



HAL
open science

Design principles for short informal statistical inferences activities for primary education

Arjen de Vetten, Aisling Leavy, Ronald Keijzer

► To cite this version:

Arjen de Vetten, Aisling Leavy, Ronald Keijzer. Design principles for short informal statistical inferences activities for primary education. Eleventh Congress of the European Society for Research in Mathematics Education (CERME11), Utrecht University, Feb 2019, Utrecht, Netherlands. hal-02411579

HAL Id: hal-02411579

<https://hal.science/hal-02411579>

Submitted on 15 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Design principles for short informal statistical inferences activities for primary education

Arjen de Vetten¹, Aisling Leavy² and Ronald Keijzer³

¹Leiden University, Leiden, The Netherlands; a.j.de.vetten@fsw.leidenuniv.nl

²Mary Immaculate College, University of Limerick, Limerick, Ireland; aisling.leavy@mic.ul.ie

³University of Applied Sciences iPabo, Amsterdam, The Netherlands, r.keijzer@ipabo.nl

This paper presents the following set of eight design principles for informal statistics inference (ISI) education for upper primary education in settings where limited time is available for ISI: Create awareness of inferential claims before engaging students in ISI via statistical investigations; use a context for which the students have no clear expectations of the outcome or which challenges current beliefs; use a question that highlights the need for an inferential claim; use a large population that makes sampling inevitable; use activities with tangible populations and samples in which generating representative samples and recording data is easy and quick; use data that require little descriptive analyses; and use the principles of growing samples and repeated sampling in order to create sampling variability. Three exemplar ISI activities illustrate how the design principles can be used to design new ISI activities and evaluate existing ISI activities.

Keywords: Primary education, informal inferential reasoning, informal statistical inference, statistics education.

Introduction

An introduction to informal statistical inference (ISI) in primary education has been advocated by various researchers (Makar, 2016; Makar, Bakker, & Ben-Zvi, 2011). The literature provides examples of activities that introduce ISI in upper primary education (De Vetten, Schoonenboom, Keijzer, & Van Oers, 2018a; Kazak, Pratt, & Gökce, 2018; Leavy, 2009). While design principles for statistics education in general are available (Ben-Zvi, Gravemeijer, & Ainley, 2018), the literature does not report on a comprehensive list of design principles for ISI education for upper primary education. The aim of this paper is to bring together a set of design principles that have been reported in the literature alongside design principles that are distilled from recent research with primary pre-service teachers (De Vetten et al., 2018a; De Vetten, Schoonenboom, Keijzer, & Van Oers, 2019) with a view to generate design principles that apply to upper primary education. The design principles presented in this paper can be used for the design of ISI activities that require relatively little time (e.g. up to an hour), but still allow for an introduction of all key concepts of ISI.

Theoretical background

ISI can be seen as part of the broader field of statistics. For statistics education in general, it is suggested to focus on developing central statistical ideas, to use real, or realistic, and motivating data, and to integrate the use of technological tools that allow students to explore and analyze data (Ben-Zvi et al., 2018). Most of these principles are useful also for ISI education in primary

education. The design principles we describe below therefore build on these principles. However, for eliciting IIR, these principles are not sufficiently specific.

A number of design principles have been targeted specifically towards ISI. First, activities where students are asked to make predictions about another or a larger sample based on the sample at hand, such as the growing samples heuristic (Bakker, 2004), appear to be an effective principle for activities that aim to engage students in IIR (Ben-Zvi et al., 2018; Makar, 2016). Some recent research involving pre-service teachers (De Vetten et al., 2019) and upper primary school students (Kazak et al., 2018) utilizing the growing sample design principle indicates mixed results. The participants used the data as evidence in making predictions of the shape of the distributions of larger samples and thus looked beyond the data. However, pre-service teachers also tended to neglect uncertainty and make deterministic predictions. Moreover, most predictions made of other samples mimic the shape of the original sample, thus suggesting that students did not have a picture in their mind of a sample distribution being the result of an overarching population distribution. Rather, they may have thought that the original sample “produces” the predicted sample. De Vetten et al. (2019) suggest that this could be due to the absence of repeated sampling elements in the activities used. Including such elements confronts students with different sample results, which may make them aware that an inference based on one sample includes some uncertainty.

