceptual understanding of statistics and probability and enable them to explore data by deploying techniques for using the tool. They also offer possibilities to visualize concepts that previously could not be seen, such as random behavior (Pfannkuch, Ben-Zvi, & Budgett, 2018). Such educational digital tools, for example, TinkerPlots, provide opportunities for statistical reasoning with data, as students build statistical models and use these models to simulate sample data (Biehler et al., 2017).

Modeling processes with a digital tool such as TinkerPlots require the development of digital techniques. Digital TinkerPlots techniques for setting up statistical models and simulating data are helpful to introduce key statistical ideas of distribution and probability (Konold, Harradine, & Kazak, 2007). The research by Garfield, delMas, and Zieffler (2012) suggests that students can learn to think and reason from a probabilistic perspective—or, as the authors

call it, “really cook” instead of following recipes—by using TinkerPlots techniques to build a model of a real-life situation and to use this model for simulating repeated samples. This way to understand the probability involved in inferences is also reflected in our previous study (Van Dijke-Droogers, Drijvers, & Bakker, 2020) in which an approach based on repeated sampling from a black box filled with marbles seemed to support students in developing statistical concepts. In this approach, students developed TinkerPlots techniques to investigate what sample results would likely occur by chance. Statistical modeling in the study presented here requires TinkerPlots techniques for building a model by choosing a graphical representation (e.g., a bar or pie chart), entering population characteristics (e.g., population size, attributes, and proportions), and entering the sample size, of a real-life situation from a given context to solve a problem. The next steps include TinkerPlots techniques for simulating repeated samples by running the model and visualizing the results in a sampling distribution, for enabling to reason about probability—taking into account number of repetitions and sample size—and to answer the problem using simulated data.

Statistical modeling processes with TinkerPlots also require, in addition to the development of TinkerPlots techniques, an understanding of the concepts involved. The literature elaborates several viewpoints on statistical modeling. We discuss three viewpoints and indicate how we incorporated them in our study. First, Büscher and Schnell (2017) argue that the notion of emergent modeling (Gravemeijer, 1999)—the conceptual shift from a model *of* a context-specific situation to a model *for*—can also be applied to statistical reasoning in a variety of similar and new contexts. Second, statistical modeling involves the interrelationship between the real world and the model world. This relationship is elaborated in Patel and Pfannkuch’s framework (2018) that displays students’ cognitive activities about understanding the problem (real world), seeing and applying structure (real world–model world), modeling (model world–real world), analyzing simulated data (model world), and communicating findings (model world–real world). Third, for reasoning with models and modeling, Manor and Ben-Zvi (2017) identify the following dimensions: reasoning with phenomenon simplification, with sample representativeness, and with sampling distribution. Statistical modeling includes the process of abstracting the real world into a model and then using this model for understanding the real world. In short, Büscher and Schnell (2017) emphasize the importance of developing context-independent models for statistical modeling processes, Patel and Pfannkuch (2018) outline the interaction between the real and the model world, and Manor and Ben-Zvi (2017) address the different dimensions when reasoning with models. These viewpoints provide insight into the development of concepts for statistical modeling. In the study presented here, we embodied the viewpoints in the design of students’ worksheets. On these worksheets, students are requested to build and run a model of a real-world situation in TinkerPlots and to use this model, by simulating and interpreting the sampling distribution of repeated samples, to understand the real-world situation.

Understanding and reasoning with the simulated sampling distribution from repeated samples are, as mentioned by Manor and Ben-Zvi (2017), essential for statistical modeling. However, the concept of sampling distribution is difficult for students. The study by Garfield, delMas, and Chance (1999) focused on the design of a framework to describe stages of development in students’ statistical reasoning about sampling distributions. Their initial conception of the framework identified five levels that evolve from (1) idiosyncratic reasoning—knowing words and symbols related to sampling distributions, but using them without fully understanding and often incorrectly—through (2) verbal reasoning, (3) transitional reasoning, and (4) procedural reasoning, towards (5) integrated process reasoning—

complete understanding of the process of sampling and sampling distributions, in which rules and stochastic behavior are coordinated. In our study, these levels will be used to indicate students' conceptual understanding of statistical modeling. Students' difficulties in reasoning with the sampling distribution are often related to misconceptions about basic statistical concepts such as variability, distribution, sample and sampling, the effect of sample size, and confusion of results from one sample with the sampling distribution. According to Chance, delMas, and Garfield (2004), ways to improve students' level of understanding statistical modeling include techniques for exploring samples, comparing how sample behavior mimics population behavior, and for both structured and unstructured explorations with the digital tool. As such, conceptual understanding of statistical modeling involves the building, application, and interpretation of context-independent statistical models—in our study the sampling distribution of repeated sampling—to answer real-life problems.

1.2 Instrumental genesis

Using digital tools in a productive way for a specific learning goal requires insight into the intertwined relationship between emerging digital techniques and conceptual understanding. A useful perspective to grasp the intertwining of learning techniques and concepts is *instrumental genesis*. A fundamental claim in this theory is that learning can be seen as the intertwined development, driven by the student activity in a task situation, of techniques for using artifacts—for example, a digital tool—and cognitive schemes that have pragmatic and epistemic value (Artigue, 2002; Drijvers et al., 2013). In this perspective, the conception of “instrument” and instrumental genesis are used in the sense described by Artigue (2002):

The instrument is differentiated from the object, material or symbolic, on which it is based and for which is used the term “artefact”. Thus an instrument is a mixed entity, part artefact, part cognitive schemes which make it an instrument. For a given individual, the artefact at the outset does not have an instrumental value. It becomes an instrument through a process, called instrumental genesis, involving the construction of personal schemes or, more generally, the appropriation of social pre-existing schemes. (p. 250)

According to Vergnaud (1996), a scheme is an invariant organization of behavior for a given class of situations. Such a scheme includes patterns of action for using the tool and conceptual elements that emerge from the activity. In the study presented here, the tasks on students' worksheet intend to construct personal instrumentation schemes consisting of TinkerPlots techniques and conceptual understanding of statistical modeling. The identification of schemes can structure and deepen the observation of students' emerging technical actions and statistical reasoning, and hence provides insight into the intertwined development of techniques and concepts.

As the application of instrumental genesis within the field of statistics education hardly exists, we present an example from a study within the context of algebra. Table 1 shows an instrumentation scheme concerning the use of a symbolic calculator for solving parametric equations, from a study by Drijvers et al. (2013). The intertwined relationship can be seen, for example, in scheme D. Here, students were asked to solve the parametric equation with respect to x . On the one hand, in order to use the correct techniques, students must be able to identify the unknown in the parameterized problem situation to enter the correct command “solve with respect to x ” into their computer algebra calculator. On the other hand, the available options of

Table 1 Example of an instrumentation scheme for solving parametric equations with a computer algebra system (Drijvers et al., 2013, p. 35)

Digital techniques	Conceptual understanding
A. Use the Solve-option of the Graphing Calculator and enter the given function	Knowing that the Solve command can be used to express one of the variables in a parameterized equation in other variables
B. Enter the '=0' sign	Knowing the difference between an expression and an equation
C. Enter the unknown to solve (x)	Realizing that an equation is solved with respect to an unknown
D. Solve the equation with respect to x	Being able to identify the unknown in the parameterized problem situation
E. Give the solution for the parametrized equation	Being able to interpret the result, particularly when it is an expression, and to relate it to graphical representations

Indicate the unknown to solve

An equation should contain an = sign

Notice the scope of the square root sign

'solve with respect to x' = 'express x in terms of b'

A solution can be an expression

the tool invite students to distinguish between the parameter and the unknown. In the study by Drijvers et al., the identification of students' instrumentation schemes provided insight into how the learning of techniques for using a computer algebra system and the conceptual understanding of solving parametric equations emerged in tandem. Furthermore, the identified schemes helped the researchers to reveal several conceptual difficulties students encountered while solving parametric equations with the digital tool.

As a second example, we present the findings from one of the scarce studies on instrumental genesis within the field of statistics education, conducted by Podworny and Biehler (2014). In their study, within a course on hypothesis testing and randomization tests with *p* values, university students used simulations with TinkerPlots. Students noted their own schemes to plan and structure their actions. These schemes drawn up by students proved useful as a personal work plan; however, it was difficult to identify common instrumentation schemes and to unravel how TinkerPlots techniques and conceptual understanding emerged together. Our study differs from theirs, as we identified our students' schemes by observing their actions and reasoning.

