Journal of Mathematical Behavior 56 (2019) 100723

Contents lists available at ScienceDirect

Journal of Mathematical Behavior

journal homepage: www.elsevier.com/locate/jmathb

Lower secondary school students' reasoning about compound probability in spinner tasks

Kjærand Iversen[a,*], Per Nilsson[a,b]

[a] The Faculty of Education and Arts, North University, Norway
[b] Örebro University, School of Science and Technology, Sweden

ARTICLE INFO

Keywords:
Compound stochastic events
Reasoning
Sample-space
Product law of probability

ABSTRACT

In this paper we investigate the different ways in which students in lower secondary school (14–15 year-olds) reason about compound stochastic events (CSE). We ask students during clinical interviews to respond to CSE-tasks in a spinner context, where two linked spinners display equal or different sizes of red and white areas. We seek to enrich our knowledge of how students make sense of CSE by not focusing exclusively on sample-space grounded reasoning. We open up the analysis to how students' reasoning can reflect aspects of multiplicative reasoning in relation to The Product Law of Probability. Our results show that students have difficulty in applying well-grounded combinatorial reasoning as well as multiplicative reasoning to the tasks, but they do show intuitive reasoning that reflect aspects of The Product Law of Probability. Two ways of reasoning identified in the current study are area-based part-whole reasoning and lowest-chance reasoning.

1. Introduction

Compound stochastic events (CSE) refers to a two-stage or a two-dimensional random experiment such as the simultaneous rolling of two dice (Polaki, 2005). Learning to model and predict the probability of CSE is central to school curricula all over the world (Jones, Langrall, & Mooney, 2007). Compound stochastic events are complex, addressing many essential concepts of probability theory such as sample space, order of the single outcomes making up the two-dimensional experiment, independent events and conditional probability (Iversen & Nilsson, 2007). From such a complexity it is understandable why many students find it hard to make sense of CSE, a fact widely documented in the literature (Langrall, Makar, Nilsson, & Shaughnessy, 2017). In this study we take a further look at how students reason about and try to make sense of CSE.

A *Laplace perspective* has been dominating research on students' reasoning in and about CSE. From a Laplace perspective focus is on describing how students are able to identify and use sample-space composition in a part-whole model (number of favorable outcomes out of the total number of outcomes) for predicting probabilities of CSE (Chernoff & Zazkis, 2011). In this study we seek to enrich our knowledge of how students make sense of CSE by not focusing exclusively on a Laplace perspective. In order to broaden the research discourse, we open up the analysis to how students' attempts to make sense of CSE can reflect aspects of multiplicative reasoning in relation to *The Product Law of Probability* (PLP) in a two spinner context.

The investigation follows the principle of clinical interviews (Ginsburg, 1981), where the researcher interacts with the student in order to seek further information about students' reasoning. We address the following research questions:

* Corresponding author.
E-mail addresses: kjarand.iversen@nord.no (K. Iversen), per.nilsson@oru.se, per.nilsson@nord.no (P. Nilsson).

https://doi.org/10.1016/j.jmathb.2019.100723
Received 19 October 2018; Received in revised form 5 July 2019; Accepted 9 July 2019
Available online 09 August 2019

- How do students reason about CSE represented by the combination of two spinners?
- What aspects of PLP can be discerned in students' reasoning?

2. Theoretical background

2.1. A socio-constructivist approach on meaning-making

In the present study we use the socio-constructivist framework of contextualization (Halldén, 1999; Nilsson, 2007) to frame our analysis of how students make sense of CSE. The framework rests on the idea that, as human beings, we cannot describe or understand even the simplest thing in a completely exhaustive way. We always experience a phenomenon in a certain way, from a certain set of premises and assumptions, whereby some aspects of the phenomenon become salient and made available for reflection, but many other aspects are left out (Säljö & Wyndhamn, 1990). Talking about students' contextualization is a way of organizing and conceptualizing such principles of experiencing and reasoning (Nilsson & Ryve, 2010). Taking into consideration how students contextualize their reasoning in a learning task is about rationalizing the student behavior and sense-making. Say a student is expressing a solution to a task that does not follow the expected solution. Instead of classifying the solution as irrational and as an expression of a misconception that must be removed in favor of the 'correct' conception, from a contextualization perspective the solution is considered rational based on the student's perspective. In this way we focus on and become alert to what students actually do and can do, given the personal contexts in which the students are operating and of which their understandings are the products (Hung, 1997).

2.2. Perspectives on probability theory

There has been extensive research on students' understanding of theoretical probability across a variety of context and problem types (Langrall & Mooney, 2005). In this research, students' reasoning and understanding have been analyzed and interpreted in relation to theoretical constructs like sample space, independency between events and the role of ordered outcomes. In particular, research has focused on teaching and learning probability from a Laplace perspective (Chernoff & Zazkis, 2011). From this perspective the probability of an event is found by following two steps. First, you generate an equiprobabilistic sample space of all possible outcomes of a random situation. Second, you apply proportional reasoning to this sample space to calculate the probability of different random events of the situation. Applying proportional reasoning means that the probability is calculated as the ratio found by dividing the number of favorable outcomes by the number of all possible outcomes. This procedure applies to both simple and compound stochastic events. For illustration, say that we are throwing two dice and are interested in the probability of the event "double six". In the first step we generate an equiprobabilistic sample space, composed of the 36 ordered pairs. In the second step, we note that one of the 36 possible outcomes are favorable for "double six", reaching the probability for "double six" as 1/36. Application of PLP is an alternative approach to solving CSE tasks. In its simplest form PLP applies to two independent events A and B, where the probability of the combined event "A and B" is found by multiplication of the probabilities p(A) and p(B): p(A and B) = p(A)·p(B). Conceptually, the idea of the PLP follows the general multiplicative principle of operation (Schmidt & Weiser, 1995). In this case the event B represents a proportion p(B) of the sample space to the CSE event, and p(A) times p(B) represents a proportion p(A) of the event B. In other words, the probability p(A and B) is reached by seeing p(A) as a multiplication operation on p(B) (Schou, Jess, Hansen, & Skott, 2013). Hence, applying the PLP to the throwing of two dice, the event "double six" is found by p(6)·p(6) = 1/6·1/6 = 1/36. Note the difference between this approach and application of the Laplace perspective. In the latter the number of possible outcomes is calculated to be 36 (6·6) since each dice has six outcomes. Next, the probability of the combined event "double six" is found from the ratio of one favorable outcome for double six, over all 36 possible outcomes. On the other hand, applying PLP means first to use the Laplace perspective on determining probabilities for the respective single events using sample spaces. The sample spaces of the single events are generally more simple than those of the CSE. The probability for the CSE is then found by multiplication of these probabilities. Solving CSE tasks by application of PLP rests on a deeper knowledge about fractions and multiplication of fractions than application of the Laplace perspective. We now review some of the earlier research on students' interpretation and ways of solving CSE tasks from these two perspectives.

2.3. Research on CSE from a Laplace theoretical perspective

What research in mathematics education has found is that many students find it difficult, or unnecessary to take into account theoretical constructs of probability theory, and three issues identified in the earlier research that relates to this are equiprobabilistic bias, the role of ordered outcomes and over-generalization. Applying the equiprobability bias (Lecoutre, 1992) to a random experiment means that the outcomes of the experiment are considered to be equally likely due to chance (Nilsson, 2007). For instance, Fischbein, Nello, and Marino (1991)) asked 278 students (11–14 years old) to compare the probabilities for the events i) 5 on one die and 6 on the other and ii) 6 on both dice, in the context of throwing two dice. The task represents a situation of two events with none-equal chance, and where the order of the outcomes is needed to be taken into consideration to arrive at the proper sample space. What the researchers found was that many students just saw the situation as a matter of chance, without involving any sample space considerations. In the words of the researchers: "One might assume that some children have considered the pairs as ordered pairs and then the correct answer would be that the couples (5,6) and (6,6) are equiprobable….Analyzing the justification, it became clear that no order was considered and that the very frequent 'equiprobable' type of answer was mostly justified by the effect of chance …" (p.