A second design principle for ISI activities is to have students compare two sample distributions and draw a conclusion about the difference between the means in the population (Zieffler, Garfield, delMas, & Reading, 2008). However, research evidence shows mixed results in terms of the success of eliciting IIR, due to a number of (contextual) factors. First, without proper teacher intervention, such tasks appear not to sufficiently stimulate students to engage in IIR (De Vetten et al., 2019; De Vetten, Schoonenboom, Keijzer, & Van Oers, in press; Pfannkuch, 2006), and, in our research, pre-service teachers tend to restrict their attention to descriptive analyses. Another factor that may contribute to the absence of IIR in activities where students compare two distributions (De Vetten et al., 2019, in press; Pfannkuch, 2006) was that these activities did not involve repeated sampling, thereby missing the opportunity to view sampling variability. Third, while the sample distributions are visible, the populations remain abstract. Finally, tasks that compare sample distributions require substantial time for the necessary descriptive analyses, which goes at the expense of the available time and energy for inferential reasoning (De Vetten et al., 2018a).

A third design principle is to use a modelling approach in which a population is modelled (often using software, such as Tinkerplots). Samples are drawn from this hypothetical population distribution in order to develop an appreciation for the likelihood of various sample outcomes and an understanding of sampling variability (Manor & Ben-Zvi, in press). This approach requires more time and puts demands on the computer skills of the teacher and the student, and may therefore be less suitable for general upper primary students in settings with limited time and support.

In sum, findings from recent studies points towards a number of conclusions regarding effective design principles for ISI education and requirements for effective design principles: First, the use of growing samples and repeated sampling is recommended in order to create sampling variability and to show that usually larger samples reduce sampling variability and thus reduce the uncertainty of

the inference. Second, it appears to be important to make students aware of the distinction between sample and population and when an activity requires a generalization beyond the data. Third, an overemphasis on descriptive analyses, arising from activities involving the comparison of sample distributions, may reduce the opportunity to engage in IIR. Fourth, in many primary school settings with limited available time and support the extensive use of technological tools may not be feasible.

Design principles

Using the above conclusions about (requirements for) effective design principles for statistics education in general and ISI education in particular we formulated eight principles for the design of ISI activities for primary education.

In particular, seven of these design principles are specifications of how statistical investigations activities should look like when the goal is to introduce students to ISI. The first design principle stands apart as it is undertaken prior to students being actively involved in conducting statistical investigations.

1. Create awareness of inferential claims and of the distinction between sample and population, before engaging students in ISI via statistical investigations

De Vetten et al. (2019) and De Vetten et al. (in press) showed that most pre-service teachers appear to only describe the data, rather than engage in IIR. In De Vetten et al. (2018a) we first engaged pre-service teachers in an activity that was designed to make them aware of the existence and use of inferential claims. This activity appeared to promote an awareness of inference. Throughout the intervention the participants retained this awareness.

2. Use a context for which the students have no clear expectations of the outcome, or for which the outcome challenges students' current beliefs.

When the outcome is unpredictable or challenges students' current beliefs, students are stimulated to engage in inferential investigations and to search for explanations (Makar et al., 2011). When the outcome is predictable students may not engage in critical IIR (see for an example De Vetten et al., 2018a).

3. Use a question that clearly highlights that an inferential claim is required and use a population with a sufficiently large number of elements that surveying the entire population is impossible within a limited time frame.

The question posed should lead the student to conclude that merely analyzing the sample data descriptively is insufficient. However, even when considerable care is taken to formulate a good inferential question (e.g. using the guidelines from Pfannkuch, Regan, Wild, & Horton, 2010), other factors may cause students to restrict their attention to descriptive analyses (De Vetten et al., 2019). Therefore, combining a tangible population (see principle 6) with a population that consists of too many elements to survey in its entirety makes drawing a sample natural and generalization inevitable (De Vetten et al., 2018a).

4. Use activities that support the generation of a representative sample and that stimulate critical reflection on the representativeness of the sample.

The effect of this principle is that students may be willing to accept the sample data as reliable evidence for making inferential claims. In case an ISI activity does not permit students to draw representative samples, but only non-representative ones (such as convenience samples from their own class), students may be reluctant to engage in IIR because generalization based on such non-representative samples is not possible anyway. Ideally, the activity is such that it is possible to draw both representative and non-representative samples, with the effect that discussions are elicited about ways to draw a representative sample.

5. Use activities in which generating samples and recording data is easy and quick and in which students preferably collect data themselves.