In general, instrumental genesis is considered an idiosyncratic process, unique for individual students. Yet, it takes place in the social context of a classroom, and as researchers, we are interested in possible patterns. As such, identifying instrumentation schemes concerns the complexity of unraveling patterns in the diversity of individual schemes that students develop. The study presented here seeks to identify common instrumentation schemes by observing students' actions and reasoning when statistically modeling in TinkerPlots, and then to use these schemes to zoom in on the genesis of the schemes to reveal how emerging TinkerPlots techniques and the conceptual understanding of statistical modeling intertwine.

2 Research aim and question

To explore the applicability of the instrumental genesis perspective to statistics education, and to statistical modeling, in particular, we conducted an explorative case study. This study focuses on 14-to-15-year-old students' intertwined development of learning techniques for using TinkerPlots and conceptual understanding of statistical modeling. We address the following question: Which instrumentation schemes do 9th-grade students develop through statistical modeling processes with TinkerPlots and how do emerging techniques and conceptual understanding intertwine in these schemes?

3 Methods

This study is part of a larger design study on statistical inference. Our previous study focused on the design of a learning trajectory in which students were introduced to the key concepts of sample, frequency distribution, and sampling distribution, with the use of digital tools (Van Dijke-Droogers et al., 2020). As a follow-up, this study focuses on the specific role of digital techniques on conceptual understanding by examining how 28 9th-grade students work on TinkerPlots worksheets, which were designed to engage students in statistical modeling.

3.1 Design of student worksheets

A suitable stage to investigate students' instrumental genesis—their development of schemes that include TinkerPlots techniques and conceptual understanding of statistical modeling—is after the introduction of the tool and the concepts, when they engage in the emergent modeling process of applying gained knowledge in new real-life situations. Prior to working with the TinkerPlots worksheets, students had a brief introduction to the tool and concepts. These preparatory activities were designed within the specific context of a black box with marbles and involved three 60-min lessons. Two of these lessons concentrated on physical black box experiments and one on simulations. Both the physical and simulation-based preparatory activities introduced students to statistical modeling by addressing concepts such as sample, sampling variation, repeated sampling, sample size, frequency distribution of repeated sampling, and (simulated) sampling distribution (Chance et al., 2004). The introduction of TinkerPlots techniques was done in the third 60-min lesson through a classroom demonstration of the tool by the teacher, followed by students practicing themselves using an instruction sheet. On this instruction sheet, the TinkerPlots techniques for making a model, simulating repeated samples, and visualizing the sampling distribution were listed. The brief introduction on techniques and concepts focused on the black box context only.

For the study reported here, we designed five worksheets. In one 60-min lesson per worksheet, we invited students to apply and expand their emerging knowledge from the preparatory black box activities in new real-life contexts. The design of the worksheets was inspired by studies from Patel and Pfannkuch (2018), Manor and Ben-Zvi (2017), and Chance et al. (2004). In each worksheet, the students were asked to build and run a model of a real-world situation in TinkerPlots and to use this model, by simulating and interpreting the sampling distribution of repeated samples, to understand the real-world situation. The structure of these worksheets is shown in Table 2. In each worksheet (W1 to W5), a new context was introduced. We chose contexts with categorical data to minimize the common confusion between the distribution of one

Table 2 Structure of designed worksheets 1–5

Worksheet component	Student activity
a. Explore and identify important factors from a given real-life problem to build a population model in TinkerPlots	Try to understand the situation and the data collection by defining the problem, making predictions, and considering variation. Apply structure by identifying all known real-world factors, considering model tools in TinkerPlots to represent real-world factors, and evaluating whether all relevant factors are included in the model
b. Build and run the model by simulating sample results and examine the behavior of the model	Use TinkerPlots to examine the behavior of the model by visualizing single sample results and repeated sample results in respectively sample and sampling distributions by, checking variation in simulated data distributions at varying sample sizes and varying number of repeated samples, comparing these data with the contextual knowledge, evaluating model fit by checking how simulated data mimic the model
c. Examine and interpret the simulated results by using the sampling distribution	Interpret the results of simulated data by identifying and using TinkerPlots tools to answer specific tasks and, in addition, by considering (the probability of a specific) range of outcomes, sampling variability, effect of sample size, and number of repeated samples
d. Answer the problem using the simulated data	Communicate findings by stating background to the problem, making model informed decisions, recognizing effects of underlying randomness, and stating limitations of the decisions

sample and sampling distribution (Chance et al., 2004) and to optimize the similarity with the black box context in the preparatory activities. When carrying out the tasks on W1 to W5, students could use the TinkerPlots instruction sheet from the preparatory activities. The aim of the worksheets was to expand students' understanding of statistical modeling—that is, the building, application, and interpretation, of context-independent statistical models; in our study, the sampling distribution of repeated sampling—by using TinkerPlots as an instrument.

3.2 Participants

We worked with two groups, each consisting of fourteen 9th-grade students. Group 1 consisted of students in school year 2018–2019 and group 2 of students in school year 2019–2020. All students were in the pre-university stream and thus belonged to the 15% best performing students in our educational system. The students were inexperienced in sampling and had no prior experience in working with digital tools during mathematics classes.

The students in group 1 went through the preparatory activities described earlier during the regular math lessons in school. Their teacher had been involved in the research project and had already carried out these lessons several times. All twenty students from the class were invited to participate in the session at Utrecht University's Teaching and Learning Lab (a laboratory classroom) and fourteen of them applied. During the lab session, these students worked on worksheets 1–3 (W1 to W3), the initial phase of the teaching sequence. For practical reasons—such as missing regular classes and travel time—multiple research sessions with the same students were not possible.

Therefore, one school year later, we performed lab sessions again, but with a different group of students, here called group 2. These students from the same school and with the same teacher as

Table 3 Overview of participants and data collection

Participants				Data collection	
Group	School year	Average math grade	Average grade preparatory tasks	Data from students' work	Total time duration recordings
1 ($n=14$)	2018–2019	6.9	8	W1–W3, initial phase	28 h
2 ($n=14$)	2019–2020	6.6	8	W4–W5, more advanced phase	17 h

group 1 went through the same preparatory lessons and W1 to W3 during their regular math lessons at their school. Again, fourteen students applied to participate in the research sessions at the university. These students in group 2 were similar to those in group 1: They performed at a similar level in mathematics, as their overall grades for the school year averaged 6.6 on a scale of 10, which was comparable to 6.9 in group 1. In addition, students' performance in the preparatory tasks averaged 8 on a scale of 10 in both groups. The teacher judged the starting level of the two groups to be similar. During the research sessions, the students of group 2 worked on W4–W5, the more advanced phase of the teaching sequence. An overview of participants can be found in Table 3.

3.3 Data collection

The data consisted of video and audio recordings from two classroom laboratory sessions. During the first 5-h session in Utrecht University's Teaching and Learning Lab, the fourteen students of group 1 worked in teams of two or three on the designed W1–W3. The advantage of this lab setting over a classroom environment was that detailed video recordings could be made of students' actions in TinkerPlots and their accompanying conversations. The students were specifically asked to express their thoughts while solving the problem, the think-aloud method (Van Someren, Barnard, & Sandberg, 1994). The teams worked on a laptop, the screen of which was displayed on an interactive whiteboard. Figure 1 shows the setup in the lab.



Fig. 1 Students working with TinkerPlots in Utrecht University's teaching and learning lab

During the second lab session, we collected video and audio recordings from fourteen students of group 2, while working on W4–W5.

3.4 Data analysis

The data analysis consisted of three phases: (1) identifying common instrumentation schemes, (2) examining the global scheme genesis process during the work, and (3) examining the scheme genesis process in depth (Table 4).

