Comparing probabilities in urns: A study with primary school students

Comparación de probabilidades en urnas: Un estudio con estudiantes de Educación Primaria

Comparação de probabilidades nas urnas: um estudo com estudantes do ensino fundamental

Luis Armando Hernández-Solís[2] • **Carmen Batanero**[1] • **María M. Gea**[1] • **Rocío Álvarez-Arroyo** [1]

Abstract

The study aimed to explore the strategies used by Costa Rican primary school students when comparing urn probabilities. The sample was intentional and consisted of 55 6th graders. Using an interpretive approach, we analyzed the children's responses to a questionnaire of five probability comparison items taken from previous studies, including different levels of proportional reasoning. Results indicate that problems in the first levels of proportional reasoning were solved easily by students, while problems at higher levels increased in difficulty. One-variable strategies, which compare only the favorable or unfavorable cases in both urns, are predominant; although there are correspondence strategies, few students show complete proportional reasoning. Aside from slight variations, results are similar to previous studies, which suggests that this task is influenced more by the child's level of maturity than by the instruction received.

Keywords: Probability comparison; proportional reasoning; primary education; statistical education; urns; Costa Rica.

Resumen

El objetivo del estudio fue explorar las estrategias que emplean niños y niñas de 6º curso de educación primaria costarricenses al comparar probabilidades en contextos de urnas. La muestra participante fue intencional y estuvo formada por 55 estudiantes de primaria. La investigación tiene un enfoque interpretativo, donde se analizan sus respuestas a un cuestionario formado por cinco ítems de comparación de probabilidades, tomados de investigaciones previas, que tienen en cuenta diferentes niveles de razonamiento proporcional. Los resultados indican que el estudiantado de la muestra resuelve con facilidad los problemas que corresponden a los primeros niveles de razonamiento proporcional, y aumentan su dificultad en los niveles superiores. Predominan las estrategias de una variable, en que se comparan solo los casos favorables o desfavorables de las dos urnas, y aunque aparecen estrategias de correspondencia,

Luis Armando Hernández-Solís ✉ lhernandezcr@correo.ugr.es, https://orcid.org/0000-0003-2956-8102
Carmen Batanero ✉ batanero@ugr.es, https://orcid.org/0000-0002-4189-7139
María M. Gea ✉ mmgea@ugr.es, https://orcid.org/0000-0002-5229-0121
Rocío Álvarez-Arroyo ✉ rocioaarroyo@ugr.es, https://orcid.org/0000-0002-3201-8542
1 Department of Mathematics Education, Universidad de Granada, Granada, Spain.
2 Faculty of Exact and Natural Sciences, Universidad Estatal a Distancia, San José, Costa Rica.

es poca la cantidad de estudiantes que muestran un razonamiento proporcional completo. Los resultados son similares, con ligeras variaciones, a los estudios previos, lo que indica que en esta tarea influye más la maduración del alumnado que la enseñanza recibida.

Palabras clave: Comparación de probabilidades; Razonamiento proporcional; Educación primaria; Educación Estadística; Urnas; Costa Rica.

Resumo

O estudo teve como objetivo explorar as estratégias aplicadas por estudantes do 6° ano do ensino fundamental costarriquenho ao comparar probabilidades no contexto das urnas. A amostragem participativa foi intencional e composta por 55 meninos e meninas do ensino fundamental. A pesquisa tem uma abordagem interpretativa, em que são analisadas as respostas desse público em um questionário conformado por cinco itens de comparação de probabilidades, coletados de pesquisas prévias que levam em consideração diferentes níveis de raciocínio proporcional. Os resultados indicam que os estudantes da amostra solucionam facilmente problemas que correspondem aos primeiros níveis de raciocínio proporcional, e aumentam a dificuldade em níveis superiores. São predominantes as estratégias de uma variável, nos quais são comparados somente os casos favoráveis ou desfavoráveis das duas urnas e, ainda que aparecem estratégias de correspondência, é pouca a quantidade de estudantes que demonstram um raciocínio proporcional completo. Os resultados são semelhantes, com ligeiras variações, aos estudos prévios, indicando que nesta tarefa a maturidade da infância influencia mais do que o ensino recebido.

Palavras-chave: comparação de probabilidades; raciocínio proporcional; ensino fundamental; educação estatística; urnas; Costa Rica.

INTRODUCTION

At present, the teaching of probability acquires great relevance as being necessary to develop a probabilistic culture in citizens (Batanero, 2006; Gal, 2005). In addition to being a relevant part of mathematics and applicable to other curricular topics, probability is necessary today in many fields of science, where it allows describing the laws governing random phenomena (Borovcnik, 2011).

An important and recognized consequence of the above is the inclusion of probability content in the primary education curricula in countries such as Spain (Ministerio de Educación, Cultura y Deportes, 2014) or the United States (National Council of Teachers of Mathematics,

2000). In Costa Rican schools, current mathematics programs grant greater value to statistics and probability (MEP, 2012). Specifically, the following expectations are indicated about what students must learn throughout primary education in relation to probability (MEP, 2012):

- *First cycle* (*Primer ciclo*) (1rst to 3rd school year/grade): Identify random and deterministic situations within daily life and associated events. Classify random events as more or less probable for particular situations or experiments. Identify events according to the simple results associated with them (p. 147).
- *Second cycle* (*Segundo ciclo*) (4th to 6th school year/grade): Identify more

probable, less probable or equally probable events according to the number of simple results belonging to each event. Determine elemental probabilities associated with particular events. Pose and solve problems related to random situations (p. 247).