The effect of this design principle is that an excessive focus on data collection is avoided. At the same time, the data generation and recording process should allow students to make sense of the data and should facilitate mathematization (Treffers, 1987) and engagement in IIR. Therefore, it is preferable that students collect their own data.

6. Use activities with a tangible population and sample.

The effect of this design principle is that students remain aware of the distinction between sample and population and that they are required to make a claim that pertains to the entire population.

7. Use data that require little descriptive analyses.

In the PPDAC-cycle, data analyses comes before interpretation of the results, including making inferences (Wild & Pfannkuch, 1999). Spending too much time on descriptive analyses of the sample data leaves little time and energy for inferential reasoning (De Vetten et al., 2019; Leavy, 2009). The effect of this design principle may be that sufficient time, energy and attention is left for IIR, while at the same time engagement in descriptive analyses helps students to make sense of the results and facilitates further mathematization (De Vetten et al., 2018a).

8. Use the principles of growing samples and repeated sampling in order to create sampling variability.

The effect of using the principle of repeated sampling is that students will consider the uncertainty of their inferences, because sample results will vary (Saldanha & Thompson, 2002). The effect of using the principle of growing samples is that students will see that larger samples usually reduce sampling variability and thus reduce the uncertainty of the inference (Bakker, 2004).

Applying the ISI design principles – Three example activities

This section describes three ISI activities and how the design principles can be used to design new ISI activities and evaluate existing ISI activities. The first activity was designed with the design principles in mind, while the second and third were not. We evaluate the extent to which they are in line with the design principles.

Activity 1: What is the most frequently used word?

The first activity (De Vetten et al., 2018a; De Vetten, Schoonenboom, Keijzer, & Van Oers, 2018b) was implemented in two settings: in a class of second-year pre-service primary school teachers and

in three primary classrooms (grade 3, 5 and 6). The activity centers round a pile of children’s novels, that constitutes a population, and the question “Which word is most frequently used in this pile of novels?” First, students make a list of the top 5 words they think are most frequently used in the pile of books. Next, students discuss how to find out which is the most frequently used word. Taking a sample is the expected response. The sampling procedure is discussed. Small groups of students draw small samples and then pool their results in multiple large samples. The inferences for small and large samples are compared. In this way, it could be shown that the proportion that yields the same most frequently used word is greater for large samples than for small samples. If the proportion of samples with the same most frequently used word is high, students could be willing to accept the possibility of making uncertain inferences based on one sample.

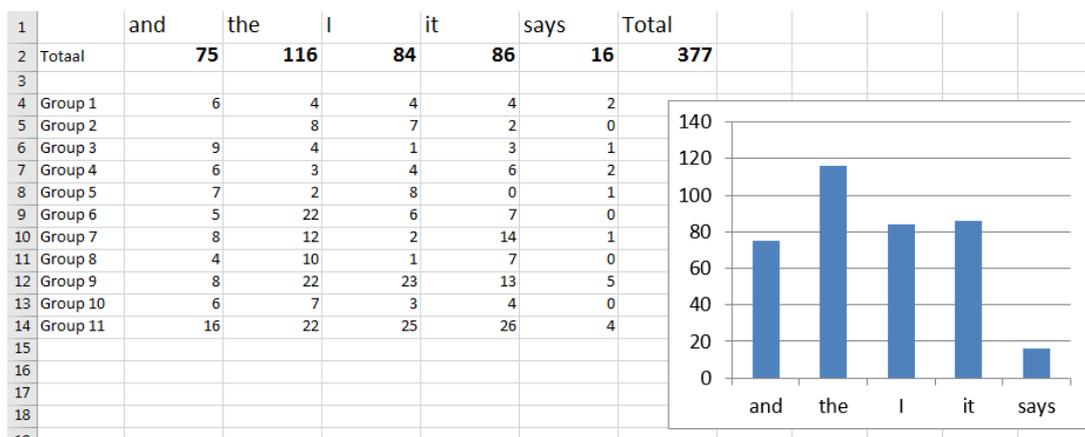


Figure 1: Pooled sample results in the activity “What is the most frequently used word?”