In phase 1 of the analysis, we used a combined approach of theory-driven (prior to data collection) and bottom-up (based on the data) to identify emerging instrumentation schemes. The final results can be found in Table 7. To identify the schemes, we conducted qualitative data analysis as defined by Simon (2019): A process of working with data, so that more can be gleaned from the data than would be available from merely reading, viewing, or listening carefully to the data multiple times (p. 112). In step 0, prior to the data collection, we defined preformulated schemes. These schemes were based on the theories on statistical modeling (Büscher & Schnell, 2017; Chance et al., 2004; Gravemeijer, 1999; Manor & Ben-Zvi, 2017; Patel & Pfannkuch,

Table 4 Overview of data analysis

Phase and objective	Outline	Data-driven/theory-driven
Phase 1: Qualitative data analysis Identification of students' instrumentation schemes	Step 0 (Prior to data collection): Preformulating instrumentation schemes based on theories on statistical modeling	Theory-driven
	Step 1: Observing each student at a certain local segment of the teaching sequence	Data-driven
	Step 2: Categorizing data from step 1, by using preformulated instrumentation schemes	Data- and theory-driven
	Step 3: Identifying global patterns of instrumentation schemes for more students, by using the categorized data of step 2	Data- and theory-driven
Phase 2: Interpretive content analysis Examination of students' scheme development by identifying their TinkerPlots techniques and levels of understanding statistical modeling during the work	Defining technical levels for TinkerPlots techniques and conceptual levels for understanding statistical modeling	Theory-driven
	Refining and specifying technical and conceptual levels in each instrumentation scheme	Data- and theory-driven
	Assigning levels to student actions when working on W1 and W5	Data- and theory-driven
Phase 3: Case study More detailed examination of students' scheme development	Further examining how TinkerPlots techniques and understanding statistical modeling intertwine in the schemes students develop, that is, how techniques may support conceptual understanding and the other way around, by zooming in on developing personal schemes of students	Data-driven

2018), and instrumental genesis (Artigue, 2002), and on expertise developed in previous interventions (Van Dijke-Droogers et al., 2020). In these preformulated schemes, specific TinkerPlots techniques were related to students' understanding of statistical modeling. In step 1, we observed each student at a certain local segment of the teaching sequence, for example, building a model of the population (W1 Task 5), and analyzed the techniques and concepts that were manifested in students' actions and reasoning at that local segment. In step 2, we categorized the data from step 1 by using the preformulated schemes. To do this categorization, at the same time as preformulated schemes were assigned, we expanded, refined and adjusted them to include the observed data. In step 3, we used the categorized data of step 2 to identify patterns for more students. By systematically and iteratively going through the categorized data, both within one student over several schemes and across students, we identified global patterns in emerging instrumentation schemes. These global patterns occurred to a certain extent in every student and across students while working on each worksheet. By adapting the preformulated schemes to the global patterns, we identified students' instrumentation schemes.

Concerning phase 2 of the analysis, the data for examining students' instrumental genesis, that is, their scheme development during the work, we used interpretive content analysis (Ahuvia, 2001). This variant of content analysis allowed us to identify both explicitly observed and latent content of students' technical actions and reasoning. To identify possible progress in students' TinkerPlots techniques, we defined five technical levels of proficiency. These levels were based on Davies' (2011) levels of technology literacy and refined by both our experiences from previous research and the collected video data. Davies defined six levels of users (non-user, potential user, tentative user, capable user, expert user, and discerning user) each of which corresponds to ascending levels of use: none, limited, developing, experienced, powerful, and selective. For our study (Table 5), we merged the last two levels of technology literacy, as our students were unable to reach the highest level in the short period of time working on W1 to W5. Based on the observed data, we specified the five technical levels for using TinkerPlots, for each instrumentation scheme. The specified technical levels were used to analyze students' scheme development during the work on W1 and W5, respectively. To identify possible progress in students' understanding of statistical modeling, we defined five conceptual levels. The conceptual levels are displayed in Table 6. The conceptual levels were merely based on the previously described levels by Garfield et al. (1999). These conceptual levels for understanding statistical modeling were further specified for each instrumentation scheme on the basis of the observed video data and prior experiences.

Table 5 Technical levels for using TinkerPlots

Level	Description
1. Non-user	Students are not able to carry out the TinkerPlots techniques
2. Limited user	Students carry out the TinkerPlots techniques by following the instructions stepwise; the techniques are still carried out hesitantly; haphazard trial and error or simply trying something
3. Developing user	Students carry out correct TinkerPlots techniques by following the instructions most of the time; incorrect techniques are used but later corrected
4. Experienced user	Students carry out correct TinkerPlots techniques fluently—that is, fast and without mistakes—sometimes augmented by newly explored TinkerPlots techniques
5. Expert/discerning user	Students make well-considered decisions for correct TinkerPlots techniques

Table 6 Conceptual levels of understanding statistical modeling

Level	Description
1. Incorrect reasoning	Wrong statements and/or incorrect using words and symbols related to the specific item of conceptual understanding
2. Idiosyncratic reasoning	Knowing words and symbols related to the specific item of conceptual understanding but using them without fully comprehending and often incorrectly
3. Verbal reasoning	Verbal understanding of the item but unable to apply it to the actual behavior. For example, the student can reproduce that results from a larger number of repeated samples lead to a better estimate of the population but does not understand how key concepts such as variability and range are integrated
4. Transitional reasoning	Correctly identifying one or two features of the item without fully integrating these features. For example, identifying, and relating just one or two of the features (1–3) in understanding that results from more repeated samples lead to a better estimate of the population: more repeated samples lead to (1) a smoother sampling distribution, without local peaks, with (2) a peak at the population proportion and (3) to an average that resembles the population
5. Integrated process reasoning	Understanding of the process of sampling and sampling distributions. For example, understanding the effect of simulating a larger number of repeated samples to the shape, peak, and average of the sampling distribution and that, as a consequence, more repeated samples lead to a better estimate of the population

The specification of both the technical and conceptual levels for coding the data was discussed in-depth with experts in this domain. Although the students worked in teams of two or three, we analyzed their proficiency levels individually. We did so as we noticed considerable differences in individual proficiency within one team and also because there was cooperation and consultation between teams. To check the reliability of the first coder's analysis, a second coder analyzed the video data of students' activities with W1 and W5, for both the coding of technical and conceptual levels. A random sample of 5% of the data (30 out of 600 fragments) was independently rated by the second coder. The second coder agreed on 85% of the codes. Deviating codes, which were limited to two levels difference at most, were discussed until agreement was reached.

In phase 3, to further examine the intertwined relationship between developing TinkerPlots techniques and understanding statistical modeling, that is, how techniques may support conceptual understanding and vice versa, we used case studies to investigate students' instrumental genesis. In these case studies, we zoomed in on developing personal schemes of students in both the initial and more advanced phase.

4 Results

In this section, we first present the six instrumentation schemes we identified, each including TinkerPlots techniques and conceptual understanding of statistical modeling (Table 7). Second, we describe students' global scheme development during the work, by presenting the levels at which the students used the techniques and concepts while working on worksheets 1 and 5. Third, we describe two students' cases to reveal in more detail the intertwining of emerging techniques and conceptual understanding in the personal instrumentation schemes that students developed.

Table 7 Instrumentation schemes for statistical modeling through simulating repeated samples with TinkerPlots