It is important, then, to analyze whether children have the necessary skills to successfully address such content, as well as their possible difficulties, so that educators can take them into account in their educational planning. Until now, no investigation has been carried out on the probabilistic reasoning of Costa Rican pupils of this age range, and although the existence of previous studies in other countries is documented, these were done with students without instruction in this subject matter. This takes great relevance in the current curricular implementation context, where specifically, probability content is given a great deal of weight within primary education. Thus, we consider that the present work can provide initial information regarding the personal meaning that Costa Rican children assign to the probabilistic concepts developed since the approval of the current curriculum.

Consequently, the objective of this exploratory study is to provide information on the strategies that children apply when comparing probabilities in two urns and the difficulty that this task has for this age group depending on the level of proportional reasoning required. We focus on students in the 6th grade of primary education (ages 11-12 years) as this is the school year where primary education ends. Furthermore, the participating students have followed the current mathematics curriculum, which was fully applied to all primary education levels as of 2015. A second objective is to compare our results with those of previous research on the subject matter.

THEORETICAL FRAMEWORK

Our work is based mainly on the studies of Piaget and Inhelder (1951) on the comparison of probabilities. Piaget (1975) describes the process of childhood learning through action and assimilation-accommodation. When children face a mathematical problem, they try to solve it (action) using the knowledge they already possess, using existing conceptual schemes that allow them to anticipate and use strategies and representations that they already know. If the child is unable to resolve the problem, a cognitive conflict is presented that he or she addresses through the processes of assimilation and accommodation. Assimilation consists of the incorporation (acceptance), by the subject, of new data or ideas, and accommodation is the change or restructuring of existing ones. The aforementioned authors suggest that knowledge progresses in stages of development that have an established order, although the precise age at which a specific stage is reached may vary. Piaget (1975) also indicated that the subjects who are in the same stage experience similar reasoning.

To study the ability and reasoning of children when comparing simple probabilities, Piaget and Inhelder (1951) conceived an experimental game employing white tokens, marked or not with a cross, of which they introduced a small number into transparent urns (boxes). Children were then asked to select between two urns the one from which they considered more likely to draw a marked token. The authors then continuously changed the number of marked

(favorable event) and unmarked (white) tokens (unfavorable event) in both urns and conducted interviews with students whose age ranged from 3.5 to 13-14 years. By comparing similar responses among same-age groups, they obtained a description of three stages in the development of their reasoning pertaining to these types of situational problems.

The first stage (I) is subdivided into two levels. Level IA is characterized as comprised of an age group where children lack the logical schemes that allow them to understand: a) the inclusion of the part to whole; b) the disjunction between two types of elements; and c) the conservation of quantities (for example, when both types of tokens are selected and separated from each other). For this reason, Level IA children are only capable of solving comparison problems involving two probabilities such as double impossibility (all tokens in both urns are white), double certainty (all tokens are marked) or certainty-impossibility (one urn containing white tokens only while the other urn only marked tokens). This is explained by acknowledging that these children only compare favorable events without taking into account all possible events. At level IB, only one type of token is compared (favorable or unfavorable), and children are yet incapable of conceiving favorable events as part of all possible events (comparison of the *part to whole*); neither are they able to compare favorable to unfavorable events (part-part comparison). However, level IB children begin to understand that the probability depends on the number of favorable or unfavorable events.

The second stage (II) is also divided into two sublevels. At level IIA, children can solve probability comparison problems that involve only one variable, that is, when only favorable or unfavorable events need to be compared. They use additive comparisons, for example, subtracting the number of favorable from unfavorable events, or vice versa, in each urn and comparing the differences. They commence to understand the disjunction (each event is favorable or unfavorable), but provide wrong solutions in cases in which the composition of favorable and unfavorable events in both urns is proportional, since at this age they have not acquired the idea of fraction or proportion. At level IIB children begin to solve the problem when the composition in the urns is proportional. To do this, they establish a correspondence between the favorable and unfavorable events resulting from one of the urns (for example, there are two favorable events for each unfavorable one) and compare it with the existing correspondence in the other urn.

In stage III, children are able to easily solve the proportionality case and manage to think of a general solution, if the number of favorable and unfavorable events is small and the ratio between them is simple, for example, double, triple, etc. A solution is generalized with advancement of age upon acquiring sufficient knowledge of fractions.