This activity was designed with the design principles in mind and it aligns well with all but one of the design principles. The activity is clearly inferential and uses a tangible and large population and visible sample (design principles [DP] 3 and 6). It requires that students consider how to take a representative sample and it allows for generating a representative sample (DP 4). Students take a book from the pile and count the occurrence of the top 5 words on a number of lines or a page, making sampling, recording and analyzing simple (DP 5 and 7). Without the need to rely on technological tools growing samples and repeated sampling can be achieved by having individual students draw samples and by combining students’ samples (DP 8). The number of students may, however, limit the possibilities for repeated sampling for larger samples. Our research shows that pre-service and primary school students were engaged in inferential reasoning (De Vetten et al., 2018a, 2018b). However, the activity included repeated sampling only for smaller samples. It thus missed the opportunity for students to see the sample results converge for larger samples. Repeated sampling for various sample sizes appears crucial to teach students to balance between sampling representativeness and sampling variability. Finally, the activity might not completely align with DP 2, because students may have beliefs about the outcome, based on information they have about word frequency in books in general (De Vetten et al., 2018a).

Activity 2: How many yellow balls are in the Black Box?

The activity “How many yellow balls are in the Black box?” (Van Dijke, Drijvers, & Bakker, 2018) was implemented in a grade 9 pre-university secondary classroom. The activity has students

investigate and estimate the number of yellow balls in a black box, filled with a mix of 1,000 yellow and orange balls, by looking through a small viewing window. It is expected that students will shake the bottle in order to take samples repeatedly and that they will use the average of their sample as an estimate. Findings are recorded and discussed by students as part of a whole class discussion. Following this, a larger viewing window is opened on the black bottle, the experiment repeated, and findings recorded. The contention is that experiences with the physical black box experiment support the development of an informal understanding of samples and sampling variability and promote investigation of the effect of repeated samples and larger sample size.

While this activity was not designed based on the list of design principles presented in this paper and its effects on students' reasoning has not yet been reported, it appears to align well with the principles. As the number of yellow balls is determined by the educational designer, students cannot have any expectations of the outcome (DP2). Generalization is natural, as it is obvious that the sample results are only a means to make a claim about the population (DP 3). It has a visible and large population and a visible sample (DP 3 and 6). It is easy to obtain a representative sample as yellow and orange balls are mixed and shaking the bottle results in a random sample. Since obtaining a non-representative sample appears to be hard in this activity, it might not stimulate critical reflection on sample representativeness. However, sampling variation may lead to discussing representativeness (DP 4). The context, question, sampling, recording and analyzing are simple (DP 5 and 7). Repeated sampling is achieved by repeatedly shaking the bottle and counting the yellow balls and by combining small group results in a whole class graph, while growing samples is possible by enlarging the viewing window from 40 to 80 balls (DP 8).



Figure 2: Black box, used in the activity “How many yellow balls are in the Black Box?”

Activity 3: Comparing math attitude of men and women

This activity is a standard activity in which students are asked to compare two samples and draw a conclusion whether in the population the groups differ (see for example De Vetten et al., 2019; De Vetten et al., in press). One example of such an activity is shown below, where the activity also involves a growing samples activity. It was implemented and evaluated in three groups of first-year pre-service primary school teachers (De Vetten et al., 2019). The driving question was whether in the population there is a difference in math attitude between men and women. During three rounds the students answer this question and make predictions of the shape of the distributions of larger samples.

This activity was not well-aligned with a number of design principles, in particular with the visibility of the population (DP 6), the simplicity of the analyses (DP 7), and the possibilities for repeated sampling (DP 8). Indeed, we showed among the pre-service teachers engaged in this activity, there was a strong tendency to restrict attention to descriptive analyses, rather than to engage in inferential reasoning (De Vetten et al., 2019). The teacher also had strong ideas of the difference between the groups in the population and, therefore, some had problems treating the data as evidence (DP 2). We also contend that, when the population elements are humans, it may be complex to obtain a representative sample, thus leading students to put less trust in the sample and reducing their willingness to make generalizations (DP 4 and 5). On the positive side, the complexity of the sampling process stimulates critical reflection on sample representativeness (DP 5) and the immensity of the population encourages students to take a sample (DP 3). Although the design principles offer suggestions to improve this activity, in particular using repeated sampling, some elements, such as a visible population from which it is easy to generate a representative sample, might not easily be rectified.