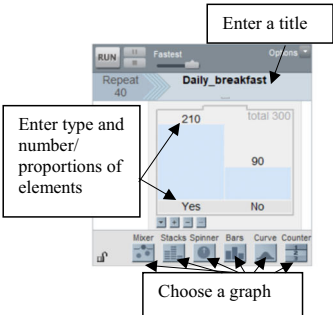
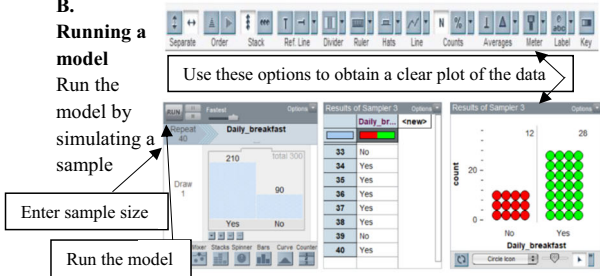
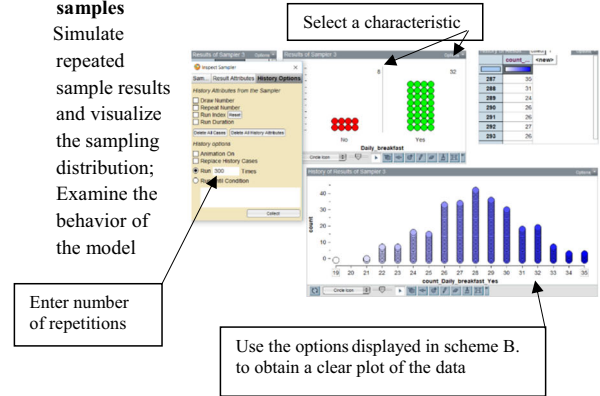
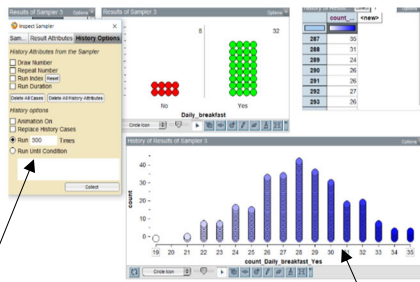
Scheme	TinkerPlots techniques	Conceptual Understanding
<p>A. Building a model Explore and identify important factors from a given real-life problem to build a population model in TP</p>		<p>A population model can be built by reducing/filtering the given data and using a suitable graphical representation (depending on the description/context given), for example a bar or pie chart, that incorporates the important factors, such as the total number of elements (population size) and the number or proportion of elements with a specific characteristic.</p>
<p>B. Running a model Run the model by simulating a sample</p>		<p>Collected results from a sample can be sorted/grouped and summarized in a suitable graph/visualization containing important sample features such as absolute and/or relative numbers for each characteristic.</p>
<p>C. Visualizing repeated samples Simulate repeated sample results and visualize the sampling distribution; Examine the behavior of the model</p>		<p>To answer a real-life problem, the behavior of the model with respect to a specific characteristic (related to the real-life data)—for example the number of students having daily breakfast—can be examined by simulating repeated samples and displaying the results in a sampling distribution. A sampling distribution displays the results of each sample in one bar chart (or stacked dot plot). Along the horizontal axis are the possible sample results displayed and along the vertical axis the frequency with which each result occurs; sample results from the same population model will vary due to chance, with results close to the population model occurring more frequently than strongly deviating sample results.</p>

Table 7 (continued)

D. Exploring repeated samples
 Simulate more repeated samples and interpret the results using the sampling distribution

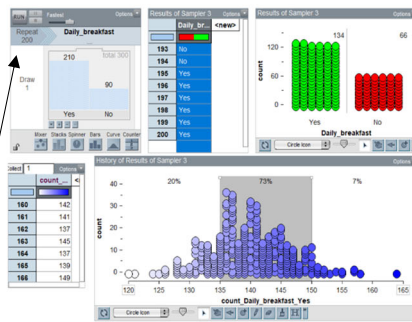


Enter a larger number of repetitions

Use the options displayed in scheme B to obtain a clear plot of the data

Results from a larger number of repeated samples lead to a better estimate of the population, that is, more repeated samples lead to (1) a smoother sampling distribution—without local peaks— (2) with a peak corresponding with the population proportion and (3) to an average that better resembles the population, and these three features ensure that results of more repeated samples lead to a better estimate of the population.

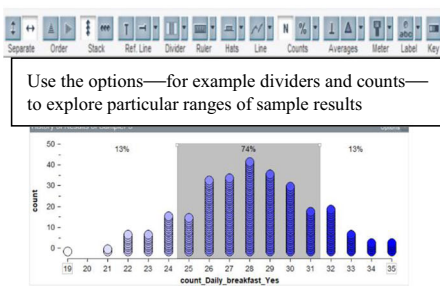
E. Exploring sample size
 Simulate repeated samples at varying sample sizes and interpret the results using the sampling distribution



Enter larger sample size and explore the new obtained results

Results from repeated samples with a larger sample size lead to a better estimate of the population, that is, a larger sample size leads to (1) less variation in the corresponding estimates of the population and (2) to an average that better resembles the population, and these two features ensure that a larger sample size leads to a better estimate of the population.

F. Interpreting sampling distribution
 Interpret the simulated results using the sampling distribution



Use the options—for example dividers and counts—to explore particular ranges of sample results

The proportion of sample results belonging to a particular range of results indicates the probability that a random sample result falls into this particular range. In this way, the probability of a given sample result can be interpreted and used to make a statement about a given real-life problem.

4.1 Identified instrumentation schemes

In observing students’ work, we identified six instrumentation schemes, called (A) building a model, (B) running a model, (C) visualizing repeated samples, (D) exploring repeated samples, (E) exploring sample size, and (F) interpreting sampling distribution (see Table 7). Column 1 provides a description of each scheme, column 2 displays a screenshot of students’ TinkerPlots

techniques for the scheme at stake, and column 3 shows students' understanding of statistical modeling that we could distill from their reasoning during these actions. Each instrumentation scheme incorporates a specific modeling process, ranging from building a population model by exploring and identifying important information in a given real-life problem in scheme A, to answering a given problem by interpreting the simulated sampling distribution from repeated sampling in scheme F. As such, the identified instrumentation schemes display how specific TinkerPlots techniques occurred simultaneously with particular elements in students' understanding of statistical modeling.

4.2 Students' global scheme development

We now describe students' instrumental genesis by presenting the observed levels at which the students used the TinkerPlots techniques and demonstrated their understanding of statistical modeling in their reasoning, throughout the teaching sequence. In each Worksheet (W1 to W5), instrumentation schemes A to F were addressed. Data from group 1 while working on W1 were used to indicate students' level in the initial phase of the teaching sequence, and data from group 2 while working on W5 for the more advanced phase. Students' technical actions with TinkerPlots were coded in technical levels for each scheme and for each student. For example, concerning W1 scheme A, five students out of fourteen in group 1 were unable to build a population model. They encountered difficulties in finding the input options for the parameters of the population model or for graphical representations of the model and, therefore, we coded their actions for W1 scheme A as technical level 1. As another example, concerning W5 scheme A, six students out of fourteen in group 2 were capable of making well thought out choices from newly explored TinkerPlots options to make a model, for example, by using non-instructed options for graphical representations like pie chart or histogram, and we coded their actions technical level 5.

Students' average technical levels for each instrumentation scheme on respectively W1 and W5 was calculated to identify students' global development during the work. For example, students' average technical level score on W1 in scheme A was calculated from five students whose actions were coded technical level 1, seven students who scored level 2, and two students on level 3, which resulted in an average technical level score of 1.8. Likewise, we coded students' reasoning and calculated their average conceptual levels for each scheme from A to F, while working on W1 and W5. The change in performance on technical and conceptual levels from the initial phase in W1 to the more advanced phase in W5 is visualized in Fig. 2. When comparing students' work on W1 to W5, students showed an improving level of proficiency in their application and control of the tool as well as in their usage and expression of statistical concepts in their accompanying reasoning. As students' development of TinkerPlots techniques and conceptual understanding of statistical modeling was observed simultaneously, the results show a co-development of techniques and concepts.