In the previous exposition it is observed that success in comparing probabilities, in the general case, supposes an adequate proportional reasoning, whose developmental stages have been analyzed in several studies (e.g, Karplus, Pulos and Stage, 1983; Noelting, 1980a; 1980b) and summarized by Behr, Harel, Post, and Lesh (1992) and Ben-Chaim, Keret, and Ilany (2012). The most relevant reference work for our study is that of Noelting (1980a; 1980b), who based on a problem of comparing two mixtures (water and orange juice), analyzed the stages initially proposed by

Piaget and Inhelder (1951) for comparison of probabilities and extends them to problems of proportionality. Concluding from his work, Noelting considers an initial stage (occurring at 2 years of age) in which the elements of the fraction are identified; and also, further subdivides some of the stages described by Piaget and Inhelder (1951):

- The intuitive stage is subdivided into three levels: lower (IA), middle (IB) and higher (IC), according to the terms of the fraction being compared (first term; second term with the same first term; or an internal relationship is established between terms, either from one fraction with the other fraction or between fractions, respectively).
- The concrete operational stage is subdivided into two levels, according to the equivalence class of the fractions being compared, whether of ratio 1 (stage IIA) or any other class of equivalence (stage IIB).
- The formal operational stage is subdivided into two levels, according to the proportionality between the terms of the fraction. The lower level is determined when the terms are multiples (IIIA) and the higher level in cases where a ratio exists (IIIB).

Another work that will be considered is that of Pérez Echeverría, Carretero and Pozo (1986), who adapted the tasks of Noelting (1980a; 1980b) and applied them to 20 students of basic general education (13 years old) and 20 high school students (16 years old). They defined different levels of difficulty in the problems, according to the required strategy:

- Level 1: problems where the number of favorable events or unfavorable events, in both urns, is the same; therefore, the use of fractions is unnecessary to solve the problem.
- Level 2: problems where proportionality between favorable and unfavorable events in the same urn or between favorable and unfavorable events in both urns exist. They can be solved by establishing a correspondence in one urn and observing that the relationship is the same in the other urn.
- Level 3: problems that present proportionality only between favorable events in both sets (or urns in our case) or only between unfavorable events, or between favorable and unfavorable events within a single set. Once the ratio between favorable or unfavorable events has been established, it can be compared with the ratio existing between other terms and determined if the value is lower or higher.
- Level 4: problems without proportionality relationships among the four members (favorable and unfavorable events in each set). It requires operating with fractions, converting them to a common denominator.

Background

Piaget and Inhelder's (1951) research inspired a series of works on probabilistic reasoning in children which are described in detail in Jones, Langrall and Mooney (2007) and Langrall and Mooney (2005). Next, we summarize those most relevant to our study.

Falk, Falk and Levin (1980) presented children between the ages of 4 and 11 with the task of comparing probabilities by varying the number of favorable and possible events, under two contexts: urns

and roulettes. They considered nine types of tasks, taking into account the following variables: a) the number of favorable events is less, greater or equal in the set with the highest probability; b) the number of favorable events is less, greater or equal in the set with the lowest probability; c) both sets are equiprobable and the number of favorable events is less, greater or equal in a set. A systematic error was to always choose the set with the most favorable events.

Truran (1994) conducted an investigation with 32 students between the ages of 8 and 15 also comparing the probability in urns. His results identify new strategies that expand those described in the research by Piaget and Inhelder (1951), such as the description of the urn's content without making a selection, giving a correct answer without justification, preference for the smallest number of the total of tokens, comparison with known simple proportions and comparison between ratios of favorable and unfavorable possibilities.

The most relevant study on the subject was carried out by Green (1982), who made an evaluation of probabilistic reasoning in English children aged 11 to 16 years with a test that reproduced, with paper and pencil, the experiments of Piaget and Inhelder (1951). Some items proposed the comparison of probabilities in urn contexts, in which they found the following strategies: a) choose the urn with the greatest number of possible events; b) choose the urn with the greatest number of favorable events; c) choose the urn that provides the greatest difference between favorable and unfavorable events; d) choose the urn with the highest proportion between favorable and unfavorable events.

Cañizares (1997) undertook a study with 320 Spanish students aged 10 to 14, and among other problems, proposes those of comparing probabilities in the context of urns. Cañizares and Batanero (1997) in a study with 134 students aged 10 to 14 considered tasks corresponding to levels I to IIIB described by Noelting (1980a; 1980b). They registered the strategies used by children, classifying them into one- and two-variable strategies. One-variable strategies are those in which only favorable, unfavorable or possible events are compared; two-variable strategies are at hand when favorable and possible events are compared in an additive or multiplicative manner. Cañizares (1997) deduces that the most frequent levels of reasoning in sampled students belong to IB through IIB, with few reaching IIIB. Some variables that influence the response are the composition of the urns (number of favorable and possible events) and the existence of possible biases in the context (for example, beliefs about equiprobability).

METHODOLOGY

The approach of the investigation is interpretive, since it focuses on understanding educational phenomena (in our case the probabilistic reasoning in children) through the analysis of quantitative and qualitative elements reflected in the responses to a questionnaire (Cerrón, 2019; Gil, León, and Morales, 2017). The research work is exploratory, since the sample is intentional and moderate in size; and according to Bisquerra (1989), it conforms to applied research, since it seeks to use the theory developed by other investigations (Cañizares, 1997; Fischbein and Gazit, 1984; Green, 1982) in the context of Costa Rica, in order

to provide knowledge to guide the teaching and learning of probability at the primary education level.

The sample used was non-random (with intentional selection) and consisted of 55 children in the 6th grade of primary school with the following characteristics: 40 children were 11 years old and 15 were 12 years of age; 29 were studying in a private institution and 26 in a public institution located in the province of Cartago, Costa Rica. The mathematics teachers indicated that the pupils received statistics and probability education since 2016 in accordance with MEP's (2012) study programs. The treatment of the subjects was based on the official textbook, and there is shortage of evidence suggesting activities being carried out using physical experimentation.