532). If we assume that students understand the role of the underlying sample space in CSE then, making accurate probability predictions of CSE becomes a question of being able to outline all ordered pairs of the sample space (Polaki, 2005). However, what research has come to observe is that students often face difficulties in taking into account the order of the single outcomes, when generating all possible configurations of a CSE (e.g., Alston & Maher, 2003; English & Watson, 2016; Konold & Kazak, 2008; Nilsson, 2007, 2009; Speiser & Walter, 1998). For illustration, Vidakovic (1998) reported on a group of grade 8 students, who were asked to discuss the role of the order of outcomes in predicting the outcomes of the totals of two dice. The discussions showed that a majority of the students thought the outcomes (5,4) and (4,5) were the same outcome, arguing that "They're the same numbers, just written down in a different order….' cause five and four both equal nine." (Vidakovic, 1998, p. 72).

Lastly, we give some examples from studies showing how students can come to over-generalize methods and reasoning from simple stochastic events to CSE. The methods work in the case of simple stochastic events but not in CSE. Lysoe (2008) asked 76 pre-service teachers to respond to the following CSE task:

> Anna has three red, two green and one blue pencil in her pen case. She asks Maria to pick out two pencils without looking. Anna thinks that the probability that both of the pencils are red is 1/5 but Maria thinks that the probability is 1/3. Does either of them have the correct answer?

Of the 76 pre-service teachers, about 20% voted for 1/3 as the correct answer. Lysoe argued that these students transformed the situation to a one-step problem: There are a total of six pencils, from which two pencils are to be drawn. In transforming the situation to a one-step problem, the students neglected the compound nature of the situation by over-generalizing the application of a Laplace perspective. The situation is not very easy to deal with from such a perspective on probability. Understanding how PLP work might be a more applicable method of dealing with the situation and so retain the two-step structure of the task.

Shaughnessy and Ciancetta (2002) provided insight into how students can come to struggle with the transition from simple stochastic events to CSE in the context of spinners. Students in grade 6 to 8 were first asked to respond to the chance of arriving at black when spinning just one spinner, displaying an equal area of black and white. Later the students were asked to respond to the two-spinner situation: "A player wins a prize only when both arrows land on black after each spinner has been spun once. Jeff thinks he has 50-50 chance of winning. Do you agree?". Around 80% of the students agreed with Jeff and justified their answers by referring to the chance of the individual spinners: "They're half of a circle, so ½ is 50%";" Both spinners have an equal amount of B and W so it's 50-50"; "The circles are exactly in half". Using the same tasks as Shaughnessy and Ciancetta (2002), Watson and Kelly (2004) also observed students' difficulties in responding properly to the question of Jeff's chance to win the game. To account for the students' difficulties, Watson and Kelly (2004) made the reflection that, for the students, "the idea of the independence of the operation of two 50/50 spinner was not at all intuitive" (p. 390). In both studies, the students dealt with the CSE by applying methods working in a simple stochastic experiment, with limited consideration of the relationship between the two spinners.

2.4. Research on CSE from a PLP perspective

When approaching a CSE from a PLP perspective, a tree diagram is often used to model the situation (Konold, 1996). With such a model, a CSE is considered as a series of simple stochastic events, one event happening after the other in a consecutive order. For example, the throwing of two dice is viewed as one die being thrown after the other. Several studies on PLP have used a certain type of task where an *object* (marble or robot) travels through a system of junctions, to end up in one of several outcomes (endpoints). Iversen and Nilsson (2007) refer to such systems as *One-Object Stochastic Phenomena* (OOSP). Compared to other CSEs, the question of the order of the single outcomes disappears in OOSP since the junctions (the single outcomes) are already structured in a consecutive order. The ordered structure of an OOSP implies that the student can focus on how to connect the single outcomes making up the CSE.

Two seminal studies on OOSP are the studies by Green (1983) and Fischbein, Pampu, and Minzat, 1975). In Green's study students were asked to predict the outcomes of a robot's random walk. A robot walking in a maze encountered one or several junctions before ending up in one of eight rooms. The students had to judge whether the rooms were equally likely for the robot to arrive at. Only 7% of the 2973 students of age 11–16 were able to give correct response to this task and, surprisingly for the researcher, only 13% of the oldest students (age 15–16) gave a correct response. For this latter group of students, the PLP was included in their curriculum. Green concluded: "The multiplication principle is poorly understood … Clearly little conceptual understanding has been achieved, and using a tree diagram was just a mechanical device." (p. 780). Ron, Dreyfus, and Hershkowitz (2017) provided further information on difficulties students may have to make use of visual mediators (models) (Sfard, 2008; Ryve, Nilsson, & Patterson, 2013) for modelling random situations from a PLP-perspective. They asked students to solve a sequence of CSE tasks with progressive use of a two-dimensional area-model. While the first task was accompanied by a 10×10 area-model (Fig. 1a), the last task was a plain text problem without support of any visual mediator (Fig. 1b). What was found was that, while students quite easily were able to solve tasks when the task presentation was accompanied by the area-model, they struggled to solve tasks when they had to make, or recall, the model by themselves. Hence, it seems as if students are able to use a visual mediator for outlining the sample-space of a CSE when the mediator is directly presented to them in a task, but face difficulties to see, by themselves, when and how a visual mediator can be useful if it is not explicitly presented in the task.

In the study by Fischbein et al. (in Fischbein et al., 1975) students of age 6–14 were interviewed on tasks related to concrete inclined boards with systems of progressive branching channels. The students were asked the following question: "If I drop the marble a lot of times, one after another, will it come out of each place the same number of times, or will it come out of some places more often than others?". In this case of a non-equiprobable system, 54% of the pre-school children gave a correct response. The rate of correct responses increased with age. For instance, in the group of students aged 13–14, up to 92% provided correct responses. As

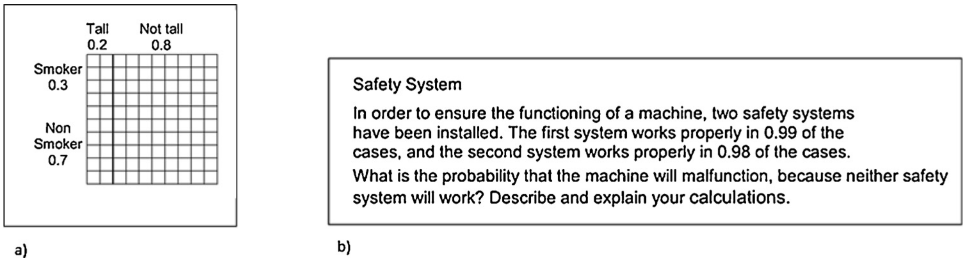

Fig. 1. In a study by Ron et al. (2017) students were asked to solve a sequence of CSE tasks with progressive use of two-dimensional area-model. In the first task a 10 × 10 area model was used (a), while the second task (b) was not accompanied by any model representation.

pointed out by Fischbein et al. this experimental setting "provides the condition for direct intuition, without recourse to actual mathematical calculation" (p. 160).

Hence, while Green's study indicated students' difficulties in mastering the PLP, Fischbein et al.'s study indicated that even young students may be able to develop an understanding of the PLP in experimental situations. Iversen and Nilsson (2007) followed up on Fischbein et al. (1975), exploring further how students can solve and reason about OOSP in a computer environment called Flexitree (Iversen, 2003). In this digital context, Iversen and Nilsson (2007) identified two informal strategies that students used for deciding on the probabilities to arrive at the different outcomes in the OOSP represented in Flexitree: a *main-road strategy* and a *division strategy*. The main-road strategy is an intuitive, qualitative strategy. Following the main-road strategy, a junction of the OOSP is not modelled as a uniform probability distribution. Instead, one of the paths leading out from a junction is considered to be the main road, while the other path out is seen as a leakage and, consequently, the probability to follow the "main road" is considered to be higher than dropping into the leakage path. Applying the division strategy students accept the symmetry of the junctions, perceiving the outcomes out from a junction to be equally likely. The division strategy relates to the *splitting model,* proposed by Confrey and Smith (1995), as an informal model for a more formal multiplicative way of thinking. Students applying the division strategy start out with an entity such as 1, 100% or a certain number of marbles, and split this quantity repeatedly into two equal parts at each junction. In the case of starting with 100% this entity is split into two entities of 50% after the first junction and split again into entities of 25% after a second junction, and so on. Hence, based on the conceptual relationship between division and multiplication, the division-strategy described by Iversen and Nilsson (2007) shows how students can develop an informal capacity of the PLP in the context of OOSP-systems.

2.5. The aim of the study

The aim of the present study is to enrich our knowledge of how students make sense of CSE according to the following research questions:

- How do students reason about CSE represented by the combination of two spinners?
- What aspects of PLP can be discerned in students' reasoning?