It is interesting to note that in schemes C and D, we observed more progress in students' average conceptual level score than for their technical level score. Both these schemes required more complex TinkerPlots techniques than the other schemes. In the initial phase, concerning these two schemes, most students of group 1 worked carefully according to the TinkerPlots instruction sheet, which enabled them to use the correct techniques. For example, concerning students' TinkerPlots techniques in scheme C during the initial phase with W1, all fourteen students had difficulty using the history option in TinkerPlots to visualize the sampling distribution. They all followed the instruction stepwise. Seven of them made mistakes in their

Co-Development of Techniques and Concepts for Statistical Modeling using TinkerPlots

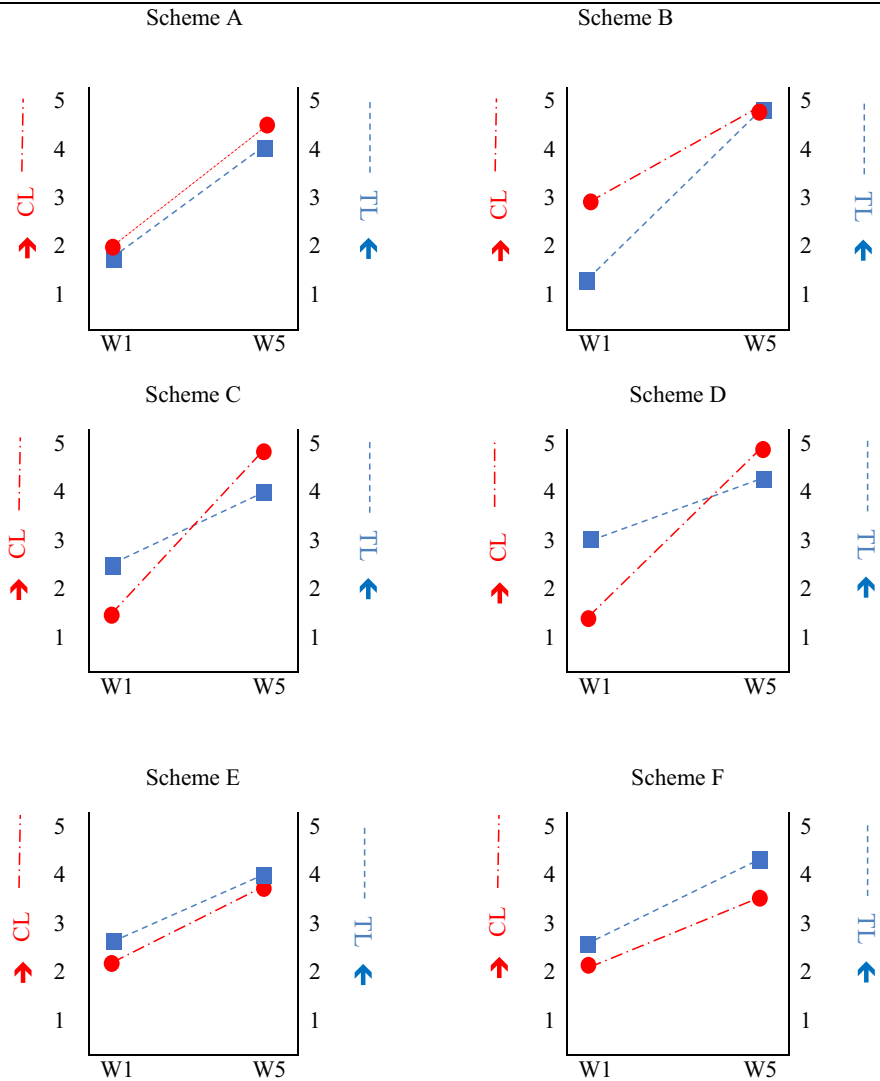


Fig. 2 Development of students’ average technical level (TL) and conceptual level (CL) for each instrumentation scheme from worksheets 1 to 5

actions—for example, not knowing how to enlarge the history window to enter all required information or not being able to select a useful characteristic for the history option—which made it difficult for them to visualize a correct sampling distribution, and as such, their actions were coded technical level 2. The other seven students made a correct visualization, although they encountered problems with displaying a clear bar chart or entering the correct values, and, as such, their actions were coded technical level 3. Students’ reasoning during the initial phase focused on the correct technical actions. For the seven students that were unable to visualize a correct sampling distribution, we also observed incorrect reasoning, that is, wrong statements

or incorrectly using words and symbols related to sample, variability, repeated samples, and sampling distribution, and we coded their reasoning conceptual level 1. In the data of five of the seven students that managed to display a correct sampling distribution, we observed superficial but correct reasoning, that is, noticing that the graph looks more or less the same as on the instruction sheet and reading the values on the horizontal axis for common sample result; as such, we coded their reasoning conceptual level 2. The two other students that visualized a correct sampling distribution were in one team. They discussed that the shape of the sampling distribution was not in line with their expectations, as they expected a smooth bell curve. Later in this section, we present in detail the work of these two students.

In the more advanced phase (W5), with regard to students' technical level in scheme C, all fourteen students in group 2 displayed a correct sampling distribution and were coded level 3 or higher. Four of them explored a quick start for simulating and adding repeated sample results to a sampling distribution; their actions were coded technical level 5. Concerning students' conceptual level in scheme C with W5, all students' reasoning was coded level 4 or higher, as they correctly stated that more repeated samples led to a smoother shape of the sampling distribution with a peak and average that resembled the modeled population proportion. For example, a student quoted:

This is in line with our expectations. Most of the sample results seem to be in between 43 and 47. This bar at 42 is a bit high [local peak], but yeah, it can happen that within these 100 repeated samples, there are incidentally more with 42...

With regard to the intertwinement of developing techniques and conceptual understanding, based on our findings in schemes C and D, it appeared that for schemes that required complex TinkerPlots techniques, a strong technical focus in the initial phase occurred together with less proficiency on a conceptual level, and, additionally, that in the more advanced phase within those schemes, students' statements shifted from discussing techniques to reasoning with concepts, which resulted in more progress for conceptual understanding. In schemes A, B, E, and F, we observed a more balanced co-development. Lastly, it is worth mentioning that in the advanced phase, most students (10 out of 14) were capable of using the simulated sampling distribution from repeated sampling as a model for determining the probability of a specific range of sample results, and, as such, to interpret the statistical model to solve a given problem.

4.3 Two cases

In this section, we present two cases of students as illustrative examples of how we zoomed in on the observed data to reveal the intertwining of emerging TinkerPlots techniques and conceptual understanding of statistical modeling in the personal schemes students develop. First, we present the case of Elisha and Willie (all student names are pseudonyms) while working on W1, as it illustrates how conceptual understanding influenced TinkerPlots techniques and vice versa in the initial phase. Second, we present the case of William and Brenda while working on W5 in the more advanced phase.

4.3.1 The case of Elisha and Willie

We focus on Elisha and Willie's work on W1 tasks 5 and 8 (Fig. 3). We start by highlighting some of their actions and reasoning, followed by an evaluation of how their developing TinkerPlots techniques influenced their conceptual understanding and vice

Breakfast worksheet

Introduction: Research at a primary school into the breakfast habits of pupils showed at the beginning of the school year that 210 of the 300 pupils eat breakfast daily. The school management wants to investigate again the number of students having daily breakfast at the end of the school year. However, asking all pupils is a lot of work and therefore they decide to take a sample of 30.

.....

Task 5: Assume that the number of pupils having daily breakfast remained the same during the school year. Use TinkerPlots to simulate sample results (number of pupils having daily breakfast) from the given population and sample size. Fill in the table below, based on the simulated results.

Sample size 30	Simulated sample results in interval notation [...:...]]
Most common results	
Exceptionally low results	
Exceptionally high results	

.....

Task 8: In the past school year, a lot of attention has been paid to stimulating 'daily breakfast' at the school. The school management wants to use the results from the sample of 30 to determine whether the breakfast behavior of pupils has improved. Suppose the sample shows that 23 out of 30 pupils have breakfast daily. Can the school management, based on this result, conclude that the pupils' breakfast behavior has improved? Support your answer with the simulated sample results.

Fig. 3 Tasks 5 and 8 from worksheet 1

versa. To answer W1 task 5, we expected students in scheme A of the instrumentation scheme to develop TinkerPlots techniques for entering the population characteristics as shown in Fig. 4(a). However, when entering the sample size in scheme B, Elisha and Willie incorrectly entered 100. Later on, when they arrived at scheme F—interpret the

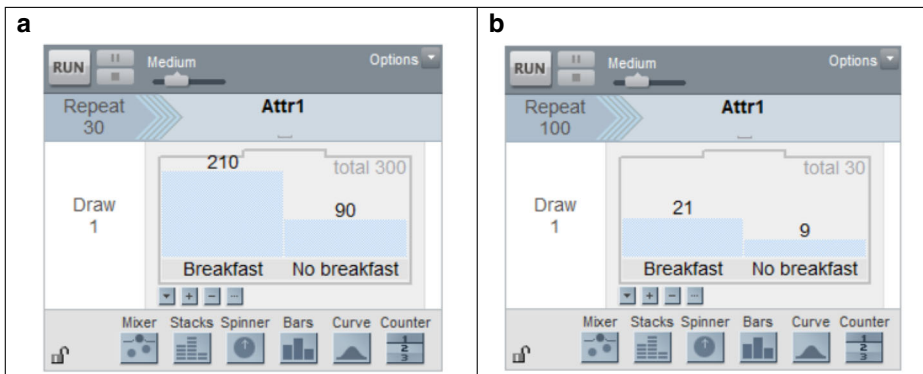


Fig. 4 Building a population model on worksheet 1 task 5: the expected model (a) and Willie and Elisha's model (b)

results using the sampling distribution—the following discussion in excerpt 1 took place while the two students were looking at the simulated sampling distribution on their screen (see Fig. 5(a)).