A questionnaire was designed (Annex), based on some items employed in Green's (1982) research, which Cañizares (1997) also used. It is intended to use items taken from validated questionnaires and to compare the probabilistic reasoning of Costa Rican pupils with that of students at a similar educational level in other countries. Green (1982) first conducted a study of the validity of the content of his questionnaire, with the help of the validation carried out by expert teams and analysis of included items. He also analyzed the reliability of the instrument using Cronbach's alpha coefficient (α), for which he obtained a value of $\alpha=0.88$.

Cañizares (1997) performed a factorial analysis of the responses of the sampled population to Green's questionnaire, and achieved a total of 15 factors. Therefore, she calculated Carmines' theta reliability coefficient (θ), which is preferable to Cronbach's alpha coefficient (α) when factorial analysis data, as in her case, show a multidimensional structure; her result was $\theta=0.8242$. All items propose the use of urn experiments employing black and white tokens, and request is made for selecting one of two given urns wherein the event "to draw a black token" is more likely; the composition of the urns is different among items.

Item 1 corresponds to the lower intuitive level (IA) described by Noelting (1980a; 1980b), because both urns contain the same number of unfavorable events and an unequal number of favorable events. Item 2 accommodates equality of favorable events and inequality of unfavorable events, thus corresponding to the middle intuitive level (IB). Therefore, in these two items it is unnecessary to use the four data of the statement, as it suffices to compare or to know the number of favorable (item 1) or unfavorable (item 2) events. Both items correspond to the first level of difficulty described by Pérez Echeverría et al. (1986) as they do not require proportional reasoning.

In items 3 to 5, the number of favorable events is a multiple of the number of unfavorable ones. In item 3 (lower concrete operational level (IIA), according to Noelting 1980a; 1980b), the number of favorable and unfavorable events is the same in both urns; in item 5 (higher concrete operational level (IIB)), the ratio between favorable and unfavorable events present in both urns is 3; and in item 4 (lower formal operational level (IIIA)), the ratio in one urn is 3 while in the other is 2. In these last three items it is necessary to use the four data of the statement as well as proportional reasoning to establish a ratio in one of the urns and compare it with the other. Items 3 and 5 are assigned difficulty level 2 as per Pérez Echeverría et al. (1986) classification, since proportionality exists between favorable and unfavorable events in each urn or between favorable and

unfavorable events in both urns. Item 4 is classified in level 3 because there is a simple relationship in the first urn (3 favorable events for each unfavorable event) which can be compared to the relation existing in the second urn (two to one). One difference in our study, compared to those of Green (1982) and Cañizares (1997) is that we provided the graphical representation of the urns in items 4 and 5, which was not given in the referenced studies.

Table 1 summarizes the classification of items according to the level of proportional reasoning required for their resolution as established by Noelting (1980a; 1980b), and their level of difficulty according to Pérez Echeverría et al. (1986). It also presents the composition of the two compared urns, (a_1, b_1) and (a_2, b_2), with "a" and "b" being the favorable and unfavorable events, respectively, in the proposed experiments of items 1 through 4.

We consider that Table 1 will help us analyze the strategies employed by children, confronting them with what is expected from a problem of comparison of fractions of the same level of difficulty (Pérez Echeverría et al., 1986) in Noelting's categorization (1980a; 1980b).

ANALYSIS AND RESULTS

Questionnaire responses were pooled and analyzed through a content analysis (Krippendorff, 2013), which allows us to establish categories that emerge objectively as a result of the systematic analysis performed. This analysis is complemented with numerical information in tables indicating the percentage of correct answers and percentages of each strategy in each item.

Table 2 exhibits the percentages of correct answers obtained in the items of the questionnaire; it also considers the results

Table.1. *Stages of proportional reasoning* (Noelting, 1980a; 1980b) *and corresponding level of difficulty* (Pérez Echeverría et al., 1986) *required per item.*

Stage	Proportional reasoning level	Item	Composition (a_1,b_1) vs (a_2,b_2)	Difficulty level
IA	Lower intuitive	1	(3,1) vs (2,1)	1
IB	Middle intuitive	2	(5,2) vs (5,3)	1
IIA	Lower concrete operational	3	(2,2) vs (4,4)	2
IIB	Higher concrete operational	5	(3,1) vs (6,2)	2
IIIA	Lower formal operational	4	(12,4) vs (20,10)	3

Table.2. *Percentage of correct answers in the present work compared to those obtained by* Cañizares (1997) *and* Green (1982) *for 6*[th] *grade students (ages 11-12 years).*

Item N°	Costa Rica (Present study)	Spain (Cañizares, 1997)	United Kingdom (Green, 1982)
1	83.6	70.3	88.0
2	72.7	67.6	55.0
3	41.8	54.1	43.0
4	50.9	27.0	38.0
5	16.4	-	20.0

obtained with students of the same age in previous investigations: Cañizares (1997) in Spain and Green (1982) in the United Kingdom. These are simple probability comparison items in the context of urns, where the only variable is urn composition.

