MODEL-ELICITING ACTIVITY WITH CIVIL ENGINEERING STUDENTS TO SOLVE A PROBLEM INVOLVING BINOMIAL DISTRIBUTION

MARTHA ELENA AGUIAR BARRERA
Cucei-Universidad de Guadalajara
martha.aguiar@academicos.udg.mx

HUMBERTO GUTIÉRREZ PULIDO
Cucei-Universidad de Guadalajara
humberto.gpulido@academicos.udg.mx

VERONICA VARGAS ALEJO
Cucei-Universidad de Guadalajara
veronica.vargas@academicosudg.mx

ABSTRACT

This research presents the results of the implementation of a model-eliciting activity called Brickyards, designed to promote the learning of the binomial distribution. The theoretical framework used was the Models and Modeling Perspective, and the participants were undergraduate students enrolled in a probability and statistics course of the Bachelor Civil Engineering Program at the University of Guadalajara, Mexico. The activity was refined during three semesters, and here we report the models generated by the students in the fourth implementation. In the first stage of the activity of this implementation, students proposed wrong solutions, which were based on ideas of proportionality and linear thinking. The activity was designed to inhibit these types of solutions and to encourage students to realize when they are dealing with a random phenomenon, and that they need a probability distribution to solve the activity. The students used RStudio software to calculate probabilities.

Keywords: *Statistics education research; Linear thinking; Proportional model; Random model; Probabilistic model*

1. INTRODUCTION

Statistics is gaining importance and social recognition as a fundamental tool for science, economics, and politics. Also, quantitative information is omnipresent in media and in the everyday lives of citizens worldwide (Ben-Zvi & Makar, 2016). Over the past few decades, statistics have become an increasingly essential tool for companies to control and improve the quality of their products and services. This has also led to the promotion of evidence-based decision-making as one of the key principles of quality management, allowing companies to make informed decisions based on data (International Organization for Standardization, 2015). In these tasks, probability distributions become the mathematical tool that facilitates the identification of patterns of variation and the characterization of the uncertainty (randomness) inherent in the data.

Accounting for variability with the use of distributions is the key in the analysis of data (Franklin et al., 2005). The understanding of randomness, therefore, is considered a statistical competence that a citizen must have today, to understand the behavior of random samples and be able to interpret a margin of sampling error (Franklin et al., 2005). This implies that attention should be paid to the development of statistical competence to understand sample variation in the teaching of probability and statistics. Additionally, this makes this subject increasingly relevant in the school curriculum (Franklin et al., 2005). Its teaching, however, faces several obstacles, one of them is that problems that are random are

Statistics Education Research Journal, 22(3), Article 3. https://doi.org/10.52041/serj.v22i3.431

not identified as such by the students, who in most cases try to analyze and solve them with proportional or linear reasoning (Dooren et al., 2003).

The illusion of linearity of phenomena leads to a series of erroneous concepts and ideas in mathematics and probability (Dooren et al., 2003), such as the misconception that random samples accurately represent the characteristics of the population: “People have erroneous intuitions about the laws of chance. In particular, they regard a sample randomly drawn from a population as highly representative, that is, similar to the population in all essential characteristic” (Tversky & Kahneman 1971, p. 105). This misconception of random samples leads to interpreting the sample statistics, such as mean and proportion, as if they were equal to their population counterparts (Gutiérrez, 2020). This limits the student’s ability to visualize the whole variability of the random phenomenon under study, and they only focus on the values with the highest probability.

Part of the problem is that teaching probability distributions often fails to emphasize the context in which problems arise, focusing instead on mathematical procedures and calculations. As a result, students are unable to identify which problems in their environment or profession can be modeled using probability distributions. Furthermore, it is not emphasized that probability distributions are the models of population measurements, and that probability seeks to quantify uncertainty, which is inherent to almost all phenomena (Franklin et al., 2005).

To address this teaching/learning problem, different alternatives have been proposed, for example, the use of educational software to analyze and simulate different random phenomena (Pfannkuch et al., 2018) and the use of problems “that touch the lives of students and are linked to their specific problems in the real world” (Pfannkuch et al., 2018, p. 1115). There are several reports that analyze the teaching of probability distributions through activities relevant to students, which make use of simulators to facilitate the visualization of random phenomena related to distributions such as normal, Poisson, or binomial (Batanero et al., 2001; Bill et al., 2009; Budgett & Pfannkuch, 2018). It is necessary, however, to generate more proposals of this type, for example, using problems close to the professional interest of students that allow them “to understand the behavior of a real-world system and, consequently, to develop a deeper contextual knowledge and an understanding of the real-world situation” (Pfannkuch et al., 2018, p. 1115).

This report presents the solution paths of students of an undergraduate level course when they tried to solve a problem in the civil engineering context. The problem involves the use of the binomial distribution and was posed based on criteria taken from the Models and Modeling Perspective ([MMP]; Lesh, 2010; Lesh & Doerr, 2003; Lesh et al., 2000). The solution strategy is organized in three stages, starting with individual work, then working in teams, and concluding with a plenary session. Some teams relied on *RStudio* software (https://www.rstudio.com) for some calculations and decisions.

The question to be answered in this study is:

> What are the models a group of civil engineering students generate when they try to solve a problem related to real life, where binomial distribution is required to model the random phenomenon involved?

For this purpose, a Model Eliciting Activity (MEA) was designed and applied from the MMP, which is related to the field of study of civil engineering and is close to the students’ reality as advised by Lesh (2010). The activity also allowed for the consideration of the use of technology to facilitate the solution. The design of the MEA was refined during three semesters of a probability and statistics course of the Bachelor Civil Engineering Program at the University of Guadalajara, Mexico. The fourth implementation is reported here, and it was carried out in the first semester of 2021. This research report contributes to the study of the learning process of random phenomena.

2. BACKGROUND

As Florensa et al. (2020) pointed out, in the last two decades there has been a substantial increase in the number of studies focused on modeling perspectives, employing different approaches. This trend is also manifest in the field of modeling problems associated with probability and statistics. It includes research on the teaching and learning of random phenomena. This research emphasizes the execution of activities pertaining to real-world problems and the utilization of software that enables the simulation of these phenomena.