Willie: According to this graph, the sample results vary between 58 and 80 pupils who have breakfast every day [silence]. But... how is this possible? We only have 30 pupils in one sample.....

Elisha: Yes, but we have already filled in 100 [points to the input option ‘repeat’ on the screen, see Fig. 4(b)] and we should have entered 30.

Willie: But why, we do it [simulating repeated samples] 100 times, don’t we? We do it 100 times with 30 pupils.

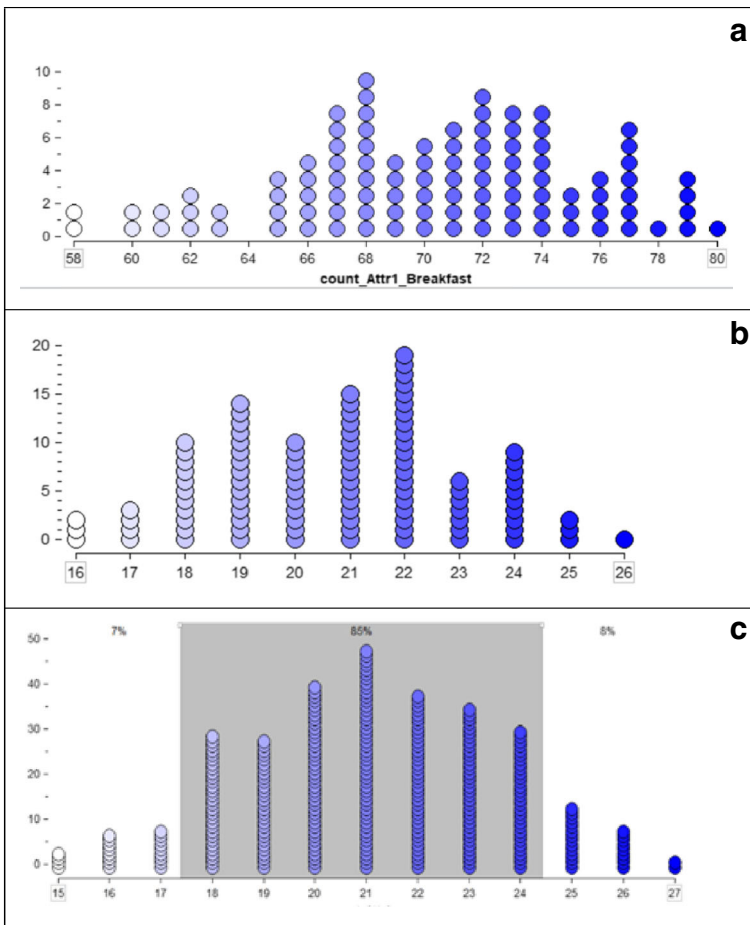


Fig. 5 Willie and Elisha’s simulated sampling distributions from repeated sampling for worksheet 1 task 5. (a) Simulated sampling distribution for worksheet 1 with an incorrect sample size of 100 instead of 30. (b) Simulated sampling distribution for worksheet 1 with sample size 30 and 100 repeated samples, showing a “bumpy” shape. (c) Simulated sampling distribution for worksheet 1 with sample size 30 and 300 repeated samples, showing a “smooth” shape

Elisha: Yes, exactly. We repeat it 100 times with 30 pupils. And now, we get for one such thing [points at the visualization of *one sample* on the screen] a result of 73 pupils who eat breakfast daily and 27 not, that is not correct. So, here [points again to the input option ‘repeat’ on the screen, see Fig. 4(b)], we should have entered the sample size, which is 30, instead of entering 100.

Following this discussion, they deleted their work and started again by entering a population model, but now with a correct sample size of 30. This time, in scheme C, they entered 100 for the number of repeated samples. This results in the simulated sampling distribution of Fig. 5(b). Here, the discussion in excerpt 2 took place.

Willie: This graph looks weird. What went wrong? Look at all those bumps.

Elisha: Let’s do it again [more repeated sampling]. And maybe, we should simulate more than 100 repeated samples. The more, the better, right?

Willie: [After simulating 200 extra repeated samples, their simulated sampling distribution looked like Fig. 5(c)]. Yes, that’s the way it should look like. Next time we just have to enter more repetitions right away. That’s simply the best. So, for now, most of the samples are between 18 and 24.

After the discussion in excerpt 2, they used the simulated sampling distribution (Fig. 5(c)) to correctly answer task 5 and 8. For task 5, they stated that most common sample results will vary from 18 to 24 out of 30; they indicated sample results varying from 15 to 17 out of 30 as exceptionally low and results varying from 25 to 27 exceptionally high. They ignored the possibility of sample results below 15 and above 27, probably as these results were not displayed on the *x*-axis of their simulated sampling distribution. For task 8, they stated that 23 out of 30 seemed to be better than 21 out of 30; however, 23 was not exceptionally high and, therefore, the school management could not conclude that the breakfast habits of pupils have improved. Elisha added that she regarded a sample of 30 as very small in this case.

In summary, the case of Elisha and Willie showed how their conceptual understanding and TinkerPlots techniques co-developed and influenced each other. From excerpt 1, it seems that they mixed up the option in TinkerPlots for entering sample size with entering the number of repeated sampling. When the (incorrect) simulated sampling distribution was displayed on their screen, this sampling distribution did not correspond to their conceptual expectations. The mismatch led them to investigate the options available to see what the problem was, which resulted in applying the correct technical option for entering sample size. Here, their conceptual understanding fostered their technical actions. From excerpt 2, we see how Elisha and Willie used the technical options for repeated sampling to get a better, less “bumpy and smoother” representation of the sampling distribution. The technique of increasing the number of repeated samples helped them understand the effect of more repeated samples by giving them a better picture of the sampling distribution. In this way, the technique of repeated sampling fostered their conceptual understanding of the effect of adding more repeated samples on the sampling distribution in scheme D.

From excerpt 2, it was difficult to distill the depth of the students’ conceptual understanding about adding more repeated samples in scheme D. Although they stated that they should enter a larger number of repetitions next time, and that a larger number of repetitions would lead to a better graph of the sampling distribution, they did not express clearly how they thought these two were related. However, later on, in W1 task 14, they explicitly mentioned that next time

they should simulate a larger number of repeated samples at once in order to reduce the influence of possible outliers and to achieve a well-shaped sampling distribution. Combining students' statements over several tasks and schemes helped us to identify their understanding of specific concepts.

4.3.2 The case of William and Brenda

We focus on the work of William and Brenda on W5 tasks 7 and 9 (Fig. 6). Instead of getting started with TinkerPlots after reading task 7, these two students started a 5-min discussion about possible answers. Excerpt 3 presents a small part.

William: 42 out of 50, that's not 90%, because then it should be 45, this is not enough. So don't buy it.

Brenda: I agree. 42 is not sufficient. Don't do it.

William: Or... (silence)... it's just a small sample size, only 50. In our earlier social media task with a sample of 50, there was a lot of variation, then 42 is not that unusual.

As the discussion progressed, they decided to model the task in TinkerPlots. Without discussing the TinkerPlots techniques, they succeeded within a few minutes and without any hesitations to display the sampling distribution as shown in Fig. 7(a). Their goal was to determine the most common results—in their strategy, the middle 80% of the samples—by placing borders for the lowest and highest 10%. After they moved the lower border of the gray area back and forth a number of times, it turned out to be impossible to get exactly 10% into the left part of the sampling distribution. On that point, the following discussion took place.

Worksheet 5: LED lights

Introduction: A do-it-yourself shop is not satisfied with the quality of LED lights they sell. There are too many complaints from customers about defective lights. They are therefore considering a switch to supplier B. This supplier guarantees that at least 90% of the lights will function. The do-it-yourself shop has therefore ordered a batch of 10,000 LED lights. Before selling them in the shop, they use a sample to verify whether the supplier's claim is correct.