In order to address the research questions, the students' perspective needs to be accounted for. In the present study we account for the students' perspective from the socio-constructivist framework of contextualization focusing on how the students express aspects of a Laplace perspective or of the PLP in line with the personal context within which they operate. The first research question focuses on how the students reason about CSE from a rather broad perspective, whereas the second research question addresses what aspects of PLP could be discerned in the students' contextualization of CSE.

3. Method

3.1. The participants

Eighteen students in ninth grade formed the population of this study. In Norway, students are introduced to probability from grade 6 with emphasize on basic probabilistic concepts. Students typically work with simple equiprobabilistic stochastic events but are also introduced to some CSEs like throwing two coins or two dice. Probability calculations are typically done by application of Laplace perspective at this level in school, while PLP is introduced at a later stage in the educational system.

3.2. Clinical interview

To examine students reasoning about CSE requires collecting data that provides opportunities for a fine-grained analysis of the students' ways of making sense of a task and the content involved in the task. The value of clinical interviews as means for such analysis has been discovered by researchers (Zazkis & Hazzan, 1998). Clinical interviews provide an opportunity to collect the words and actions needed for such analyses (Zazkis & Hazzan, 1998) and can be a powerful instrument in eliciting narrative data that allows

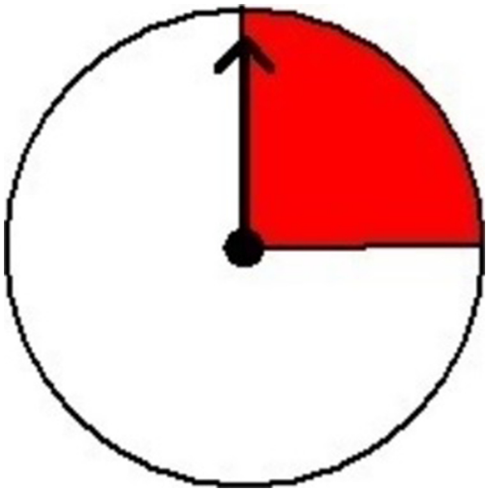

Fig. 2. A model of the single concrete spinner used to introduce the spinner context. The spinner is spun, and a fixed arrow points to the result, either white or red. Students were asked to predict the probability for the result to be red (task 0).

researchers to investigate people's views in great depth (Kvale, 1996). In the current study, we followed the logic of clinical interviews, which provided the opportunity to add follow-up questions to gain further insight into students' reasoning. Each interview began with a task, accompanied by concrete material (spinners) or written representation of the task. The role of the interviewer was to follow up students' responses, by asking them to explain their answers. The contingency of questions is a central element in clinical interview, aiming for more reflection on the part of the subject. The students were interviewed individually in a separate room close to their regular classroom. Each interview lasted for about 30 min.

3.3. The tasks

The interview protocol included two types of tasks: *likelihood tasks* and *comparison tasks*. Although the expected solutions to the likelihood tasks were numerical values, the comparison tasks gave the student the possibility to offer qualitative argumentation, or even intuitive responses. A student who was unable to reach a numerical value from a calculation on a likelihood task was asked to estimate a numerical value of the probability. The interview included several probability contexts, such as dice, boxes with marbles, and spinners, supported by concrete materials. In this article we will limit the results and analysis to students' work with the spinner tasks. In order to introduce and familiarize the students with the spinner context, the interviewer (the first author of the paper) began each interview by showing the student a concrete spinner (Fig. 2 shows a model of this spinner). The interviewer conducted a few experiments with the spinner to show how it worked and then the students were asked to predict the probability of the outcome "red" from a single spin with the spinner. None of the students had any problem giving a correct numeric value (1/4) to this task.

3.3.1. Likelihood tasks – two spinners

Following the same procedure as in the introduction task with one spinner, the two-spinner context was introduced to the student by showing two concrete spinners, each with a red and a white sector (Fig. 3). The situation was explained to the student as a game, with the result "red on both spinners" meaning a win for the player. Students played this game a few times. Each time, they had to tell

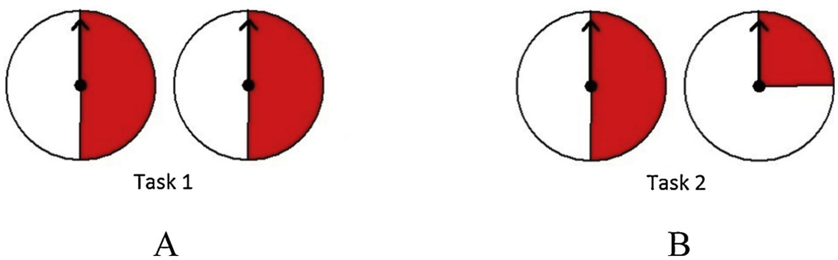

Fig. 3. A picture of the spinners in the two likelihood tasks. The students were asked the following question:" What is the chance of winning this game? What is the chance for both spinners to show red?". The spinners in task 1 will be denoted 1 L and 1R, and similarly, 2 L and 2R will be the denotation for the spinners in task 2, where L and R stands for Left and Right respectively.

i. The chance to win is highest if I choose setup A.
ii. The chance to win is highest if I choose setup B.
iii. There is an equal chance to win with either setup A or B.

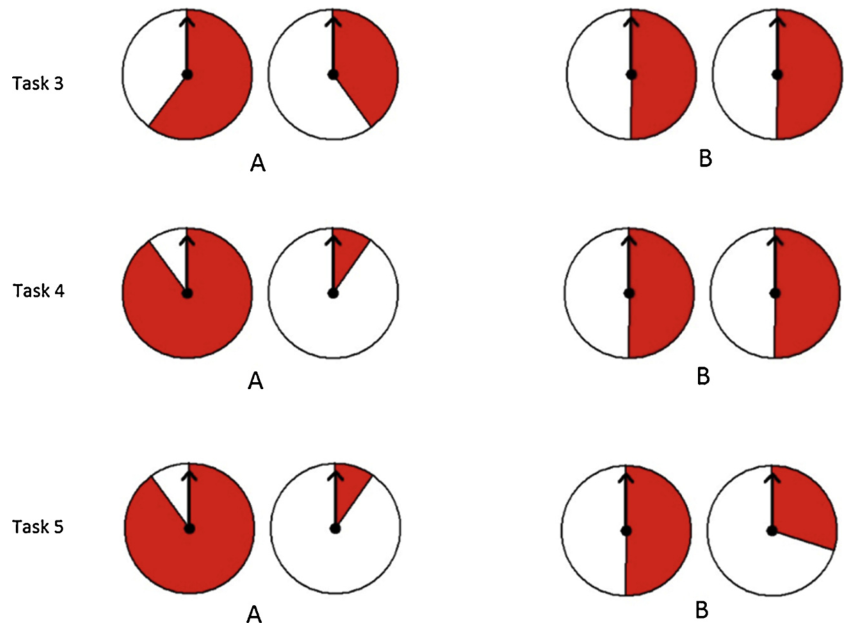

Fig. 4. A picture of the spinners in the three comparison tasks. The students were asked to consider if any of the setups A and B gave a higher chance to win (arriving at two red). Again, L and R denote Left and Right spinner of a setup. For instance, we name the spinner on the left-hand side of setup 3A as spinner 3A-L and the spinner on the right-hand side as spinner 3A-R. The same logic is used for all setups in task 3, 4 and 5.

what the outcome was and evaluate if it was a win or not. This experimentation with the two spinners was meant to introduce the context to the students. Next, they were asked to reflect on the following questions: "*What is the chance of winning this game? What is the chance for both spinners to show red?*" The students responded to two different likelihood tasks (task 1 and 2, see Fig. 3). The design of task 1 (Fig. 3A) is a stochastic situation similar to flipping two coins: Each spinner has two outcomes of equal probability. In task 2 the students had to predict the probability for "two red" in a setup with one uniform and one non-uniform spinner (the latter with 25% red and 75% white, see Fig. 3B). In moving from task 1 to task 2 the interviewer actually removed one of the concrete spinners and replaced it with a new spinner. Both tasks can be solved by application of PLP: ½ x ½ = ¼ to task 1 and ½ x ¼ = 1/8 to task 2. Alternatively, the tasks can be solved by application of Laplace perspective, giving respectively 4 and 8 equally likely outcomes, and the solutions ¼ and 1/8 for the one favorable ("two red") outcome. The multiplicative relationship between the two tasks is important. That is, the probability of obtaining "two red" in task 2 is half the probability of obtaining "two red" in task 1 (1/8 and ¼ respectively) due to the effect of the second spinner (the second spinner is changed from ½ to ¼). The underlying rationale for choosing these tasks was to explore how the students were able to exploit this multiplicative relationship between the tasks, or if their contextualization of the situation led them to other interpretations.