Different authors have made developments regarding the teaching and learning of statistics in the framework of the CATALST (Change Agents for Teaching & Learning Statistics) curriculum (e.g., Garfield et al. 2012; Noll et al., 2018). It was one of the first simulation-based introductory statistics curricula to be developed for radical change in introductory statistics courses (Justice et al., 2020). The CATALST curriculum "uses the ideas of chance and models, along with simulation and randomization-based methods, to enable students to make and understand statistical inferences" (Garfield et al., 2012, p. 883). This curriculum combines several elements: simulation and randomization through *TinkerPlots*™ software and the transition to thinking about real problems. Garfield et al. (2012) reported that the CATALST program has been successful and indicate that over a four-year period it has been replicated 15 times at 10 different institutions.

A continuation of the CATALST project was reported by Noll et al. (2018), who, through narrative, studied the development of statistical models created by students as they tried to solve a real-life problem with the support of *TinkerPlots*™ software to simulate data. Their results contribute to the conception of the curriculum in probability and statistics regarding the construction of statistical models through the analysis of the students' narrative.

Budgett and Pfannkuch (2018) indicated that simulations, properly designed, allow students to experience random phenomena: "For example, by simulating a Poisson process, it would be possible to make connections between the rate at which events occur, and the time that elapses between events" (p. 1283). These authors accompanied their proposal with the design of a task that follows six principles. One of the principles is related to context and Budgett and Pfannkuch used "data relating to FIFA World Cup soccer tournaments … to promote student engagement" (p. 1284). They pointed out their proposal could support the understanding of randomness and the probabilistic process of the Poisson distribution.

Regarding *R,* a programming language for statistical computing and graphics, more studies are needed to analyze its impact on the learning of probability distributions. On this topic, we found a study by Stemock and Kerns (2019), which compared *R* with *SPSS*. The study did not identify significant differences in terms of students' opinion on the ease of use, impact on course quality, and recommendation for future use. Which is a positive aspect, since *R* has the advantage of being cost-free.

No studies were found at the undergraduate level on how students internalize the probability model for binomial distribution, so this is one of the main contributions of the present study. In addition, there is little literature on probability and statistics in the field of civil engineering at the undergraduate level. In relation to this, three classic textbooks of this discipline for engineers were reviewed (Devore, 2011; Hines et al., 2008; Walpole et al., 2012), and it was found that these books contain very few problems related to civil engineering. Furthermore, these textbooks emphasize the application of algorithms, and place much less emphasis on understanding and modeling the problems. Therefore, another contribution of this study is the design of a MEA, related to a problem in the context of civil engineering, to promote the learning of binomial distribution.

3. MODELS AND MODELING PERSPECTIVE

The theory that underpins this research encourages students to model a problem instead of focusing on the answers produced to solve it. The model is understood as the creation of:

> … conceptual systems (consisting of elements, relations, operations, and rules governing interactions) that are expressed using external notation systems, and that are used to construct, describe, or explain the behaviors of other system(s)—perhaps so that the other system can be manipulated or predicted intelligently. (Lesh & Doerr, 2003, p. 10)

In order to develop these models, the starting point is choosing problematized situations that are close to the students' reality. These situations can be real or appear to be real to provoke the need to create a mathematical model (Lesh et al., 2000; Noll & Kirin, 2017).

Each individual generates their own models (conceptual systems) that enable them to explain and give meaning to specific situations (Lesh & Doerr, 2003). These models are created in constant confrontation with their previous knowledge, with the ideas of their peers and with the participation of teachers or coaches. This exchange of ideas leads to "building shared meanings and solutions to problems" (Noll et al., 2018, p. 1269).

MEAs are a design tool that MMP uses to enable the generation of models. These activities are "simulations of 'real life' in which students' conceptual understandings can be directly documented and assessed" (Sevinc & Lesh, 2018, p. 302). The MEA enables the generation of models by recreating real problems that can be solved in a short time (90 min.). In addition, MEA creates conditions so these models can be "documented so as to provide a trail of evidence of learners' changing ways of thinking" (Brady, 2018, p. 46).

Lesh et al. (2000) pointed out that students have both school abilities and real-life abilities, and the two often function almost completely independently. Since these skills are not linked, the brain makes the necessary adaptations to work with the former in school situations and the latter in family and friend situations. Achieving the linking of both skills will allow school knowledge to remain and become part of the students' daily life. MEAs are the bridge that favors the linking of these two abilities (school and real-life). Through engaging in the activity, students construct mathematical models that leave a trail of evidence for educators to better understand students thinking (Garfield et al., 2012).

4. METHODOLOGY

The MEA design was refined during three semesters in a probability and statistics course of the Bachelor Civil Engineering Program at the University of Guadalajara, Mexico. The process followed in the fourth implementation is detailed here, which was carried out in three stages, starting with individual work, then working in teams, and concluding with a plenary session.

The models generated by the students when they addressed the problem posed in the MEA called Brickyards, were analyzed qualitatively. The activity involves the binomial distribution and was designed under the six principles proposed by Lesh and collaborators (2000):

1) Construction of models—the way in which the activity is presented should provoke the need to look for a mathematical solution.
2) Reality—the problem raised is real or could be real.
3) Self-evaluation—allows the student to evaluate the proposed solutions for themselves.
4) Documentation—explicitly reveals how students are thinking about possible solution paths.
5) Shareable and reusable construction—the proposed model is easily modifiable and reusable in other situations.
6) Effective prototype—students will remember the model when they encounter other structurally similar situations.

The design of the MEA Brickyards starts from raising an important real problem, which refers to waste due to critical defects in the production of fired clay bricks for the construction of walls. This artisanal production is carried out by small companies, whose average percentage of waste is 7% (Instituto Nacional de Ecología y Cambio Climático, 2016). Based on this reality, an imaginary contest was conceived to encourage brickyards to reduce their waste (Figure 1). The requirement for a brickyard to participate in the contest is that it achieves less than 7% waste (which is the industry average in the state of Jalisco, Mexico). The winning factory will be the one that demonstrates the greatest waste reduction with respect to 7%.