Items 1 and 2 were relatively straightforward for the study subjects, as more than two-thirds of pupils performed the task correctly. These items are of difficulty level 1 as per Pérez Echeverria et al. (1986) and are assigned levels IA (item 1) and IB (item 2) according to Noelting's (1980a; 1980b). These items can be solved correctly just by comparing the favorable or unfavorable cases (one-variable strategies). It can be stated that most of the study subjects have reached these first levels of reasoning.

Item 4, corresponding to proportional reasoning level IIIA (Noelting, 1980a; 1980b) and difficulty level 3, according to Pérez Echeverría et al. (1986), exhibited moderate difficulty as only half of the students (50.9%) solved it correctly, thus only half the pupils reached this level of reasoning. The items with the greatest level of difficulty (difficulty level 2, according to Pérez Echeverría et al. (1986)) were 3 and 5; the latter item (level IIB, according to Noelting 1980a; 1980b) was only answered correctly by 16.4% of the students.

There is an inversion in the expected difficulty in item 5, which theoretically should be easier than item 4 as expected from the classification promoted by Pérez Echeverría et al. (1986) (Table 1); however, experimental results demonstrate otherwise. We believe that for children it has been easier to compare the equality between favorable (item 2) and unfavorable (item 1) events. The comparison of two similar ratios (3/1 and 6/2, item 5) has been made difficult by the fact that the set has also taken into account the number of favorable events, which is greater in the second urn. When compared to results from previous studies, the level of difficulty of each item was similar, although our results in items 1, 2 and 4 were slightly better than those obtained by Cañizares (1997) and markedly better than those reported by Green (1982) for items 2 and 4. Item 3 results were worse than those of Cañizares (1997) but similar to Green's (1982); item 5 was found very difficult, also in the previous research.

Student strategies

From the analysis of the students' arguments when justifying their answers, resolution strategies were classified based on the classification established by Cañizares and Batanero (1997). These strategies were as follows:

A. Comparison of possible events. Although circumstantially correct answers could be generated, it lacks a logical basis and originates from the impossibility of comparing the total set with a subset (part-whole).

E38: "There are more tokens than in E" (answer B, item 3).

B. Comparison of favorable events. Selecting the urn with the greatest favorable events. Correct answers are generated when there is equality of unfavorable events, as in item 1.

E8: "Because there are more black balls in A than in B" (answer A, item 1).

C. Comparison of unfavorable events. Selecting the urn with the fewest unfavorable events. It represents an advance with

respect to strategy B, as it is recognized that the number of unfavorable events diminishes the probability of winning. Correct answers are generated when there is equality of favorable events, as in item 2.

E54: "Because there is one less white token than in the other urn" (answer A, item 2).

Although there is preference for the previous one-variable strategies, typical of the preoperational stage, other two-variable strategies were used.

D. Additive comparison of favorable and unfavorable events. It consists in comparing the difference between favorable and unfavorable events in both urns. In the example that follows, the difference would be zero.

E16: "Because both urns contain the same amount of white and black tokens" (answer C, item 3).

E. Correspondence. A criterion of proportionality is established in one fraction to be applied in the other. This type of reasoning is typical of a higher level of development and, according to Piaget and Inhelder (1951), is associated with the formal operational period, although it could also appear during the concrete operational period in simpler cases of proportional composition of tokens between urns. Some examples are:

E7: "The same possibility because if we double Ex2, the same result as F is obtained" (answer C, item 3).

E14: "Because urn E contains the same number of black tokens as white tokens and the same for urn F" (answer C, item 3).

In the previously cited examples it is appreciated how pupils establish a criterion of proportionality in one of the urns in order to apply it to the other; this is natural when the individual still lacks the knowledge to perform calculations using fractions. According to Noelting (1980a; 1980b), this type of strategy is associated with stage IIA, where children internally relate the terms of the fraction, differentiating the concepts of ratio and quantity. The author makes the distinction between "within"- and "between"- type strategies to compare two fractions (a_1/b_1 vs a_2/b_2), so it is important to identify that E7 performs a "between" strategy, when comparing the terms of one fraction with those of another (a_1 with a_2 and b_1 with b_2), while E14 uses a "within" strategy, because the individual compares the terms within the same fraction to establish a ratio ($a_1/b_1 = 1/1$) and then compares it with the ratio in the other fraction ($a_2/b_2 = 1/1$).

F. Multiplicative (ratio comparison). The number of favorable events is related to the number of possible events, that is, the *part to whole*, or also fractions formed by the number of favorable and unfavorable events, to later compare them by applying Laplace's rule. Few individuals in the study group use multiplicative strategies, which are undoubtedly the most elaborate and require mastery of calculation with fractions:

E14: "Because urn J has a third as many white tokens as there are black tokens and the same for urn K" (answer A, item 5).

E17: "Because urn H has half as many white tokens as there are black tokens and urn G has one third as many white tokens as there are black tokens" (answer B, item 4).