... **Task 7:** Suppose there are 42 functioning LEDs in the sample (size 50). What advice would you give the shop about the purchase of the batch of LED lights? Justify your answer.

... **Task 9:** On closer inspection, the shop doubts whether a sample size of 50 would be appropriate. Which sample size would you recommend? Justify your answer.

... **Task 12:** Because the shop has doubts about the quality of the LED lights from supplier B, they ordered a batch of 10,000 from supplier C. They also examine this batch with a sample of 50 to determine whether this batch may be better than the batch of supplier B.

Open the file "LED lights supplier C" with TinkerPlots. Simulate repeated samples from the hidden population model of supplier C. Based on the simulations, estimate the number of functioning LEDs in the total batch of 10,000 LEDs. Justify your answer.

Fig. 6 Tasks 7, 9, and 12 from worksheet 5

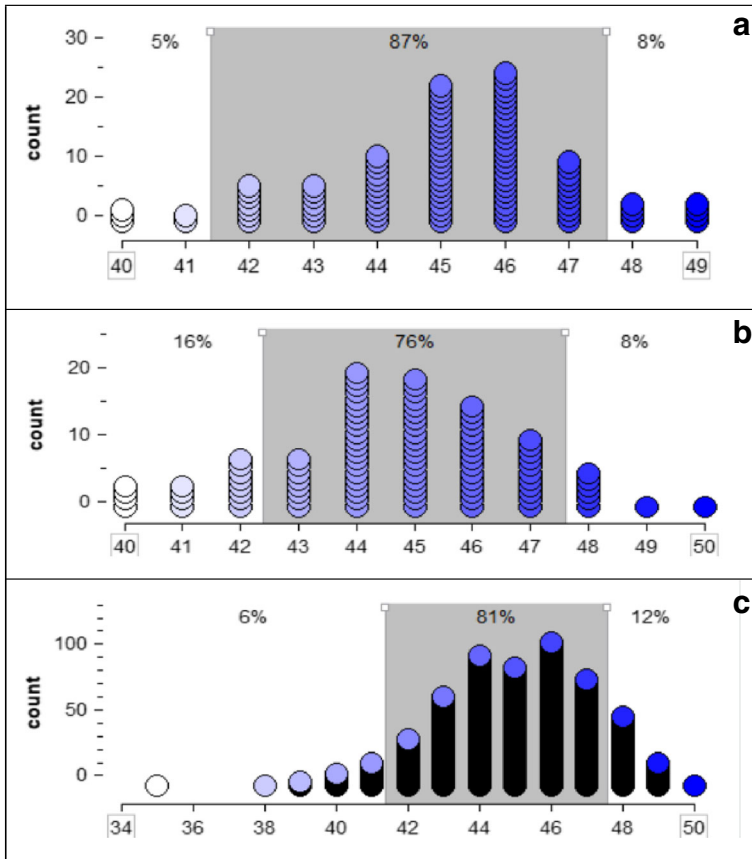


Fig. 7 William and Brenda’s simulated sampling distributions from repeated sampling for worksheet 5 task 7. (a) Simulated sampling distribution for 100 repeated samples in worksheet 5 with a left border of the gray area at 5%. (b) Simulated sampling distribution for 100 repeated samples in worksheet 5, second attempt, with a left border of the gray area at 16%. (c) Simulated sampling distribution for 500 repeated samples in worksheet 5 with a left border of the gray area at 5%

William: This is not a good sampling distribution.

Brenda: How is that?

William: It is not possible to get 10% here [pointing to the left part in the sampling distribution]. It is either 5% or 13%

Brenda: And now what? There’s not much we can do with this. Can’t we do it again? Then maybe it will be better.

They decided to delete everything and start all over again. This resulted in the sampling distribution of Fig. 7(b). Then the discussion in excerpt 5 took place.

William: This isn’t much better... now we have 8% or 16%...

Brenda: Let’s just do it again.

William: Again? Wait, I think we can do this again faster. We can leave this [points to sub screen 1, 2 and 3] and only have to do the repeats again.

[...]

- William: We should have discovered this earlier, that would have saved us a lot of work with the previous worksheets. In fact, we always investigate the same thing, but with a different subject.
- Brenda: How do you mean?
- William: Well, we investigate possible sample results with a given samples size to answer the questions. It doesn't really matter whether it's breakfast, social media or lights.

This third attempt also resulted in a left area smaller than 10%. At that moment, they decided to increase the number of repetitions, as that usually gives a better picture. After William said: "You can probably add samples in a quick way, without starting all over," they explored the techniques and soon found out how to add samples. This resulted in the sampling distribution of Fig. 7(c).

- William: I don't think there is any point in adding more repetitions, it remains the same. 42 is apparently exactly at the border of common results. And now what?
- Brenda: The sampling distribution hardly changes, so there's no need for more repetition. I think 42 is not much. Most results are higher
- William: Okay, based on these sampling distributions we find 42 to be too few. So our advice is not to buy!

Later on, when they worked on W5 task 9, they fully agreed that the sample size was too small. Brenda stated: "The larger sample the better the results, but very large is not convenient," at which point William proposed to pick a sample size of 200. As with task 7, they wanted to explore a fast way in which they did not have to remove all the sub-screens. To this end, they discussed the views on each sub-screen and finally decided that only sub-screen 1 could remain. Here, they discussed concepts such as sample size, difference between sample size and number of repeated samples, and the relationship between the tables and dot plots. When using a fast method for larger sample size, the effect of larger sample size confirmed their conjecture.

In summary, the case of William and Brenda in the more advanced phase showed a focus on conceptual understanding when reading the task, a focus that we saw in almost all students in W5. Excerpt 4 illustrates how the two students, after reading task 7, discussed concepts such as variation, probability, and sample size. Moreover, in this excerpt, they related this task to a previous task and context (the context of W3). This also appeared in the second part of excerpt 5; here, we saw how the use of similar TinkerPlots techniques in different worksheets and contexts enabled William to discover a context-independent pattern. The TinkerPlots techniques helped him to identify technical patterns in the modeling process and thus to view the concepts involved at a more abstract—context-independent—level. Regarding the intertwined relationship between TinkerPlots techniques and conceptual understanding, excerpt 5 showed how their understanding—in this case, their overestimation of variation in many repeated samples—triggered them to explore new techniques. Also, the other way around, how in excerpt 6 the techniques helped them to understand that the sampling distribution of many repeated samples remains stable. Also, their work on W5 task 9 showed, as in the case of Willie and Elisha, a two-directional relationship between TinkerPlots techniques and conceptual understanding. Their understanding of statistical modeling concerning a general approach and patterns resulted in a search for more advanced TinkerPlots techniques by using already

modeled parts of the process in their sub-screens, and also, the techniques strengthened them in their conjecture about the effect of sample size.

5 Discussion

The aim of this study was to explore the applicability of the instrumental genesis perspective to statistics education, and to statistical modeling in particular. We identified six instrumentation schemes for statistical modeling processes with TinkerPlots, describing the intertwined development of students' digital techniques and conceptual understanding. We noticed an increase in their mastery of the tool as well as in their statistical reasoning, evidencing students' co-development of techniques and conceptual understanding. We observed a two-directional intertwining of techniques and concepts. The two student cases showed in more detail how students' understanding of concepts informed their TinkerPlots techniques and vice versa. Although we found a two-directional intertwining in all schemes and phases of the teaching sequence, at particular moments we noticed more emphasis in one direction.

In the more advanced phase of the teaching sequence, the results show how the use of similar TinkerPlots techniques over different worksheets and contexts enabled students to discover context-independent technical patterns. Students' identification of those technical patterns in the modeling process enabled them to view concepts at a higher, more abstract level. We interpret this as emergent modeling (Gravemeijer, 1999), which involves the conceptual shift from a model of a context-specific situation to a model for statistical reasoning in a variety of similar and new contexts. Although the emphasis here was on technical patterns that informed students' conceptual understanding, we also saw the opposite direction intertwined in this process, as their conceptual understanding concerning a general approach and patterns resulted in a search for more advanced techniques by using already modeled parts of the process on their screen.