Deciding which companies meet the requirements and the winner of the imagined contest is not a simple problem. To make these decisions, it is necessary to rely on statistical sampling of the production of the participating companies. A scenario with different sample sizes and levels of defects is proposed in such a way that the decision is not obvious and encourages students to analyze the activity from a probabilistic basis (Table 1). The sample sizes were decided by assuming different company sizes and applying criteria from the Tables for Inspection by Attributes MIL-STD-105E (Gutiérrez, 2020).

The Environmentalist

Digital newspaper

WEDNESDAY, NOVEMBER 11, 2021

Contest to Reduce the Waste in Artisanal Brick Production in Jalisco has been challenged by several brickyards

The Ministry of the Environment and Natural Resources (SEMARNAT) and the Mexican Chamber of the Construction Industry (CMIC) have implemented a program in Jalisco called Good Practices and Environmental Care in the Artisanal Production of Fired Clay Bricks. There are "around 17,000 brickyards distributed throughout the country under conditions of informality in small production units, with rudimentary technology and offering products that differ regionally in terms of dimensions and appearance characteristics" (INECC, 2016)*,

An important problem for brickyards is the waste due to critical defects (molten, broken or raw bricks). As a way to improve artisanal brick production, SEMARNAT and CMIC launched a contest to reduce waste brickyards the state of Jalisco (Mexico).

However, the contest was overshadowed by challenges and negative comments from several of the contestants.

The contest rules stipulate that to be a contestant it is necessary to achieve a brick waste below 7%, which was the average for this industry in Jalisco in 2016.

The complainants claim that there is little clarity in the criteria for **disqualifying a brickyard**, as well as in the election of the **winning brickyards**.

The organizers point out that the disqualification of the contestants and selection of the winners were properly made.

To obtain evidence of the level of critical defects, the organizers took into account the size of the brickyards and the waste of bricks.

The following table shows the data considered to choose the winning brickyards, as well as those that were disqualified for not achieving a waste below the state average (7%).

Brickyard	Sample size	Brick waste	Evaluation
L1	200	10	Disqualified
L2	215	7	Second place
L3	315	13	Third place
L4	323	16	Disqualified
L5	500	22	First place
L6	898	49	Disqualified

An anonymous source has informed this newspaper that the organizers are thinking of cancelling the contest, which would be unfortunate because it is a good incentive to reduce the waste of bricks and thus reduce costs and pollute less.

Help us to demonstrate that the evaluators decided correctly in the choice of the winning and disqualified brickyards.

Let's support the continuation of the contest!

Send us a proposal describing why the organizers made the right decisions regarding:

(a) Disqualifying three brickyards for not guaranteeing that they achieved less than 7% waste.

(b) Allocating the first three places.

Figure 1. MEA Brickyards online format [English translation].

Table 1. Sample sizes and defective bricks

Brickyard	Sample size	Defective bricks	Evaluation	Sample waste (%)
B1	200	10	Disqualified	5.00
B2	215	7	Second place	3.30
B3	315	13	Third place	4.10
B4	323	16	Disqualified	4.90
B5	500	22	First Place	4.40
B6	898	49	Disqualified	5.50

Table 1 shows the data considered to choose the winning brickyards, as well as those that were disqualified for not achieving a waste below the state average (7%). The six sample sizes and the corresponding number of defective bricks are different, so the decision is not evident. On the assumption that students have a deep-rooted linear thinking, these data were chosen in such a way that for the decision analysis it was not sufficient to calculate the percentage of defective bricks in each manufacturer's sample. Indeed, applying this percentage criterion, the decisions made by the contest organizers on the disqualified and winning brickmakers appear to be incorrect. This is detailed in Table 1, which shows that brickyards B1, B4 and B6 were disqualified, even though all three have a sample waste of less than 7%. Also, the winning brickyard has a higher percentage waste (4.4%) than the second and third place (3.3% and 4.1% respectively).

The activity emphasizes that the decisions shown in Table 1 were appropriate and that the goal is to show that the evaluators decided correctly. The idea is to encourage students to go beyond the use of percentages (proportional model), and to explore alternatives based on probability. The expectation is for them to analyze the effect of using different sample sizes on both variability and uncertainty. For this task the students could use *RStudio* software, as they had been trained in its use before participating in this research.

The implementation of the activity was carried out in three stages. In each of them, the instruction was the same: showing that the contest organizers made the right decisions in both disqualifying and rewarding the brickyards (see sentences *a* and *b* at the end of Figure 1). The stages were as follows:

1. Individual work: Writing the first individual solution ideas in an online blogspot.
2. Teamwork: Sharing the individual solutions and generating a team solution proposal by consensus.
3. Plenary session: Presenting each teams' proposal.

The analysis of the models constructed by the students was carried out based on the work developed by the students during the three previous stages. Release 8 of the *Atlas.ti* (https://atlasti.com/es) software was used for the analysis of the writings and the videos of the teamwork and the plenary activity. The videos were recorded with *Google Meet*.

5. ANALYSIS OF RESULTS

The most relevant aspects of the data obtained in the three stages in which the MEA Brickyards were resolved are presented in this section.

5.1. INDIVIDUAL WORK STAGE

In this stage, the students made the first approaches to try to justify or question the decisions of the contest evaluators. Thus, the first ideas related to percentages, the rule of three, and sample size emerged. This is shown below.

Percentage and rule of three approaches. As mentioned in the methodology, the MEA Brickyards was designed to avoid the emergence of any proportional model. Even so, it appeared in this first stage. Figures 2 and 3 show quotations from the comments made by the students in the blogspot, which were assigned the code "percentage" and "sample" in *Atlas.ti.* Figure 2 shows that out of a total of 17 student comments, 13 used the sample percentages as if they were population parameters (the production of the brickyards). Two examples of this are the comments 29:1 and 29:7:

29:1 disqualified brickyards [that] did not exceed 7% of waste, so they were unfairly disqualified.

29:7 We can check by applying the rule of three, and they do not exceed 7% of waste, and checking the brickyards chosen as winners are poorly organized in their places.

In this regard, Konold (1995) pointed out that when a model such as the percentage is internalized, the transition to the random model is not easy.