It can be seen that E14 establishes 1/3 as the fraction of white (unfavorable events) to black (favorable events) tokens and compares

this value to that in the other urn. E17, in item 4, performs a comparison of fractions and relies on the graphical representation shown in Figure 1, setting favorable and unfavorable events to establish more clearly a "within" comparison through a "part-part" relation.
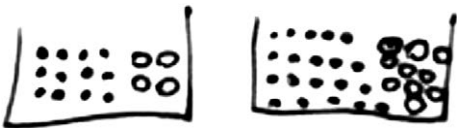

Figure 1. Graphic representation drawn by E17 in item 4.
Source: present investigation.

G. Equiprobability bias. When all the events of a random experiment are considered as equiprobable it is known as "equiprobability bias" (Lecoutre, 1992).

E27: "Because both urns have the same probability of drawing a white token or a black token" (answer C, item 1).

E30: "Any token can be drawn" (answer C, in items 1, 2 and 3).

H. Disposition of tokens. There is a notable percentage of arguments associated with beliefs that relate the spatial arrangement of the tokens in the item's graphic representation with the probability of the event; also identified in Cañizares and Batanero (1997). This occurred in all items for different students:

E52: "It has the black token on all 8 sides" (answer A, item 1).

E39: "Because the black token is also at the top" (answer B, item 3).

E48: "Because the black ball is first and the black ball is much more likely to come out" (answer C, item 5).

I. Other. These strategies arise when individuals express arguments such as attribution to luck or the manner in which the token is drawn from the urn, as registered in the following example:

E32: "But the white token can also be drawn, since it all depends on how you hold your hand" (answer A, item 1).

Table 3 summarizes the strategies involved in the different items. In item 1, just over three-quarters of the individuals made justifications associated with relevant

Table 3. *Percentage distribution of adopted strategies per item (the correct strategies have been underlined).*

Strategy	Items				
	1	**2**	**3**	**4**	**5**
A. Comparison of possible events	3.6	1.8	20.0	12.7	10.9
B. Comparison of favorable events	72.7	21.8	12.7	23.6	32.7
C. Comparison of unfavorable events	5.5	41.8	5.5	10.9	12.7
D. Additive comparison of favorable and unfavorable events	5.5	16.4	32.7	29.1	23.6
E. Correspondence	0.0	0.0	1.8	3.6	7.3
F. Multiplicative (ratio comparison)	1.8	1.8	3.6	1.8	1.8
G. Equiprobability bias	3.6	1.8	1.8	0.0	0.0
H. Disposition of tokens	1.8	10.9	7.3	10.9	7.3
I. Others	5.5	3.6	10.9	7.3	3.6
No response	0.0	0.0	3.6	0.0	0.0

Source: present investigation.

strategies, with strategy B being the most widely used (72.7%) due to the equality of unfavorable events. In Cañizares and Batanero (1997), approximately two thirds of same-age individuals used pertinent arguments and comparison of favorable events was also the most adopted strategy (48.6%), but higher in percentage of two-variable strategies (10.8%).

In item 2, 60% of students used pertinent strategies, where C was the most used (41.8%), which is natural due to the existence of the same number of favorable events. Cañizares and Batanero (1997) obtained a lower percentage of pertinent arguments (54.0%) of which approximately 40% compared the number of unfavorable events, as in our study.

For item 3, 38.1% of pertinent arguments were obtained, a little lower than the 48.6% obtained by Cañizares and Batanero (1997); and while our pupils focused on strategy D (32.7%), those of Cañizares and Batanero (1997) did so on E (43.2%), which requires a higher level of proportional reasoning.

In items 4 and 5, not even 10% of correct strategies were reached, and reasoning focused on one-variable strategies, in the comparison of favorable and unfavorable events, which is not appropriate, because the number of favorable or unfavorable events does not coincide. The same occurred with item 4 in Cañizares and Batanero (1997), where only 2.7% of pertinent strategies were obtained in children of the same age as in our study. It should be remembered that this item presents a higher level of proportional reasoning, corresponding to stage IIIA (Noelting, 1980a; 1980b). The results from Item 5 are not related to Cañizares and Batanero (1997) because they used a different item in their investigation.

CONCLUSIONS

When analyzing the responses of children in the sample regarding the comparison of probabilities in the context of urns, we obtained similar results to those of other previous investigations involving children of the same age and carried out in a historical period where elementary probability content was absent from school curricula. It is true that, at the time of applying the questionnaire, our subject group had not yet received the probability content of their 6th grade course, only the probability curricula from previous school grades. Although they had not studied the quantification of probabilities using Laplace's rule, this is not binding with the success in the items proposed in the questionnaire, since the use of such tool (Laplace's rule) is not necessary to establish the comparison of the requested probabilities, as was the case in previous reported investigations with students of the same age. On the other hand, students were not accustomed to comparing two different experiments as demanded by the proposed tasks, since they had previously only worked on events from the same experiment. Therefore, it would be important to expand our study to include a 7th grade group, to ensure that they have performed similar probability comparison tasks and thus analyze the effect of the instruction.

However, our conjecture is that the greater difficulty of some items is due to the fact that only part of the individuals has reached the corresponding level of proportional reasoning in Noelting's classification (1980a; 1980b). In the 5th grade, Costa Rican children study proper and improper, homogeneous and heterogeneous fractions, and carry out activities to compare fractions, however, the context of probability is often not used to complete their study of fractions and make comparisons.