3.3.2. Comparison tasks

The comparison tasks followed up on the likelihood tasks. Again, the context is explained to the student as a game, but now the student was asked to choose between two different setups, setup A and B before playing the game (see Fig. 4). While the likelihood tasks were accompanied by concrete material, the different setups in the comparison tasks were shown as pictures printed on paper. The students had to consider three different comparison tasks. For each of these, they were asked to reflect on three possible alternatives followed by the question "Which one of these alternatives do you believe in? And why?" (see Fig. 4). Setup B in task 3 and 4 is the same as in task 1, so the probability in these cases is still ¼ (16/64) for the favorable outcome ("red on both"). The probability of the other setups can again be solved by application of PLP or Laplace perspective, giving the probabilities 15/64 and 9/64 for the favorable outcome "two red" in setup A in respectively task 3 and 4. The probability of the favorable outcome in task 5 is calculated to be ½ x 3/10 = 3/20. That is, the probability for the favorable outcome is higher in setup B for all tasks.

3.4. Method of analysis

The method of analysis followed the principles of constructivist grounded theory (Charmaz, 2014). This means that both theory and data were central in giving an account of how students made sense of CSE. The theory of contextualization provided a logic to our analysis; a logic that is based on the assumption that both behavior and articulated expressions are rational in terms of being intentional (von Wright, 1971). Practically this means that, in deciding on how students reasoned about the spinners, we considered their behavior as rational in terms of an act being performed in order to solve a specific task in the situation in question (Halldén, 1999; Nilsson, 2007, 2009). However, because the theory of contextualization tells nothing about the content of students' meaning-making, our analysis operated also on previous research and on probability theory as such. As an overall frame of reference, the two theoretical approaches to probability – Laplace perspective and PLP – guided our analyses. In other words, our theoretical approach was made up by contextualization, the two theoretical approaches to probability, together with the bricolage (Cobb, 2007) of empirical findings from stochastics education reported on above.

In practice the analysis followed three steps. First, we categorized students' responses according to their answers. Second, we asked, what specific tasks the students were engaged in, in order to make their answers reasonable and intelligible. Third, we looked at what was the overarching idea that guided the students' way of solving the task and, connected to *The Constant Comparative Method* (Glaser & Strauss, 1967; Starrin, Dahlgren, Larsson, & Styrborn, 1997), we compared their different ways in order to reach a set of categories of how students can reason about CSE, represented by the combination of two spinners. For instance, in our analyses we noticed that several students answered 0.5 and 3/8 on task 1 and task 2 respectively. In making sense of that, we asked, "What personal context(s) were the students engaged in that could make their answers reasonable and intelligible?" With deeper investigation of the interview we saw how the students were focusing on visual features of the task situation. They contextualized the composition of two spinners within the realm of an area-context where they were engaged in finding out how much of the total area of the two spinners was red. All interviews were transcribed. Following the three steps described above, the two authors worked in parallel developing core categories from analyzing the transcripts. Next, the two authors worked together in developing the final categories by comparing, discussing and refining the core categories they had found.

It is important to point out that we have no intention of categorizing a *particular student* into a type of reasoning. Rather, we want to describe the different types of reasoning identified among the students as a group, and in particular, how they made sense of, or had difficulty in making sense of, the different CSEs in the interview situation.

4. Results and analysis

The results and analysis are structured according to our two research questions. This means that we first describe the ways in which the students were reasoning about CSE, represented by the combination of two spinners. We then address how some of these ways of reasoning reflect aspects of PLP. The interviews were conducted by the first author, referred to as "I" in the excerpts below.

4.1. How do students reason about CSE represented by the combination of two spinners?

4.1.1. Area-based part-whole reasoning

Area-based part-whole reasoning is a *visual based* type of reasoning, where students count out a number of parts of equal area, with a certain number of these is seen as favorable. In the case of one spinner (Fig. 2), the single red part is seen as the one favorable outcome out of four possible outcomes/parts. This is a proper way to solve simple stochastic phenomena tasks. What we see from our data however is how students "generalize" this one-step procedure to CSE tasks (Shaughnessy & Ciancetta, 2002), leading to improper probability predictions.

4.1.1.1. Area-based part-whole reasoning for simple stochastic phenomena. The spinner situation was introduced by showing the students one real spinner and asking them to consider the probability for the result "red". None of the students had any problem solving this task, applying area-based part-whole reasoning. This is how Casper solved task 0 (Fig. 2).

I: If I spin this spinner, what do you believe its chance is to stop on red?

Casper: It must be one fourth.

I: Ok. Why?

Casper: Because … as it looks now it seems like its divided into four parts … as if it is one fourth in a way.

Casper does not draw any helplines. However, he structures the spinner into four equally sized parts, one of which being red. Because only one area sector out of four equally sized area sector is red, Casper claims that the probability for the result red "must be one fourth".

4.1.1.2. Area-based part-whole reasoning to CSE – the case of Brit. In tasks 1 and 2 above, most students came to apply area-based part-whole reasoning (Fig. 5). For these students, the two spinners are not perceived as independent objects (Watson & Kelly, 2004). Instead, the two-spinner situation is transformed to a one-step stochastic phenomenon. Students who apply area-based part-whole reasoning perceive the two spinners as one 'whole' object, consisting of a certain number of equal-sized parts. We follow Brit's work on task 1, 2 and 3 to describe variations of area-based part-whole reasoning. Even though Brit is quite open to a different interpretation of the situation and different solving strategies, area-based part-whole reasoning influences her way of thinking throughout all three tasks. We start out by looking at her work on task 1 after the interviewer has asked her about the probability for the result two red in this task: Brit: Fifty percent.

Fig. 5. A model of area-based part-whole reasoning to the CSE two spinners. The situation is interpreted as a simple stochastic phenomenon, with a certain number of favorable outcomes. Task 1 is interpreted as a situation with one spinner consisting of 2 reds out of 4 possible, and in a similar way, task 2 is interpreted as a spinner with 3 red parts out of 8 possible parts.

I: Fifty?

Brit: Maybe twenty-five.

I: Maybe twenty-five? Just take your time to think. Brit: Twenty-five I believe.

I: Ok. Can you tell me why you believe this?

Brit: At first, when I was thinking ‘fifty’ I thought it is an equally number of red and white ones. But, to get red on both is a … so I took half of this. I: Ok. Why half?

Brit: Well, because … then we must consider each and one of [task 1], this is hundred percent all together … because it is fifty percent, if you understand what I say? I: Yes. And then it becomes twenty-five?

Brit: Yes, about this (laughs)

I: Ok. Brit’s initial response reflects area-based part-whole reasoning: there is a total (one whole) of four equal-sized areas and two of these areas are red so P(two red) = 50%. The response is rather brief; it’s more like a quick idea passing her mind. However, we will see that this way of reasoning keeps coming back in her work on subsequent tasks. When the interviewer raises a query on 50% (her first answer) she is quick to lower this value to 25%. Her reasons for this are not clear at this point, but it is reasonable to believe that she contextualizes the query from the interviewer as if her first answer is wrong and needs adjustment. That her adjustment is an effect more of social, interactive aspects than of a conceptual idea becomes rather clear in how Brit reacts when the interviewer pushes her to explain her decision on 25%. She has a hard time making sense of why she responded 25% for two red. She expresses that one must consider the contribution of each spinner, which speaks to the fact that she structures the situation in more detail than in her initial response. However, her reasoning is incomplete and unclear. Moreover, when she responds to the question “Why half?”, she reflects on hundred percent and fifty percent, which speaks to the fact that the one-step procedure involved in the area-based part-whole reasoning is still prominent in how she contextualizes the task solution. In task 2 Brit stresses the area-based part-whole interpretation further.

Brit: Three eight.

I: Three eight? Can you explain this?

Brit: First, I will divide them both into four parts. Then I will count the number of red, and the number of white ones. Easy and simple. So, its eight all together so then …. Yes.

I: Ok, but you didn’t do it like this earlier?

Brit: No.

I: If you had done so it would have been fifty?