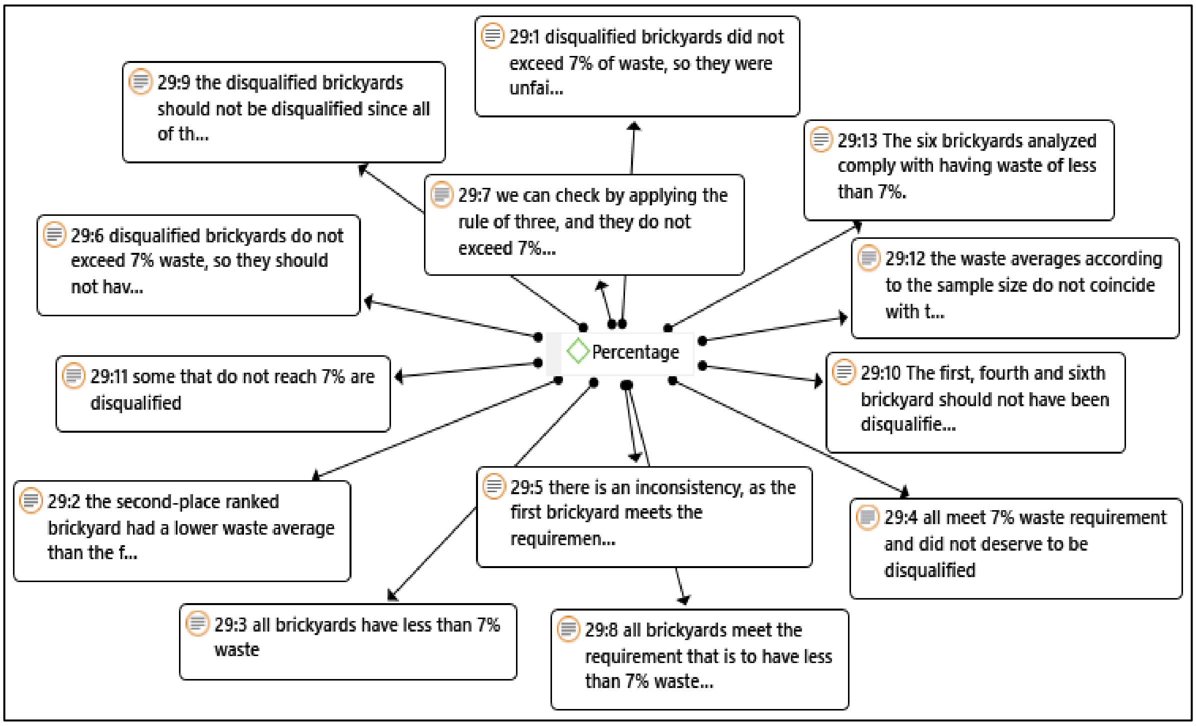

Figure 2. Comments made by students on the blogspot related to the percentage in the first stage of the MEA Brickyards.

The idea of proportionality also emerged in the students' arguments about the sample (Figure 3), when they proposed samples of the same size to achieve a fair contest:

32:3 It would be necessary to have an equal sample size for all and thus know the brickyard with the lowest waste.

32:4 We can see that the samples are very varied and in the case of the wastes they do not vary so much in the case of some, so I consider that the way to choose the first place was somewhat unfair.

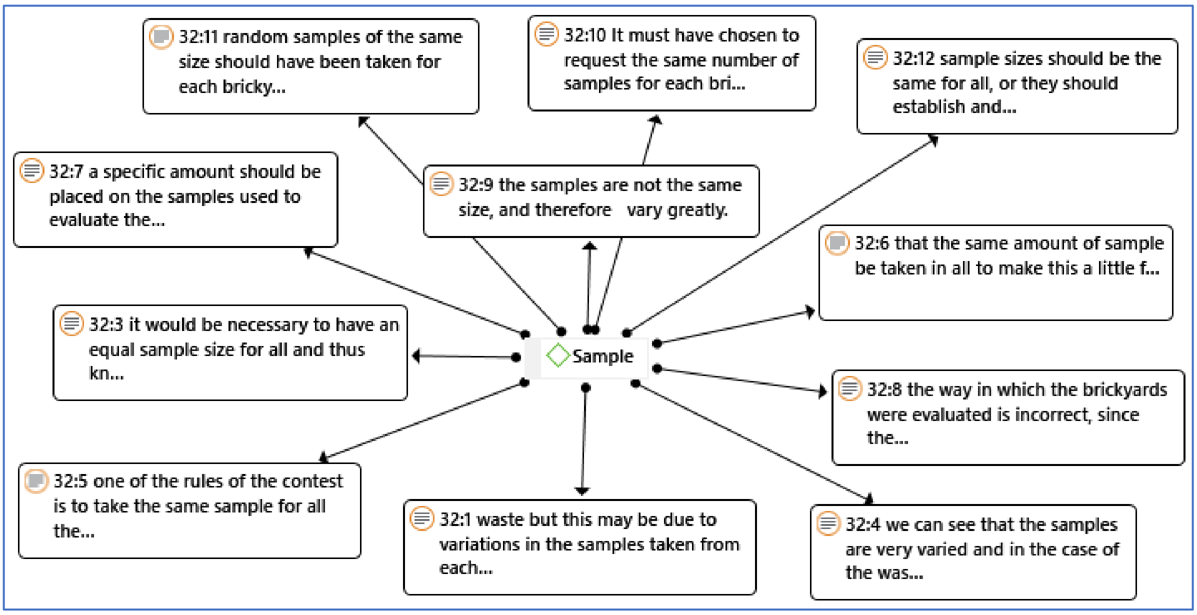

Figure 3. Comments made by students on the blogspot related to the sample size in the first stage of the MEA Brickyards.

5.2. TEAMWORK STAGE

For the team activities in the probability and statistics course, the class was organized into ten teams. Teams seven and nine did not participate in any activity of this unit, so in this part of the analysis the work of eight teams (T1, T2, ..., T6, T8 and T10) was reviewed.