Our recommendation, then, is to complement the probability content taught in primary education with exercises similar to those disclosed in the items of the present questionnaire, where, first, students should only be presented with one-variable problems requiring first level (items 1 and 2) proportional reasoning. In addition, the study of fractions can be further complemented with "part-part"-type comparisons, as currently "part-whole" comparisons predominate and, also with the application of proportionality to probability comparison situational problems.

Although Piaget and Inhelder (1951) point out that children compare possible events at first (strategy A), this strategy did not appear frequently in our study. We agree with Cañizares and Batanero (1997) who point out that in problems where favorable events are explicitly distinguishable from unfavorable events, students first compare favorable events (strategy B) prior to possible events, as there is an initial "part-part" perception. This is evidenced in all the items of our questionnaire (except item 3), where the percentages of strategy B were higher than those of strategy A. The above finding is clearly reflected in item 1, where almost three quarters of the individuals selected strategy B which only generates correct answers for this item.

It is important, however, that students move on to superior strategies in more elaborate problems. We think that this step is facilitated by exposing pupils to manipulative material, where they can recreate the situation presented, expose their initial beliefs and, with the help of the teachers, correct them through experience. According to Pratt (2000), there are many materials at our disposal that can serve as resources to support the construction of correct intuitions about probability. This recommendation follows the principle that knowledge is actively constructed by the individual and not passively received from the surroundings (Piaget and Inhelder, 1951), hence the importance of promoting the active teaching of children in the field of probability.

Our study is limited, due to the size of the sample and its intentional nature, so our purpose is to expand it in order to obtain more generalizable conclusions. However, despite these limitations, we believe that the information obtained can contribute to the training of educators involved in the teaching of this subject.

As pointed out by Alpízar et al. (2012) and Alpízar, Chavarría and Oviedo (2015), a portion of primary school teachers manifest insecurity when teaching probability either due to their weak didactic training in probability or complete lack of teaching experience on the topic. Therefore, we highlight the value of the information gathered by the present investigation, and more so in the context of curricular implementation currently experienced in Costa Rica.

The benefits spanning from the present investigation are multiple: it will assist in the development of educator training processes and teaching courses, it supports reflection on the cognitive demands of the tasks imposed upon young-age pupils, their ways of reasoning, possible biases and how such elements can guide educational planning.

ACKNOWLEDGMENTS

DECLARATION OF THE CONTRIBUTION OF THE AUTHORS

The percentage distribution regarding the contribution for the conceptualization, preparation and correction of the present article is disclosed as follows: L.H. 50%, C.B. 20%, M.G. 20%, RA 10%.

DECLARATION OF AVAILABILITY OF DATA

Data supporting the results of this study will be made available by author L.H. upon reasonable request.

REFERENCES

Alpízar, M., Barrantes, J., Bolaños, H., Céspedes, M., Delgado, E., Freer, D., Padilla, E. y Víquez, M. (2012). Aspectos relevantes sobre la formación docente en I y II ciclos en los temas probabilidad y estadística. *EDUCARE*, *16*(2), 113-129. https://doi.org/10.15359/ree.16-2.7

Alpízar, M., Chavarría, L. y Oviedo, K. (2015). Percepción de un grupo de docentes de I y II ciclo de educación general básica de escuelas públicas de Heredia sobre los temas de estadística y probabilidad. *Actualidades Investigativas en Educación, 15*(1), 1-23.

Batanero, C. (2006). Razonamiento probabilístico en la vida cotidiana: Un desafío educativo. En P. Flores & J. Lupiáñez (Eds.), *Investigación en el aula de matemáticas. Estadística y Azar*, 1-17. Granada: Sociedad de Educación Matemática Thales. CD ROM.

Behr, M. J., Harel, G., Post, T. R. y Lesh, R. (1992). Rational number, ratio, and proportion. En D. D. A. Grows (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 296-333). Macmillan.

Ben-Chaim, D., Keret, Y. e Ilany, B. S. (2012). *Ratio and proportion: Research and teaching in mathematics teachers' education*. Sense Publisher. https://doi.org/10.1007/978-94-6091-784-4_2

Bisquerra, R. (1989). *Métodos de investigación educativa*. P.P.U.

Borovcnik, M. (2011). Strengthening the role of probability within statistics curricula. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education* (pp. 71-83). Dordrecht. https://doi.org/10.1007/978-94-007-1131-0_11

Cañizares, M. J. (1997). *Influencia del razonamiento proporcional y combinatorio y de creencias subjetivas en las intuiciones probabilísticas primarias* [Tesis Doctoral]. Universidad de Granada, España.

Cañizares, M. J. y Batanero, C. (1997). Influencia del razonamiento proporcional y de las creencias subjetivas en la comparación de probabilidades. *UNO, 14*, 99-114.

Cerrón, W. (2019). La investigación cualitativa en educación. *Horizonte de la Ciencia, 9*(17), 1-8. https://doi.org/10.26490/uncp.horizonteciencia.2019.17.510

Falk, R., Falk, R. y Levin, I. (1980). A potential for learning probability in young children. *Educational Studies in Mathematics*, 11, 181-204. https://doi.org/10.1007/BF00304355

Fischbein, E. y Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? *Educational Studies in Mathematics*, *15*(1), 1-24. https://doi.org/10.1007/BF00380436

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. In G. A. Jones (Ed.), *Exploring probability in school* (pp. 39-63). Springer, Boston, MA. https://doi.org/10.1007/0-387-24530-8_3

Gil, J., León, J. y Morales, M. (2017). Los paradigmas de investigación educativa, desde una perspectiva crítica. *Conrado, 13*(58), 72-74.