Brit: Yes, but maybe it is fifty on the earlier one as well.

I: What do you believe mostly? Fifty earlier, and three eighths now? Brit: It might be that it was fifty earlier. That is, four eighths. And, three eighths now.

Brit explicitly articulates the partition of each spinner into four equal parts, with one whole divided into eight visually observable parts. Next, she counts out the number of red parts (three), and the total number of parts (eight), as an argument for her answer 3/8. She even adds the comment” easy and simple” to express that she finds this task to be easy. The strong influence of part whole area reasoning becomes even more clear when the interviewer reminds her of her answer on task 1. She turns back to her initial idea in task 1, being open to accept 4/8 as the probability for “two red” due to the idea that there are four red areas out of a total of eight areas. In task 3 Brit is asked to compare setups A and B according to the chance of arriving at two red (Fig. 3). Responding to task 3 her area-based part-whole reasoning takes yet another form.

Brit: Equally likely.

I: Equally likely. Ok. Why?

Brit: Because, if you look … here [3A-L] … if you put this one [3A-R] over there [3A-L] … then it’s one whole [3A-L]. If you move this one [3B-R] over there [3B-L] it is also one whole.

I: And, the chances to win are the same then, with these two [3A] and these ones [3B]?

Brit: Yes.

In task 3, Brit focuses on the total area of red on both setups. She describes how adding the red part of 3A-R onto spinner 3A-L makes a whole red spinner. And, in the same way, adding the red part from spinner 3B-R onto spinner 3B-L makes another whole red spinner. Her argument goes like this: because the total red area of the both pairs of spinners are the same, this implies that the probability of the result “two red” must be the same in both setups.

As many as 14 out of 18 students showed area-based part-whole reasoning on at least one of the tasks in the interviews. See Table 1 for an overview of the result for the whole group on tasks 1 and 2.

Even though area-based part-whole reasoning was typically identified as an initial response to a task, several students, like Brit, came back to this way of reasoning on several of the tasks. This shows the impact this way of reasoning had on students’ ability to

Table 1

N = 18. X indicates that the student used area-based part-whole reasoning to solve the task, while a blank slot indicate he did not. 14 of the 18 students solved at least one of the two tasks with area-based part-whole reasoning, while 3 students did so on both tasks.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
T1	X		X	X	X		X	X	X			X	X				X	X
T2		X		X							X					X	X	X

contextualize and solve CSE tasks.

4.1.2. Combinatorial sample-space reasoning

Five of the 18 students in the present study used combinatorial sample-space reasoning to solve task 1, but such reasoning was not identified in any of the other tasks. In line with earlier research (e.g. English & Watson, 2016; Konold & Kazak, 2008; Speiser & Walter, 1998; Vidakovic, 1998) the students in our study had difficulties in generating a conventional sample space. Following the notion introduced by Chernoff and Zazkis (2011) some of the students generated a sample space (RR, RW, WR, WW) to solve the task, while others used a sample set (RR, RW, WW). The dialogue with the students shows that the latter group of students perceive RW and WR to be "the same" outcome, and therefore interpret the sample set to consist of three equally likely outcomes. To illustrate the combinatorial sample-space reasoning identified in our study we look at how Hanne and Frode solved task 1. While Frode generates a sample space, Hanne uses a sample set to solve the tasks. Hanne:(task 1) It is fifty percent.

I: Fifty percent chance for the result two red?

Hanne: Yes. It can also be red and white, and two white.

I: But the chance for both spinners to stop on red is fifty percent? Hanne: Thirty-three.

I: Thirty-three?

Hanne:Yes.

I: How do you justify this? Hanne:One possibility is that they stop on two red ones, another one is one white and one red, and both could stop on white.

Hanne's initial response is area-based part-whole reasoning. But, as she continues to reflect on the task she turns to reason in accordance with combinatorial sample set reasoning, suggesting a sample set that she see as consisting of three equally likely outcomes. Frode, on the other hand, solves task 1 by generating a complete sample space:

I: What is the chance of winning in this game? What is the chance for both spinners to stop at red?

Frode: Twenty-five percent.

I: Twenty-five percent? Ok. Why?

Frode: Because it can end up on four different…no, three different…or, four…it is the same. Anyway, it can end up at red-red or white-white, or white-red or red-white.

I: So, there are four different possibilities?

Frode: Yes.

I: And, one of these is red-red?

Frode: Yes. I: And, that is why you think the chance is twenty-five?

Frode: Yes.

Even though Frode generates a complete sample space with four outcomes, his remark "…it can end up on four different…no, three different…or, four…it is the same" indicates that he is rather insecure to whether he should use 3 or 4 outcomes to calculate the probability for the result "two red". After a direct yes/no-question from the interviewer, Frode sticks to 25% as his last reply to this task.

4.1.3. Multiplicative reasoning

As mentioned before, the interview protocol was structured to trigger multiplicative reasoning, but such reasoning appeared only to a limited extent in our data. Two of the eighteen students explicitly applied PLP to the tasks, and yet two more students responded in a way that indicated intuitive multiplicative reasoning. We look more closely at how these students responded to the tasks.

4.1.3.1. Application of PLP – the case of Anne. Two students explicitly reasoned in accordance with the principles of PLP to solve the CSE tasks by consistently multiplying fractions to calculate the probability of the result two red in tasks 1 and 2 and to compare the probabilities of two red in tasks 3, 4 and 5. For example, Anne was quick to answer "one fourth" to task 1, with the argument "we have to multiply one half with one half. It is one fourth." Likewise, she solves task 2 by "One half times one fourth. It is one eight". She applies similar reasoning to solve the comparison tasks, for example task 4:

Anne: (Task 4) I will calculate with parts again.

I: Ok.

Anne: It is one multiplied by nine [4A].

I: Hundredth?

Anne: Mm. And then it's one fourth.

I: Yes.

Anne: Nine hundredth, how much is that (in percent?)?

I: Nine percent.

Anne: Nine percent … huff, yes (laughs). It is that simple. And, one fourth is twenty-five, so then I would go for B.

I: You would play with these [4B]?

Anne: Yes.

The aim of the interviews was to receive rich information on students' probabilistic reasoning in CSE. If students encountered difficulties in performing calculations, the interviewer helped leading the student forward, in order to gain further information about the students' sense-making of CSE. For instance, in the dialogue above we see that Anne is insecure about use of different representation of rational numbers, which might influence her capacity to compare the probabilities of different setups. When solving task 3 she was quick to calculate probabilities by the use of fractions. A similar approach to solve task 4 would be to multiply 1/10 (red part on 4A-L) by 9/10 (red part on 4A-R), but in this case she solved the task by first multiplying the numerators (1 multiplied by 9). When the interviewer interferes, she asks for help to converting the fraction 9/100 to percent and is rather embarrassed when she realizes how easy this is. After finishing these calculations, she has no problem comparing the probabilities of setup A and B. *Intuitive multiplicative response* Some students could not provide a numeric value to task 2. On such occasions, the interviewer asked the students to estimate a numeric value. A majority of these students then estimated the probability to be in the range 18–20%, and thus not seeing multiplicative relationship between task 1 and 2 (the probability for the result "two red" in task 2 is 12.5%, half the probability of 25% for this result in task 1). However, two students, Bjarne and Ivar, estimate a numeric value to task 2 that reflects a kind of multiplicative relationship between the two spinners. Bjarne answers 25% to task 1 based on combinatorial reasoning. But, when asked about task 2 he could not come up with a numeric value. The interviewer challenges him to estimate a value:

I: (Task 2) Can you just estimate a value, a percent that seems to fit? It doesn't have to be a precise or exact value.

Bjarne: About twelve and a half percent.

I:Twelve and a half? Why?

Bjarne: There is a rather high chance for red on that one [2A], but then it is half as much chance for red on that one [2B]. It is … yes … more difficult that the last one (task 1).

I: So you think twelve and a half then?

Bjarne: Yes, about that.

Bjarne compares the probability to get red on spinner A and B in task 2, and observes that the probability for red on spinner B is half the probability for red on spinner A. It is this observation that seems to motivate his reply 12.5% to this task.

Ivar also answered 25% to task 1, based on combinatorial reasoning. When asked to consider task 2 his first reply was: "I am quite sure it's not very high [the chance]", but he did not attempt to compute a numeric value. The interviewer then offered him the alternatives 10%, 20% or 33% for the probability of "two red" to task 2.