The analysis of this stage was based on the video recordings of the virtual meetings of each team and the first solution proposal of each team, which was presented in free format (Word or PowerPoint). The dialogues were analyzed in the *Atlas.ti* software and coded as follows: the letters TW to indicate teamwork, followed by a number indicating the team number and a letter to designate the participant (which is not related to the student's name or surname to preserve their anonymity). For example, TW1J indicates a comment made in teamwork, in Team 1, by Student J. Regarding the analysis of the documents with the first solution proposal, the following coding was used: the letters FTW for the first printed version of the teamwork, followed by the team number.

Ideas on percentage and proportionality. At the beginning of the teamwork, the teacher indicated that the simple value of the percentage of defective bricks in the samples (Table 1) did not solve the activity. Even so, the attempt to solve the activity based on percentages appeared in the dialogue of several teams. T1 recalculated the percentage of waste for each sample and corroborated that all the companies were below 7%, and they also realized that the third place in the contest had a lower percentage than the second place. Faced with this apparent contradiction, they questioned whether the activity had a trick. One student speculated, "There is something hidden in the reading, like a message to decipher or something like that" (TW1E).

The preliminary analysis of T3 was related to the percentages. To explain the decisions about the disqualified sites, the team argued that there could only be three places. Therefore, the rest were disqualified. Their way of resolving the contradiction (third place lower percentage of waste than first place) was to ignore it. Another team that also began its analysis with the percentages was T10. When they identified the contradictions, they were blocked for a long time without being able to get anywhere.

In turn, the first idea T4 explored was a relationship between the sample size and the number of defective bricks. "If you look at it, for every 20 bricks, one goes wrong. Almost all samples have like that score" (TW4J), to which their partner responded, "I think we are already following the thread of fate" (TW4O), giving continuity to the argument. The team reviewed the table again and returned to the initial idea of the percentage. "Check the table for the percentage, those that were disqualified are the ones with the highest percentage, the percentage closest to 7%" (TW4J). The idea concluded there, without corroboration. It appears that the initial ideas of these four teams (T1, T3, T4 and T10) were again about the percentages of the samples, seen as the identical reflection of the corresponding characteristic of the population. That is, without understanding the random characteristic of the samples (Tversky & Kahneman, 1971).

It is noteworthy that when the teams found that they could not solve the disqualification and the award with the sample percentage, they looked for other solution paths. This is in agreement with Lesh et al. (2000), who noted that when students have the possibility of evaluating through the indications of the activity that they are not on the right track, they can look for other possible solutions.

Trial-and-error approach with probability distributions. All the teams that started by trying to solve the problem with percentages, then changed and tried with probability distributions. For example, T1 wondered if the problem would be related to probability and mentioned that it could be with the binomial distribution, but the team do not go deeper into this option. Instead, they looked up the definition of random experiment, which had been seen in the previous unit of the course, but they could not relate it to the MEA Brickyards.

Another approach followed by several teams was to use *RStudio* commands related to probability distributions. They incorporated the data from the activity, but more with a trial-and-error method than with a clear strategy. Thus, T2 and T6 tried to solve the activity with the hypergeometric distribution but observed that they did not have N (lot size or population from which the sample was taken); subsequently, they looked for another option. In one case, T2 confused the interpretation of the p parameter of the binomial distribution by associating it with the probability of obtaining good bricks, when it should have been related to the probability of obtaining a defective brick. This confusion is a language problem, frequent with this distribution, since in the usual definition of this distribution, it is indicated that the parameter p represents the probability of the event of interest, called "success". In the application, the event of interest was defective bricks, and not good bricks, as it was interpreted by T2. Both teams (T2 and T6) left the hypergeometric distribution and chose the binomial distribution, as did T3, T4, and T10. Although they all took the path of using the binomial distribution, the way of approaching it was different as detailed below.

T2 relied on *RStudio* to apply the binomial distribution. In this process they identified the sample size, but had problems identifying p, and instead used the waste percentage of the samples. For example, for brickyard B1, they used $n = 200$ and $p = 0.05$; but to plot the distribution they used the whole domain of the probability function (x = 0,1, 2, ..., 200), so that the most probable values of x are left in a small portion of the left side of the plot, giving the impression that the distribution has a very long right tail and they wondered why the plot had that shape (Figure 4).

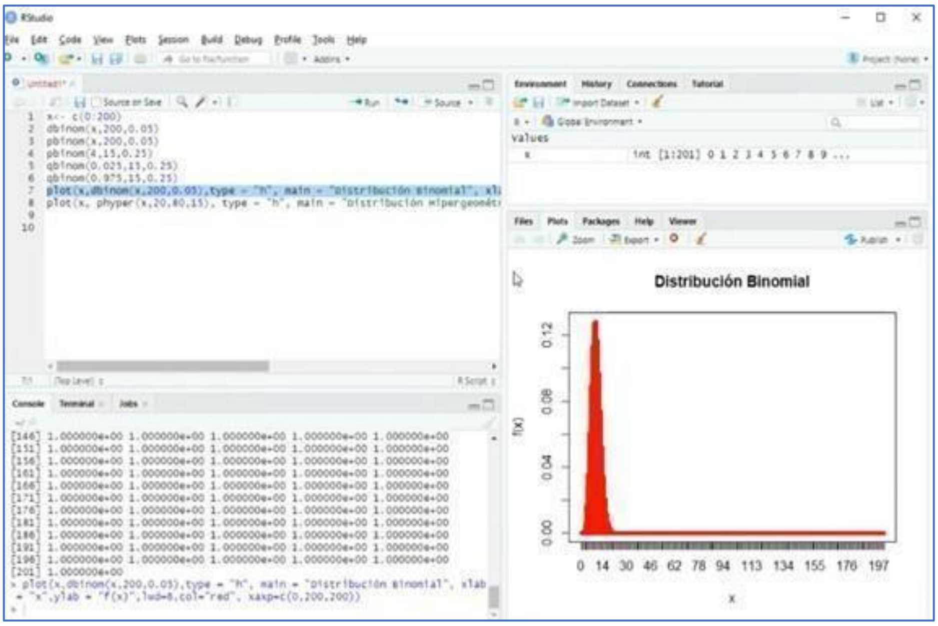

Figure 4. T2 using RStudio.