In a short period of time, students—who were inexperienced in taking samples and working with digital tools—learned to carry out the modeling processes, including interpreting the simulated sampling distribution. Regarding this promising result, we discuss some possible stimulating factors. As a first factor, it appeared from the identified instrumentation schemes that the required techniques in the digital environment of TinkerPlots strongly align with key concepts for statistical modeling. This strong alignment probably facilitated students to overcome initial difficulties concerning variability, distribution, sample and sampling, the effect of sample size, difference between the sample and sampling distribution (Chance et al., 2004), which we hardly observed in the more advanced phase. For example, concerning the common confusion between the sample and sampling distribution, the distinct visualization of sample and sampling results within the digital environment of TinkerPlots enabled students to distinguish between both distributions. As a second factor, the required TinkerPlots techniques invited students to phenomenon simplification (Manor & Ben-Zvi, 2017). For example, in the initial phase, we observed difficulties in distilling sample size and population proportion from the context given for entering the correct model, while these difficulties hardly occurred in the more advanced phase. As a third factor, applying similar statistical modeling processes in TinkerPlots to varying real-life contexts allowed students to distinguish and interact between the model world—using the same digital environment—and the real world using varying contexts (Patel & Pfannkuch, 2018).

The findings presented in this paper should be interpreted in the light of the study's limitations. First, the results of this research are based on students in a classroom laboratory instead of students' regular classroom environment. By conducting the preparatory activities in students' regular classrooms and maintaining the same student teams, lesson design, and a familiar teacher, we tried to reduce the influence of the classroom laboratory setting at the university. Second, due to practical reasons, we were confined to working with two groups of students, group 1 in the initial phase and group 2 in the more advanced phase of the teaching sequence. Differences between both groups may have affected students' global scheme development. However, the students in both groups performed at a similar level in mathematics and their performances in the preparatory tasks were comparable. The teacher judged the starting level of the two groups to be similar. Third, distilling students' conceptual understanding from their reasoning was challenging. However, by combining the sometimes flawed statements made by the students with their accompanying activities—such as their next action with the tool or their statements later on in their process—we tried to identify their understanding of the concepts. Fourth, we worked with pre-university students, the top 15% achievers in our educational system. Other students may need more time.

Although we focused on statistical modeling processes using TinkerPlots, we consider our findings on the intertwining of emerging digital techniques and conceptual understanding applicable to the broader field of statistics education, and to other educational digital tools as well. Digital tools for other areas in statistics education also structure and guide students' thinking by providing specific options for entering parameters and commands and/or by facilitating explorative options that may strengthen students' conceptual understanding.

Overall, we conclude that the perspective of instrumental genesis in this study proved helpful to gain insight into students' learning from and with a digital tool, and to identify how emerging digital techniques and conceptual understanding intertwine.

6 Implications

The study's results lead to implications for the design of teaching materials and digital tools, and for future research. In designing teaching materials, it is important to take into account the two-directional relationship between emerging digital techniques and conceptual understanding, both during instruction and during practice. Attention to digital techniques in the initial phase, especially to more complex ones, seems to have a positive effect on learning the associated concepts later on. The development of context-independent techniques and concepts requires sufficient time and practice for students with different contexts and situations. In designing digital tools, the intertwined relationship between digital techniques and conceptual understanding calls for attentive consideration of how the digital techniques are related to the concepts, to deploy the digital tool in a productive way for the intended learning goal.

This also suggests an implication for future research on statistics education using digital tools. Although we focused on statistical modeling using TinkerPlots—that is, solving real-life problems by the building, application, and interpretation, of the sampling distribution of repeated samples—we assume our global findings also hold for other statistical processes and digital tools. However, the specific intertwining of emerging digital techniques and conceptual understanding is unique for each digital tool and intended learning goal. To identify the specific intertwinement, we recommend using the perspective of instrumental genesis in analyzing video and conversation data, which can be added by using clinical interviews.

On a final note, this study gave an insight into the applicability of the instrumental genesis perspective in the context of statistics education, and statistical modeling with digital tools in particular. Instrumental genesis seems a fruitful perspective to design technology-rich activities and to monitor students' learning.

Funding This research received funding from the Ministry of Education, Culture and Science under the Dudoc program.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research*, *54*(2), 139–172.
- Artigue, M. (2002). Learning mathematics in a CAS environment: The genesis of a reflection about instrumentation and the dialectics between technical and conceptual work. *International Journal of Computers for Mathematical Learning*, *7*(3), 245–274.
- Biehler, R., Frischemeier, D., & Podworny, S. (2017). Editorial: Reasoning about models and modeling in the context of informal statistical inference. *Statistics Education Research Journal*, *16*(2), 8–12.
- Büscher, C., & Schnell, S. (2017). Students' emergent modeling of statistical measures—A case study. *Statistics Education Research Journal*, *16*(2), 144–162.
- Chance, B., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 295–323). Dordrecht, the Netherlands: Kluwer Academic.
- Davies, R. (2011). Understanding technology literacy: A framework for evaluating educational technology integration. *TechTrends*, *55*(5), 45–52.
- Drijvers, P., Godino, J. D., Font, D., & Trouche, L. (2013). One episode, two lenses. A reflective analysis of student learning with computer algebra from instrumental and onto-semiotic perspectives. *Educational Studies in Mathematics*, *82*(1), 23–49.
- Gal, I. (2002). Adults' statistical literacy, meanings, components, responsibility. *International Statistical Review*, *70*(1), 1–51.
- Garfield, J., delMas, R., & Chance, B. (1999). Developing statistical reasoning about sampling distributions. In *Presented at the First International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL)*. Kibbutz Be'eri: Israel.
- Garfield, J., delMas, R., & Zieffler, A. (2012). Developing statistical modelers and thinking in an introductory, tertiary-level statistics course. *ZDM The International Journal on Mathematics Education*, *44*(7), 883–898.
- Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning*, *1*, 155–177.
- Groth, R. E. (2015). Working at the boundaries of mathematics education and statistics education communities of practice. *Journal for Research in Mathematics Education*, *46*(1), 4–16.
- Konold, C., Harradine, A., & Kazak, S. (2007). Understanding distributions by modeling them. *International Journal of Computers for Mathematical Learning*, *12*(3), 217–230.
- Manor, H., & Ben-Zvi, D. (2017). Students' emergent articulations of statistical models and modeling in making informal statistical inferences. *Statistics Education Research Journal*, *16*(2), 116–143.
- Patel, A., & Pfannkuch, M. (2018). Developing a statistical modeling framework to characterize year 7 students' reasoning. *ZDM-Mathematics Education*, *50*(7), 1197–1212.

- Pfannkuch, M., Ben-Zvi, D., & Budgett, S. (2018). Innovations in statistical modeling to connect data, chance and context. *ZDM-Mathematics Education*, 50(7), 1113–1123.
- Podworny, S., & Biehler, R. (2014). A learning trajectory on hypothesis testing with TinkerPlots – Design and exploratory evaluation. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education (Proceedings of the Ninth International Conference on Teaching Statistics, Flagstaff, USA)*. Voorburg, the Netherlands: International Association for Statistical Education and the International Statistical Institute.
- Simon, M. A. (2019). Analyzing qualitative data in mathematics education. In K. R. Leatham (Ed.), *Designing, Conducting, and Publishing Quality Research in Mathematics Education* (pp. 111–123). Cham, Switzerland: Springer.
- Thijs, A., Fisser, P., & Van der Hoeven, M. (2014). *21e-eeuwse vaardigheden in het curriculum van het funderend onderwijs [21st century skills in the curriculum of foundational education]*. Enschede, the Netherlands: SLO.
- Van Dijke-Droogers, M. J. S., Drijvers, P. H. M., & Bakker, A. (2020). Repeated sampling with a black box to make informal statistical inference accessible. *Mathematical Thinking and Learning*, 22(2), 116–138.
- Van Someren, M. W., Barnard, R., & Sandberg, J. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London, UK: Academic Press.
- Vergnaud, G. (1996). Au fond de l'apprentissage, la conceptualisation. In R. Noirfalise & M.-J. Perrin (Eds.), *Actes de l'école d'été de didactique des mathématiques* (pp. 174–185). Clermont-Ferrand, France: IREM.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.