Ivar:I would guess about 10.

I:About 10? Could you say some words about why?

Ivar:Because … it was 25 percent on the previous one (task 1).

I:Ok.

Ivar:Then this one is half [2R].

I:Yes.

Ivar:I am sure it is not thirty-three since this one [task 2] is less. And, this one [task 2] is more than 5 percent less than [task 1]

I:Ok.

Ivar:So, it can't be twenty either.

I:So, about 10 percent?

Ivar:Yes.

Both Bjarne and Ivar gave numerical values to task 2 in accordance with the multiplicative relationship between task 1 and 2. The interviewer gives Ivar the options 10%, 20% and 33% to reflect on in task 2. Based on these options, Ivar suggests 10% for two red in task 2. 33% is not an option for him. He finds 20% being too much because he thinks the probability to obtain two red in task 2 should be significantly lower than just 5% less than obtaining two red in task 1. In this case, 10% is the alternative closest to half of his answer to task 1. Bjarne solves this task by comparing the amount of red on the two spinners in task 1 with the amount of red on the two spinners in task 2, and this leads him to say that the probability for "two reds" in task 2 is half the probability for "two red" in 1. That is, these students show intuitive responses in line with the multiplicative relationship between the probabilities in task 1 and 2.

4.1.4. Lowest-chance reasoning

In comparing two different pairs of spinners, according to the chance of obtaining two red outcomes, several students expressed a *lowest-chance reasoning*. Applying lowest-chance reasoning on tasks 3, 4 and 5 means that the students distinguished and focused on the single spinner, from which they found that the lowest chance was to obtain red. In other words, the pair of spinners containing the single spinner with lowest chance for red was claimed to have the lowest chance of generating two red outcomes. In task 4 for instance, setup A is claimed to have the lowest chance for two red because it involves the single spinner (the spinner on the right-hand side of setup A) with lowest chance for red.

Casper's first reaction to task 3 was that the chance for two red was equal between setup A and setup B. But, quite soon he starts to doubt and instead finds the chance to obtain two red from setup B larger than from setup A, "Because…here it is equally large…in a way…half on each. But here (setup A) it is less than half on one, and a bit more than half on the other again." Casper's reasoning is not very specific. What he does, basically, is to provide an observational report of the two pairs of spinners. What seems to be crucial for Casper, for deciding on setup B, is that in setup B the white and red areas are of equal size for both spinners, which is different from the spinners of setup A. But if we look at how Casper responds to task 4 and 5, there is reason to assume that he, already in task 3, adopts lowest-chance reasoning, that there is less chance to obtain two red from setup A than from setup B because, setup A contains the spinner from which it is lowest chance to obtain red. In task 4, Casper articulates this idea:

Casper: Yes, I would go for B.

I: Ok. Why?

Casper: Because here [4A-R] you see only a small part is red …

I: Mm.

Casper: … of it all.

I: So it is the highest chance to win with these two (4B)?

Casper: Yes.

Casper is zooming in on one spinner. He makes no explicit comparison between setup A and setup B. His focus is on spinner 4A-R, from which he perceives it difficult to obtain red. It is this insight that forms the basis for choosing setup B on task 4. In task 5 Casper elaborates his lowest-chance reasoning:

Casper: I would go for B here too. This one spinner (5B-R) has more red … than this one (5A-R). And, even if this one (5A-L) is quite large then, it is very little red on this one (5A-R).

I: Mm. Casper:On this one (5B-L) again, is half …. there it is half red and there (5B-R) it is … yes, it is less than half. I am not sure how much it is.

I: Yes, it is thirty percent.

Casper: Yes. Thirty percent chance, then I would rather go for B.

The small cut of red on Spinner 5A-R is in focus. However, the design of task 5 challenges Casper to be more sensitive and explicit in comparing the areas of red between the two setups. In task 5 the red area of spinner 5B-R is reduced, compared to the corresponding red area of spinner 4B-R in task 4. Casper also takes note of, “that one is rather large [5A-L]”. Moreover, it is visually rather easy to see that the proportion of red on spinner 5A-L is greater than the proportion of red on spinner 5B-L. However, lowest-chance reasoning is strong. For Casper, the difference in red between spinner 5A-R and spinner 5B-R is still large enough to give him reasons to claim that the chance for two red is highest in Setup B. So, what seems to be crucial for several students, when comparing the outcomes of CSE, is the difficulty to obtain one of the single outcomes of a CSE. More generally, if one single outcome of a CSE is hard to get and one single outcome is easy to get, then the outcome that is hard to get is perceived to have greater impact on the entire CSE than the outcome that is easy to get. Bjarne makes this explicit: Bjarne:Yes, that one…I would have chosen that one [4B]

I:Okay. Why?

Bjarne:It’s a larger (red) area on that one [4B-R]…there is an equal-sized area on both these [4B]. Here [4A-R], its much less area to target, so it’s easier to end on white color.

I:But then it is higher chance there [4A-L]?

Bjarne:Yes, but it doesn’t help with only one correct. You need to get two!

Most of the students, 13 out of 18, reasoned in the same way as Casper and Bjarne, pointing out the spinner with lowest chance for the result red. An interesting variation from this was a small number of students identifying the same spinner but expressing its high chance for the result white. For example, Ivar gave a correct response to all three comparison tasks, with this type of argument. When asked to explain his answer to task 3 he voices, “On that one [3A-R] white is largest. There [3A-L] red is largest. The result can easily turn out to be red and white”. Similarly, for task 4 he claims, “Because, the only result you get on that one [4A-R] is white.”, and in task 5, “It’s a very high chance to get white on that one [5A-R]”.

4.2. What aspects of PLP can be disclosed in students’ reasoning?

In this section we highlight how students’ reasoning, particularly cases of lowest-chance reasoning, expressed aspects of multiplicative ideas, related to the PLP. In relation to this, we also provide insight to what role task-design can have for the appearance of multiplicative ideas in CSE, represented by the combination of two-spinners.

Above all, the difference in response and reasoning between the comparison and likelihood tasks is quite striking. The success rate was higher in the comparison tasks, compared to the success rate in the likelihood tasks. What we propose as an explanation for this difference is that the comparison tasks cued more of a multiplicative contextualization than the likelihood tasks did, which cued an area-based part-whole reasoning. Lower-chance reasoning provides an example of how students were inclined to adopt to multiplicative ideas in comparing two CSEs of two spinners and is based on the idea that the result of one spinner influences the total result of two spinners. If there is a small chance of red on one spinner then students see it as unlikely that the result will be red on both spinners, even when the total area of red of the two spinners are the same (task 3 and 4). For example, Casper and Bjarne articulated how a low chance for red on one spinner implied a low chance for the result “two red”, even when the total area is the same. This way of reasoning is related to how one fraction influences the product of two fractions in PLP; if one of the fractions is close to zero it will highly influence the product of two fractions. Similarly, Ivar articulated that a high chance for white on one spinner and a high chance for red on the second spinner imply that the chance for “one white and one red” is high. It could indicate that they consider “two white” or “two red” as unlikely with the same reasoning as above, or the alternative that multiplication by two fractions close to 1 must be a product quite high (more than 50%). Or, the reasons for their responses could be mainly intuitive. Nevertheless, even when their argumentation is qualitative and incomplete these students’ reasoning could be a starting point toward understanding PLP.

In both task 3 and 4 the total area of red is the same in setup A and B. However, in the transition from task 3 to task 4, we observed a shift from area-based part-whole reasoning to lowest-chance reasoning (see (Table 2)). In this transition the “less symmetric” situation – task 4 – seemed to direct the students’ reasoning towards lowest-chance reasoning more than was the case in task 3, where the red areas were more equally distributed between setup A and setup B. The transition from task 4 to task 5 then supported a reverse shift in students’ reasoning (see (Table 2). Between task 4 and task 5, the transition goes from a situation where the setups to be compared have the same total area of red (task 4) to a situation where one setup has less total area (setup B), compared to the other setup (setup A) (task 5). This shift indicates how strong area-based part-whole reasoning is. In task 4 the two areas of red to be compared is the same and the students found it proper to distinguish between the two setups according to lowest-chance reasoning. However, in task 5, the two total areas of red in the two setups are not the same, and now half of the students argued in line with area-

Table 2
C: Correct answer, E: Equal chance response (incorrect), W: Wrong answer other than equal chance response. Four students did not respond to task 3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
3	C	E	E		C	C	E			E	E	C	C	C	C	C		E
4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
5	C	W	W	C	C	C	C	C	C	W	C	W	C	C	C	C	C	E

based part-whole reasoning (even though it is setup 5B with the least total of red that have the highest chance to get the result two reds).