One of the team members worked on his own and changed $p = 0.05$ to $p = 0.07$ (7% waste in Jalisco) and applied the inverse of the binomial distribution in *RStudio*, as *qbinom*(α, n, p), first with $\alpha = 0.05$ and then $\alpha = 0.10$. With this last value he obtained the values where x was expected to be (between 8 and 20) and he made the following interpretation:

TW2S: What I can do is change here 0.5 and 0.95, let's see, no, there is the eight, it is the minimum wow, I think that's why.
TW2C: What do you mean, 8 is the minimum? What?
TW2S: Well, it says that it has to be 7%, right? Ah, well this one got 8%.
TW2C: Aah, and that's why it's out.

They interpreted the inverse of the binomial distribution result, *qbinom*(0.05, 200, 0.07) = 8, as a percentage and compared it with 7%, which was higher, so they considered that Brickyard 1 was disqualified. They were only able to perform the analysis for the first brickyard. Perhaps, if they had had the results of all the brickyards, they could have changed this interpretation.

Teams T3 and T4 adopted a similar approach using the cumulative binomial distribution, where x = waste and $p = 0.07$ (Figure 5), but they could not relate the results to the disqualified brickyards (higher cumulative probability), nor to the awarded ones. This is an alternative solution to the activity, which could have been achieved by these teams.

Brickyard 1 > *x* < - *c*(0:200) > *pbinom*(10,200,0.07) [1] 0.1661267	*Brickyard 3* > *x* <- *c*(0:315) > *pbinom*(13,315,0.07) [1] 0.02329336	*Brickyard 5* > *x* < - *c*(0:500) > *pbinom*(22,500,0.07) [1] 0.01051955
Brickyard 2 > *x*< - *c*(0:215) > *pbinom*(7,215,0.07) [1] 0.01472532	*Brickyard 4* > *x* < - *c*(0:223) > *pbinom*(16,223,0.07) [1] 0.6060626	*Brickyard 6* > *x* <- *c*(0:898) > *pbinom*(49,898,0.07) [1] 0.03666152

Figure 5. Results obtained in RStudio with the cumulative binomial distribution with x = waste, FTW3.

In contrast T6 used the binomial distribution and the sample proportion of waste of each brickyard as *p*. When they were asked why they used this value of *p*, they changed to 0.07 but confused the result of the binomial distribution with the average waste. As they could not reach a conclusion, they looked for another way, which was to calculate *P(x = 1)* in the six brickyards, as a way to "compensate" for different sample sizes. As a result, T6 obtained very small values for the brickyards with large sample sizes. They did not, however, interpret the values (Figure 6).

```
> dbinom(x,200,0.07)
 [1] 4.972671e-07 7.485741e-06 5.606257e-05 2.785044e-04 1.032413e-03
 [6] 3.046172e-03 7.451658e-03 1.554432e-02 2.822631e-02 4.532397e-02
[11] 6.515930e-02
> x <-(1)
> dbinom(x,200,0.07)
[1] 7.485741e-06
> x <-(1)
> dbinom(x,215,0.07)
[1] 2.70949e-06
```

```
> x <-(1)
> dbinom(x,315,0.07)
[1] 2.799332e-09
> x <-(1)
> dbinom(x,323,0.07)
[1] 1.606238e-09
> x <-(1)
> dbinom(x,500,0.07)
[1] 6.56235e-15
> x <-(1)
> dbinom(x,898,0.07)
[1] 3.369607e-27
```

Figure 6. RStudio results presented in FTW6

Another team that used the inverse of the binomial distribution was the T10, with *α = 0.025* and *p = 0.07* and the sample size of each brickyard. They subtracted the sample waste (bricks with defects) from the value of *qbinom*(*α, n, p*), obtaining negative numbers for the disqualified brickyards. They complemented this analysis with a bar graph and explained, "Now we can see how the values have changed based on the binomial distribution and it is clearly seen why Brickyards B2, B3 and B5 have obtained the first three places" (FWT10, p. 11). They did not specify, however, that those with larger bars were the ones disqualified (Figure 7).

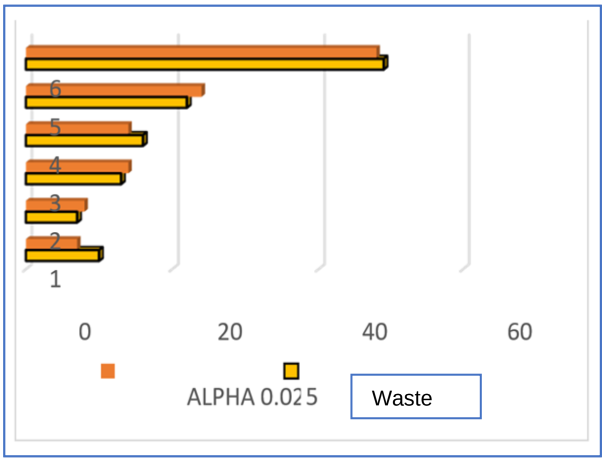

Figure 7. Bar chart to compare qbinom(α, n, p) and sample waste for each brickyard, FTW10.

In summary, the teams that explored a probabilistic analysis in *RStudio* (T2, T3, T4, T6 & T10) were unable to arrive at a clear formulation of an explanatory model for the disqualification and rewarding of brickyards. They used the *RStudio* without a clear reason about the parameters to be used in the three functions of the binomial distribution: the probability function (*dbinom*), the cumulative (*pbinom*), and the inverse (*qbinom*). They were also unable to forcefully link the results with the MEA Brickyards. This difficulty of relating a problem with a distribution that can model it is something that happens frequently and not only happens to students, but also to experts. Garfield et al. (2012) indicated that when modeling random phenomena, no matter the level of experience of the modeler, everyone faces the conflict of identifying the appropriate model for each reality.

Proportionality model or rule of three. To solve the problem of working with different sample sizes, some teams matched the sample sizes and adjusted the number of defective bricks proportionally, and thus could decide based on this adjusted value. This process overlooked the fact that the teams were working with samples, where such proportionality does not necessarily work.