Green, D. R. (1982). *Probability concepts in school pupils aged 11-16 years* [Tesis doctoral]. University of Loughborough, Reino Unido.

Jones, G., Langrall, C. y Mooney, E. (2007). Research in probability: Responding to classroom realities. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 909-955). Information Age Publishing y NCTM.

Karplus, R., Pulos, S. y Stage, E. (1983). Early adolescents proportional reasoning on "rate" problems. *Educational Studies in Mathematics,* 14, 219-233. https://doi.org/10.1007/BF00410539

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology.* SAGE.

Langrall, C. W. y Mooney, E. S. (2005). Characteristics of elementary school students' probabilistic thinking. En G. A. Jones, (Ed.), *Exploring probability in school. Challenges for teaching and learning* (pp. 95-119). Dordrech, Holanda. https://doi.org/10.1007/0-387-24530-8_5

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557-568. https://doi.org/10.1007/BF00540060

MECD (2014). *Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación Primaria*. Madrid: Autor.

Ministerio de Educación Pública (MEP). (2012). *Programas de Estudio de Matemáticas. I, II y III Ciclos de la Educación General Básica y Ciclo Diversificado*. Autor.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*. The Council.

Noelting, G. (1980a). The development of proportional reasoning and the ratio concept. Part I - Differentiation of stages. *Educational Studies in Mathematics*, *11*(2), 217-253. https://doi.org/10.1007/BF00304357

Noelting, G. (1980b). The development of proportional reasoning and the ratio concept. Part II - problem-structure at successive stages; problem-solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics*, *11*(3), 331-363. https://doi.org/10.1007/BF00697744

Pérez Echeverría, M. P., Carretero, M. y Pozo, J. I. (1986). Los adolescentes ante las matemáticas: Proporción y probabilidad. *Cuadernos de Pedagogía,* 133, 9-13.

Piaget, J. (1975). *Psicología de la inteligencia*. Buenos Aires: Psique.

Piaget, J. e Inhelder, B. (1951). *La genése de l'idée de hasard chez l'enfant*. Presses Universitaires de France.

Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education*, *31*(5), 602-625. https://doi.org/10.2307/749889

Truran, J. (1994). Examination of a relationship between children's estimation of probabilities and their understanding of proportion. En J. P. Ponte y J. F. Matos (Eds.), *Proceedings of the XVIII PME* (pp. 337-344). Universidad de Lisboa.

ANNEX: QUESTIONNAIRE

Item 1a: 3 black tokens and 1 white token are placed inside urn A. Urn B contains 2 black tokens and 1 white token. (Look at the drawing):


If you have to draw a black token to win a prize, without looking inside the urn, which urn would you choose to draw from? Indicate the correct answer:

(A) Urn A provides a greater chance of drawing a black token.
(B) Urn B provides a greater chance of drawing a black token.
(C) Both urns provide the same possibility.
(D) I don't know.

Item 1b: Why?

Item 2a: Two other urns have inside some black tokens and some white tokens (Look at the drawing):
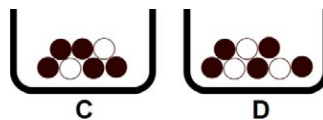
- Urn C: 5 black and 2 white.
- Urn D: 5 black and 3 white.

From which urn (C or D) is it more likely to draw a black token? Or, on the contrary, do both give the same possibility?

(A) Urn C.
(B) Urn D.
(C) The same possibility.
(D) I don't know.

Item 2b: Why?

Item 3a: Two other different urns also have black and white tokens (Look at the drawing):

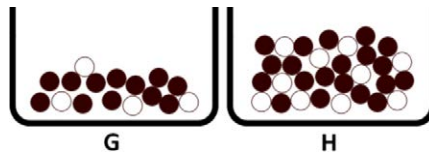- Urn E: 2 black and 2 white.
- Urn F: 4 black and 4 white.

Which urn provides the best chance of drawing a black token?
(A)   Urn E.
(B)   Urn F.
(C)   The same possibility.
(D)   I don't know.

Item 3b: Why?

Item 4a: Two other different urns have black and white tokens (Look at the drawing):

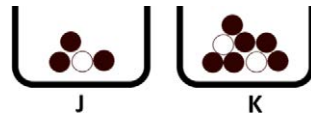- Urn G: 12 black and 4 white.
- Urn H: 20 black and 10 white.

Which urn provides the best chance to draw a black token?

(A)   The same possibility.
(B)   Urn G.
(C)   Urn H.
(D)   I don't know.

Item 4b: Why?

Item 5a: Two other urns different from the above have black and white tokens (Look at the drawing):

J   K

- Urn J: 3 black and 1 white.
- Urn K: 6 black and 2 white.

Which urn provides the greater chance of drawing a black token?

(A   The same possibility.
(B)   Urn J.
(C)   Urn K.
(D)   I don't know.

Item 5b: Why?