5. Discussion

Previous research has identified the difficulties students can have in making sense of CSE (Langrall et al., 2017). The present study enriches the picture of students' difficulties in applying the Laplace perspective to probability by identifying students' rather persistent use of area-based part-whole reasoning in solving CSE tasks, even after being taught how to solve such tasks from a Laplace perspective. The study also contributes to research in probability education by shedding new light on such difficulties and by demonstrating how students can express aspects relevant to the PLP when asked to estimate and compare probabilities of CSEs. Two ways of reasoning identified in the current study are area-based part-whole reasoning and lowest-chance reasoning.

5.1. Implications for research and practice

As described earlier, students' difficulties with taking into account the role of ordered outcomes when handling CSP are well-known (Alston & Maher, 2003; English & Watson, 2016; Konold & Kazak, 2008; Nilsson, 2007; Speiser & Walter, 1998). What the present study does is to shed new light on how such difficulties can be related to how students contextualize the composition of two spinners within the realm of an area context. Based on such a contextualization of CSE, it is understandable why students do not reflect on or take into account the order of the single outcomes of a CSE. In an area context, the order of the single contributions to the CSE does not become an issue. If the main goal is to construct a common whole for a part-whole decision, it does not matter in which order the different areas constituting the whole are put together. In other words, the contextualization gives no reasons to reflect on aspects of the Laplace perspective—generation of sample space built up from ordered pairs—since area-based reasoning has such strong impact on their reasoning.

Related to the previous comments, our study also shows a new case of how students may over-generalize a one-step transformation of the two-step structure implied in a CSE (Lysoe, 2008; Shaughnessy & Ciancetta, 2002). In the case of two spinners, the generalization is based on a contextualization where the two spinners are merged into one whole. The two spinners are not kept separated (Watson & Kelly, 2004). Instead, the areas of the two spinners are put together forming one whole, one big common area, and the situation is solved with the same part-whole procedure, as in the case of a single stochastic phenomenon.

The area-based part-whole reasoning was prominent in task 2, where the students worked on an asymmetric CSE. However, it was interesting to note that many students also applied such reasoning to task 1, where the CSE was build up by two uniform spinners. This was interesting because task 1 is a type of task that Norwegian students are quite familiar with from school. They are familiar with the task construction and educated in applying combinatorial strategies to structure and solve this kind of task. However, only five of the 18 students structured task 1 in a combinatorial sense. This finding is in line with the observation of Ron, Dreyfus, and Hershkowitz (2017), that students are able to apply combinatorial strategies, like the use of two-dimensional representations, as means for structuring the sample space and favorable outcomes of CSE. However, similar to their findings, our study confirms that, when tools for combinatorial reasoning are not directly accessible to the students in a task presentation, it is not likely that students see the use of such tools by themselves. More generally, this finding emphasizes the challenge for teaching to "avoid *inert knowledge*—knowledge students have learned to reproduce but cannot use effectively" (Bakker & Derry, 2011, p. 6). Although the area-based part-whole reasoning described above concerns some of the difficulties students have with solving CSE tasks, the lowest-chance reasoning and intuitive multiplicative response are examples of the potential students have in solving CSE tasks. Both these ways of reasoning relate to PLP, which is an alternative way of solving CSE tasks. Students using such a way of reasoning are able to see the two random objects in the two-spinner CSE and reflect on the relationship and impact each random object has for the whole situation. This way of reasoning is in contrast to the way students with area-based part-whole reasoning look at the CSE. What we see in our data are just glimpses of what students can do, so it would be interesting to see how these students reason after a well-designed teaching program with similar tasks, for example in a learning environment where they can carry out experiments with different setups with two spinners.

Tirosh and Stavy (1999) report that students often are reasoning along the intuitive rule "Same X – Same Y", where they connect two quantities that are not related in the general case. One example of this rule is how students perceive the relationship between amount of sugar and sweetness; if students were asked about what will happen to the sweetness of the water if one teaspoon of sugar is added to a full cup of water and another to a half-full cup of water, many students argued that the sweetness was the same in both cups because it was the same amount of sugar. We argue that area-based part-whole reasoning provides a new case of this intuitive rule, in how students connect the total area of red to the compound probability of “red on both spinners”. More concretely, in task 3 the two setups display the same area of red and, based on this sameness, many students conclude the CSE ‘two red’ is the same between the two setups. Tirosh and Stavy (1999) claim it is important to identify such intuitive knowledge as a source for further teaching, rather than avoiding such reasoning. They suggest that students should meet situations where “Same X – Same Y” is proper, and compare these with situations where this is not the case. From such a perspective on teaching and learning, the area-based reasoning identified in our study should be seen as a resource for curriculum design. Following up the ideas suggested by Tirsoh and Stavy, students could work with pairs of spinners to explore the relationship between total area of the combined spinner and the probability for the combined result "red on both spinners”. Such an exploration would lead in the direction of PLP rather than use of the Lapacian perspective.

That students were inclined to apply area-based part-whole reasoning to CSE may indicate aspects of school-culture. Similar to probability teaching in many countries (Langrall et al., 2017) the teaching of probability in Norway stresses the Laplace interpretation of probability. Following the Laplace perspective, teaching lets students practice on tasks where the probability is determined by calculating the ratio between the number of favorable outcomes out of all possible outcomes. The part-whole, ratio model to probability is practiced through many tasks in school. It is well grounded in tasks concerning single stochastic phenomena, like the throwing of one die or one coin. It is reinforced in the teaching of CSE, like the throwing of two dice or two coins. So, even if many students do not express full-fledged combinatorial reasoning in the present study, their area-based reasoning to a great extent involves part-whole ratio reasoning. More specifically, in area-based part-whole reasoning in the two-spinner situation, the two spinners are merged into a whole and the probability of “two red” is determined by how big part of this whole is red. That teaching seems to encourage students to seek, or be expected to seek, for a relationship between parts and whole, in order to construct a ratio for the calculation of probabilities, raises questions of how the teaching may restrict students from exploring other aspects of CSE and, in particular, how it can have consequences for the teaching and learning of structuring and understanding CSE by means of the PLP. On this account we invite future research to explore further how teaching may manifest part-whole ratio modeling to probability and, what consequences such teaching may have for the teaching and learning of structuring of CSE from a multiplicative angle. Finally, we want to offer suggestions on how our study can provide implications for task design (Watson & Ohtani, 2015). In task 3 area-based part-whole reasoning was prominent and, in task 4 several students gave priority to lowest-chance reasoning. However, in task 5, students, adopting to lowest-chance reasoning in task 4, returned to area-based part-whole reasoning. Hence, taking together task 3, 4 and 5, our results show how students can shift between area-based part-whole reasoning and lower chance reasoning and, how area-based part-whole reasoning may have a kind of priority over lowest-chance reasoning. The result shows the role task design in cueing students’ contextualization of CSE and indicates that the task combination presented by task 3, 4 and 5, can be useful in explicating students’ struggle with CSE and in confronting students with different ways of reasoning about CSE and with, eventually, lack of consistency in different ways of reasoning in situations of CSEs. More generally, our study points to the need for more research to give a more thorough account of how students contextualize and reason in situations where task-combinations are made up of comparison tasks with equal and non-equal areas of the target outcomes.

5.2. Limitations and further research

In the current study students in 9th grade responded to different CSE tasks. Two types of tasks were used; likelihood tasks and comparison tasks. What we observed was that the success rate was higher for the comparison tasks compared to the likelihood tasks. Does this imply that comparison tasks in general are more intuitive and easier than likelihood tasks and so, that teaching in probability should give priority to comparison tasks in students’ early probability learning? Based on the result of the current study we are not in the position to make such a general claim because the difference in success rate may be due to differences in task presentation. More specifically, in the likelihood tasks the students were asked to come up with a numerical value of a probability, which was not the case in the comparison tasks. In the comparison tasks it was enough to offer qualitative or intuitive responses. So, based on this insight, we invite research to take a close look at how different tasks constructions and, particularly, how different kinds of answers asked for, can support or hinder students’ sense making of CSE. The present analysis of how students make sense of CSE with two spinners is based on clinical interviews. Although clinical interviews can provide data for fine-grained analysis of students’ reasoning about CSE, the power relationship between the researcher and the student can also be critical for how students act in an interview situation. On this account, we can ask how the asymmetry between the student and the researcher influenced, or even hindered, students in their reasoning about CSE in the present study. Due to an implicit contract (Brousseau, 1997), we can ask whether the students actually were articulating their own, genuine way of reasoning or, if they were mainly focused on what they thought the teacher expected them to say or do (Millar, Leach, & Osborne, 2001). Hence, in order to provide a more complete picture of how students make sense of CSE and, particularly, apply aspects of PLP in such sense-making, we encourage research to triangulate interview data with data gained from authentic teaching situations, where students engage in CSE in the interplay between individual work, group-work and whole-class discussions.