T5 expanded the sample size, *n*, to 1000 for all brickyards, and the expansion factor for each *n* was applied to the number of defective bricks in the sample for each case (Figure 8). We have called this approach the Sample Equalization Model. The justification given by this team was as follows:

> If we matched them all to a sample quantity such as one thousand, for example, we would see that those that were disqualified were with good reason since they have a brick waste of 50 or more, relative to those that were selected as winners. (TW5I).

They concluded:

> Therefore, we can say that, although all the brickyards present a percentage lower than the 7% established to participate and not be disqualified, the decision of the organizers was taken in the right way, because, although the brickyards that were selected with the first, second and third place produce a higher amount than the disqualified ones, if we look proportionally they have a much lower average waste than the others. (FTW5, p. 7).

Brickyard	Sample size	Defective bricks	Sample waste in percentage	
B1	1000	50	5.00	%
B2	1000	33	3.26	%
B3	1000	41	4.13	%
B4	1000	51	5.08	%
B5	1000	44	4.40	%
B6	1000	55	5.46	%

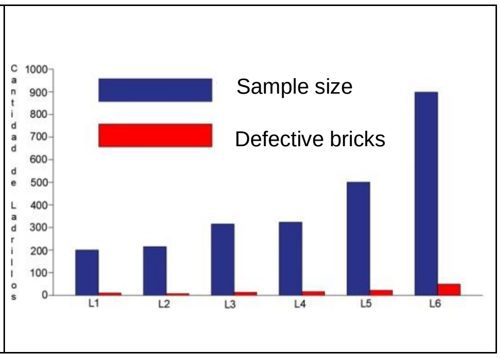

Figure 8. Model with sample sizes of 1000 for each brickyard, FTW5.

As can be seen, this team built a model that combines percent defective with a sample of 1,000. They concluded that those with 50 or more defective bricks were disqualified. Underlying this approach is the idea of applying the rule of three, where the sample size of each brick (which would correspond to 100%) was changed to 1000. The team calculated the decrease obtained by each one according to its reduction percentage. This only partially worked for the disqualified, but not for the awardees. It must be remembered that this is because the binomial probability, in this case, does not transfer linearly. This erroneous way of proceeding by this team is a reminder that problems involving samples must be approached using probabilistic tools, because in general, the calculation of probabilities is not achieved with linear thinking.

T8 proceeded in a relatively similar way to justify disqualification. They obtained the average of the sample sizes (408.5) and the average of the number of defective bricks (19.5), and from these the average percentage of waste, 4.77%. Then, T8 compared this average with the percent defective of the sample of each brickyard (see Table 1). They saw that all the disqualified brickyards have a percentage higher than 4.77%. The T8 members concluded that this was the reason for the disqualification.

It is interesting to note that because concepts such as percentage and linear thinking were internalized, the students were able to combine the ideas creatively with other concepts to reach a solution, as happened in the last two cases. These teams built two models that they considered satisfactory to identify the disqualified brickyards. In their analysis they considered the six brickyards, without visualizing a method that could be used to evaluate any brickmaker, as happens in a contest. In this regard, the MMP says that in the initial modeling processes, students focus on one part of the problem without seeing the whole or as Lesh (2010) points out, "When the functioning of a system is relatively 'uncoordinated' (i.e., non-systemic) students tend to 'see' the 'forest' and loose cognizance of the 'trees'—or vice versa. And, when they focus on one type of detail, they often lose cognizance of other details" (p. 18).

Presentation of progress (teamwork stage). Since the MMA Brickyards was not completed in the time assigned to the first teamwork session, the students were asked to continue with the activity at home, and if they needed advice from the teacher-researcher, they should request it. As a result of this advice, T2 (which in their previous work had related the wastage to the results of the inverse of the binomial distribution) was able to link this function with the disqualification.

Before continuing with the teamwork (second day), T2 presented its progress in class, as a way of supporting all the teams that could not advance in their analysis. In the presentation they related the results obtained from the inverse of the binomial distribution with the problem and identified that savings exist when the waste is less than the result of this function with an $\alpha = 0.025$ (Figure 9).

Brickyard 2

Brickyard 2 has a sample of 215 bricks, of which 7 were defective bricks.
The organizers of the contest accepted and gave the second place to brickyard 2, this is because it meets the requirement of less than 7% waste, then through the command in which we will use alpha of 0.025 and 0.975 to know our result

```
> qbinom(0.025,215,0.07)
[1] 8
> qbinom(0.975,215,0.07)
[1] 23
```

This data is what *RStudio* provides us, we simply gave the order to give us the range in which we have a waste of 7%, the evaluation is fair because what the contest is looking for is a waste less than or equal to 8 bricks, in this case were 7 bricks one less than the 8 of the range of 7% of waste so it should not be disqualified.

Figure 9. Inverse of the binomial distribution analysis and its relationship to waste, FTW2.

In T2's presentation, several doubts arose from the rest of the students in the class. This illustrates the difficulty of understanding the model of the inverse of the binomial distribution for the solution of the MEA. The starting point is to identify the possible values of the number of defective bricks, x, which have the highest probability of occurrence, under the assumption that the population waste is 7%. For example, if students want to find the interval $[L, U]$, in which the possible values of x can fall with probability 0.95, this is:

$$Pr(L \leq x \leq U \mid p = 0.07) = 0.95$$

then a good way to find the values of L and U is to match them with quantiles 0.025 and 0.975 of the corresponding binomial distribution. In *RStudio* these quantiles are obtained with the inverse of the binomial distribution by:

$$L = q_{0.025} = qbinom(0.025, n, p) \text{ and } U = q_{0.975} = qbinom(0.975, n, p)$$

So, if x falls outside that interval, it is evidence against $p = 0.07$. In particular, if $x < L$, it is evidence in favor of $p < 0.07$, which is what each brickyard should guarantee to participate in the contest.

Another alternative solution is to calculate the probability that x is less than or equal to the observed value in the sample of each brickyard, x_0; which with *RStudio* is obtained with *pbinom*(x_0, n, p). If this probability is less than 0.025, then the observed value x_0 is outside and below the aforementioned interval and, consequently, it is evidence that $p < 0.07$.