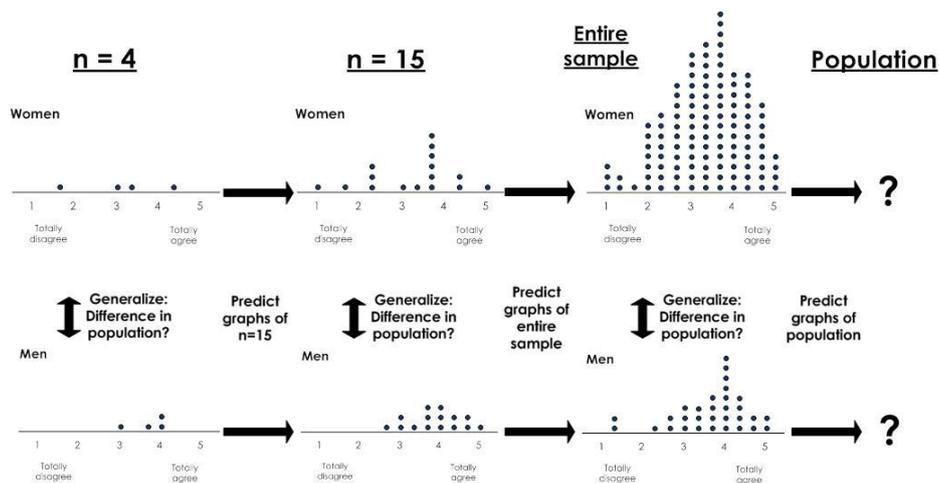


Figure 3: Overview of the activity “Comparing math attitude of men and women”

Conclusion

Designing ISI activities for primary education that align with all the design principles is not straightforward, as it demands very good knowledge of ISI and educational design. We have suggested two activities that to a large extent align with the design principles. Teachers can effectively use these activities in their classrooms. The design principles can also be used by educational designers to align their own activities with the design principles and elicit IIR more effectively. For example, the activity in which two samples are compared could incorporate repeated sampling, and the averages and measure(s) of dispersion could be provided, so that more time can be spent on IIR. Some of the design principles for ISI education presented in this paper are well-known and have been reported before, but some are based on recent ISI education research. In this way, we aim to provide a comprehensive set of design principles, based on recent evidence, that can be useful for educational designers and are ready to be corroborated by future research. Moreover, the principles are potentially applicable beyond the primary education (teacher) setting, to other settings where the design for activities is necessary when first introducing students to ISI.

References

- Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht, the Netherlands: CD-β Press.
- 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* (pp. 473–502). Cham, Switzerland: Springer.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2018a). The development of informal statistical inference content knowledge of pre-service primary school teachers during a teacher college intervention. *Educational Studies in Mathematics*, 99(2), 217–234. doi:10.1007/s10649-018-9823-6
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2018b). *Knowledge in action. The informal statistical inference content knowledge of three pre-service teachers while teaching ISI in upper-primary school*. Manuscript in preparation.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2019). Pre-service teachers and informal statistical inference: Exploring their reasoning during a growing samples activity. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research: International perspectives*. New York, NY: Springer.
- De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (in press). Pre-service primary school teachers' knowledge of informal statistical inference. *Journal of Mathematics Teacher Education*. doi:10.1007/s10857-018-9403-9
- Kazak, S., Pratt, D., & Gökce, R. (2018). Sixth grade students' emerging practices of data modelling. *ZDM*. doi:10.1007/s11858-018-0988-3
- Leavy, A. M. (2009). *Insights into Informal Inferential Reasoning in the Primary Classroom*. Paper presented at the 57th session of the International Statistical Institute, Durban, South Africa.
- Makar, K. (2016). Developing young children's emergent inferential practices in statistics. *Mathematical Thinking and Learning*, 18(1), 1–24.
- Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning*, 13(1-2), 152–173.
- Manor, H., & Ben-Zvi, D. (in press). Design for reasoning with uncertainty in informal statistical inference. In G. Burrill & D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research: international perspectives*. New York, NY: Springer.
- Pfannkuch, M. (2006). *Informal inferential reasoning*. Paper presented at the Seventh International Conference on Teaching Statistics, Salvador, Brazil.
- Pfannkuch, M., Regan, M., Wild, C., & Horton, N. J. (2010). Telling data stories: Essential dialogues for comparative reasoning. *Journal of statistics Education*, 18(1), 1–38.
- Saldanha, L., & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51(3), 257–270.
- Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics: The Wiskobas project*. Dordrecht, The Netherlands: Reidel.
- Van Dijke, M., Drijvers, P., & Bakker, A. (2018). *Repeated sampling as a step towards informal statistical inference*. Paper presented at the Tenth International Conference on Teaching Statistics (ICOTS10, July 2018), Kyoto, Japan.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–248.
- Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40–58.