Acknowledgements

We would like to thank William Gray for proofreading and assistance with our English.

References

Alston, A., & Maher, C. A. (2003). Modeling outcomes from probability tasks: Sixth graders reasoning together. N. A. Pateman, B. J. Dougherty, & J. Zilliox (Eds.). *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education (jointly with PME-NA) Vol. 2,* 25–32.

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

Brousseau, G. (1997). *Theory of didactical situations in mathematics.* N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Trans.Kluwer: Dordrecht.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). London: Sage.

Chernoff, E. J., & Zazkis, R. (2011). From personal to conventional probabilities: From sample set to sample space. *Educational Studies in Mathematics, 77*(1), 15–33.

Cobb, P. (2007). Putting philosophy to work. Coping with multiple theoretical perspectives. In F. K. Lester (Ed.). *Second handbook of research on mathematics teaching and learning* (pp. 3–38). Greenwich, CT: Information Age Publishing.

Confrey, J., & Smith, E. (1995). Splitting, covariation, and their role in the development of exponential functions. *Journal for Research in Mathematics Education, 26*(1), 66–86.

English, L. D., & Watson, J. M. (2016). Development of probabilistic understanding in fourth grade. *Journal for Research in Mathematics Education, 47*(1), 28–62.

Fischbein, E., Nello, M. S., & Marino, M. S. (1991). Factors affecting probabilisticjudgements in children and adolescents. *Educational Studies in Mathematics, 22*(6), 523–549.

Fischbein, E., Pampu, I., & Minzat, L. (1975). *The child's intuition of probability.* Appendix II in Fischbein, E. (1975)*The intuitive sources of probabilistic thinking in children.* Reidel: Dordrecht.

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics, 1*(3), 4–11.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York: Aldine de Gruyter.

Green, D. R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.). *Proceedings of the first International Conference on teaching statistics* (pp. 766–783). Sheffield, UK: Teaching Statistics Trust.

Halldén, O. (1999). Conceptual change and contextualization. In W. Schnotz, S. Vosniadou, & M. Carretero (Eds.). *New perspectives on conceptual change* (pp. 53–65). Oxford: Pergamon, Elsevier Science.

Hung, D. W. L. (1997). Meanings, contexts, and mathematical thinking: The meaning-context model. *The Journal of Mathematical Behavior, 16*(4), 311–324.

Iversen, K. (2003). *Flexitree: Software.* Nord-Trøndelag University College, Faculty of Education, Engineering and Nursing.

Iversen, K., & Nilsson, P. (2007). Students' reasoning about one-object stochastic phenomena in an ICT-environment. *International Journal of Computers for Mathematical Learning, 12*(2), 113–133.

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. Lester (Ed.). *The second handbook of research on mathematics* (pp. 909–956). Reston, VA: National Council of Teachers of Mathematics (NCTM).

Konold, C. (1996). Representing probabilities with pipe diagrams. *Mathematics Teacher, 89*(5), 378–382.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics, 2*(1).

Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing.* Thousand Oaks, CA: Sage.

Langrall, C., Makar, K., Nilsson, P., & Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: An integrated perspective. In J. Cai (Ed.). *Compendium for research in mathematics education* (pp. 490–525). Reston, VA: National Council of Teachers of Mathematics (NCTM).

Langrall, C., & Mooney, E. S. (2005). Characteristics of elementary school students'probabilistic reasoning. In G. A. Jones (Ed.). *Exploring probability in school: Challenges for teaching and learning* (pp. 95–119). New York: Springer.

Lecoutre, M.-P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics, 23*(6), 557–568.

Lysoe, K. O. (2008). Strength and limitations of informal conceptions in introductory probability courses for future lower secondary teachers. In M. Borovcnik, D. Pratt, Y. Wu, & C. Batanero (Eds.). *Research and development in the teaching and learning of probability* (pp. 1–14). Monterrey: International Association for Statistical Education.

Millar, R., Leach, J., & Osborne, J. (2001). *Improving science education: The contribution of research.* Buckingham: Open University Press.

Nilsson, P. (2007). Different ways in which students handle chance encounters in the explorative setting of a dice game. *Educational Studies in Mathematics, 66*(3), 293–315.

Nilsson, P. (2009). Conceptual variation and coordination in probability reasoning. *The Journal of Mathematical Behavior, 28*(4), 247–261.

Nilsson, P., & Ryve, A. (2010). Focal event, contextualization, and effective communication in the mathematics classroom. *Educational Studies in Mathematics, 74*(3), 241–258.

Polaki, M. V. (2005). Dealing with compound events. In G. A. Jones (Ed.). *Exploring probability in school: Challenges for teaching and learning* (pp. 191–214). New York: Springer.

Ron, G., Dreyfus, T., & Hershkowitz, R. (2017). Looking back to the roots of partially correct constructs: The case of the area model in probability. *The Journal of Mathematical Behavior, 45*, 15–34.

Ryve, A., Nilsson, P., & Patterson, K. (2013). Analyzing effective communication in mathematics group work: The role of visual mediators and technical terms. *Educational Studies in Mathematics, 82*(3), 497–514.

Schmidt, S., & Weiser, W. (1995). Semantic structures of one-step word problems involving Multiplication or Division. *Educational Studies in Mathematics, 28*(1), 55–72.

Schou, J., Jess, K., Hansen, H. C., & Skott, J. (2013). *Matematik for lærerstuderende - Stokastik. 1.-10. klasse.* Fredriksberg: Samfundslitteratur.

Säljö, R., & Wyndhamn, J. (1990). Problem-solving, academic performance and situated reasoning. A study of joint cognitive activity in the formal setting. *The British Journal of Educational Psychology, 60*(3), 245–254.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge, UK: Cambridge University Press.

Shaughnessy, M., & Ciancetta, M. (2002). Students' understanding of variability in a probability environment. B. Phillips (Ed.). *Developing a Statistically Literate Society, Proceedings of the 6th International Conference on the Teaching of Statistics, Cape Town, South Africa)*. Retrieved from https://iase-web.org/documents/papers/icots6/6a6_shau.pdf.

Speiser, R., & Walter, C. (1998). Two dice, two sample space. In L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee, & W.-K. Wong (Vol. Eds.), *Proceedings of the fifth Conference on Teaching of Statistics: Vol. 1*, (pp. 67–73). Voorburg, the Netherlands: International Statistical Institute Permanent Office.

Starrin, B., Dahlgren, L., Larsson, G., & Styrborn, S. (1997). *Along the path of discovery. Qualitative methods and grounded theory.* Stockholm: Studentlitteratur.

Tirosh, D., & Stavy, R. (1999). Intuitive rules: A way to explain and predict students' reasoning. *Educational Studies in Mathematics, 38*(1-3), 51–66.

Vidakovic, D. (1998). Children's intuition of probabilistic concepts emerging from fair play. In L. S. K. L. Pereira-Mendoza, T. Wee Kee, & W.-K. Wong (Vol. Eds.), *Proceedings of the Fifth Conference on teaching of statistics: Vol. 1*, (pp. 67–73). Voorburg, The Netherlands: International Statistical Institute Permanent Office.

von Wright, G. H. (1971). *Explanation and understanding.* London: Routledge and Kegan Paul.

Watson, A., & Ohtani, M. (Eds.). (2015). *Task design in mathematics education: An ICMI study 22*. Berlin: Springer.

Watson, J., & Kelly, B. (2004). Expectation versus variation: Students' decision making in a chance environment. *Canadian Journal of Science Mathematics and Technology Education, 4*(3), 371–396.

Zazkis, R., & Hazzan, O. (1998). Interviewing in mathematics education research: Choosing the questions. *The Journal of Mathematical Behavior, 17*(4), 429–439.