One aspect that emerged was a question related to the graphs that several teams made in *RStudio* and where they obtained a very large right tail. As seen in Figure 4 (distribution panel), where students plot the binomial distribution for its entire domain ($x = 0, 1, ..., n$). That is, they did not relate that by identifying the values of x with more probability, it also helps to define the values of x for which it is convenient to plot the probability function. This confusion was identified using the data analysis software. In this case, *RStudio* helped to understand the difficulties of students in comprehending and applying probability distributions.

At the end of this stage of teamwork, the assignment of the first three places in the activity was not resolved, so the teamwork continued, and its results will be analyzed in the next stage.

5.3. PLENARY SESSION STAGE

In the plenary session, each team delivered a document with its analysis of the previous stage, both for the disqualified brickyards and for the awarded ones. In addition, each team made an oral presentation, which was videotaped. Below we present the highlights of these documents and presentations. To identify to which team a contribution belongs, we use the following codes:

- T#P for an idea expressed during the team # presentation.
- T#D for a contribution of team # in its corresponding document.

For example, T2P is a comment made in the presentation of team 2.

At this stage, an important evolution was observed in the way students analyzed the activity. All the teams, except T1, were able to model the behavior of x with the binomial distribution, that is, they modeled the expected number of defective bricks with a higher probability of occurrence. For which they used:

- the inverse of the binomial distribution to identify disqualified brickyards; and
- the cumulative binomial distribution to identify the winning brickyards.

T5 did not participate in this stage.

The inverse binomial distribution model for brickyards disqualification. The fact that the teams left the initial ideas of percentage, proportionality and linearity, and moved to the use of the binomial distribution was a big step in the conceptualization of a random phenomenon, since it implied that they identified that with the binomial distribution they modeled the expected behavior of the waste, as a function of the sample size, n, and with the premise that the average value of the waste from which they had to start was 7%, which is the average for the state of Jalisco. Therefore, a reduction with respect to the average implied that the value of x (defective bricks in the sample) was below the most probable values, which is, below the quantile 0.025 ($L = q_{0.025} = qbinom(0.025, n, p)$).

All teams identified that they could compare the result of the inverse binomial distribution with the waste, x, of each brickyard, and with that identify the disqualified ones. T3 calculated L minus x, and indicated that those with negative results were disqualified, while T6 calculated x minus L and indicated that those with positive results were disqualified (Figure 10). Thus, both teams reached the same conclusion.

T3

Table of results

Brickyard	Sample size	Defective bricks	7% loss	Difference in shrinkage	Result
B1	200	10	7	-3	Disqualified
B2	215	7	8	1	2nd place
B3	315	13	14	1	3rd place
B4	323	16	14	-2	Disqualified
B5	500	22	24	2	1st place
B6	898	49	48	-1	Disqualified

T6

Applied binomial distribution

Brickyard	Sample size	Defective bricks	Loss evaluation	Positioning
B1	200	10	+3	Disqualified
B2	215	7	-1	2nd place
B3	315	13	-1	3rd place
B4	323	16	+2	Disqualified
B5	500	22	-4	1st place
B6	898	49	+1	Disqualified

Figure 10. Results of comparing the waste with the quantile (x vs. L) for each brickyard, performed by T3D and T6D.

Most of the teams (T2, T4, T8 and T10) were able to use the inverse of the binomial distribution to identify the expected range of the behavior of *x*, with a 7% waste. Each team interpreted this range in its own way. For example, T2 indicated, “The disqualification is fair because what the contest is looking for is a waste less than or equal to 7 bricks, in this case it was 10, which is within the range of 7% waste” (T2D, p. 6). T4 put the following annotation in relation to the disqualification of Brickyard 1, “Because we get that the range of 7%, starts at 7 and the waste we have is 10 so it exceeds with 4 and if it is within the range” (T4D, p. 4), and so on for each brickyard. For their part, T3 and T8 were creative to express in a didactic way with a graph when the brickyard should be disqualified. This was when the waste fell in the left tail (Figure 11).

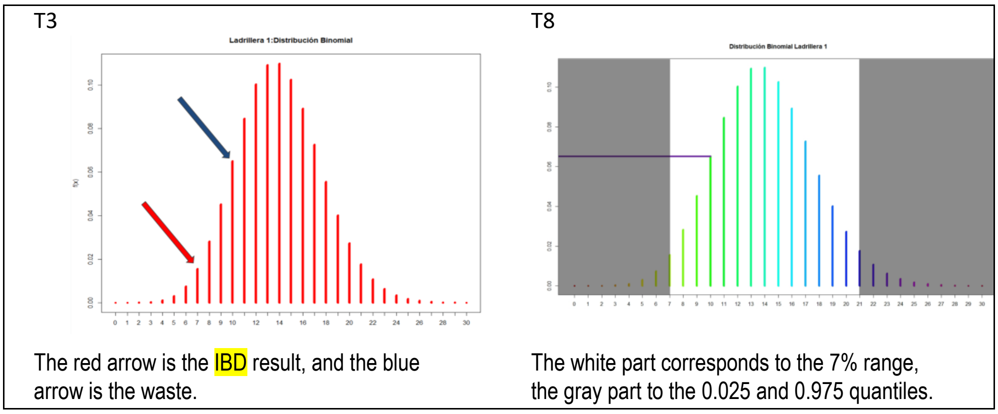

Figure 11. Graphs to explain disqualification [T3D and T8D].

Another aspect to be emphasized is the values of *x* for which it was convenient to make the plot in *RStudio*. At the beginning most of the teams obtained a graph like Figure 4 (distribution panel), where the most probable values of *x* are in a small portion of the initial part of the plot, giving the impression that the distribution has a very long tail on the right. When the teacher/researcher indicated to the teams to check the probability obtained for the *x* values of the right tail (which was zero), some teams (T2, T3, T8 & T10) understood that they only had to include in the plot the values with non-zero probability.

In this case, *RStudio* reinforced the conceptualization of the values that *x* can take, along with those that have more probability; a topic that also causes confusion for many of the students. This agrees with Dos Santos et al. (2014), who pointed out that making a good integration between the use of statistical software to perform the technical calculations facilitates the internalization of the concepts.

The cumulative binomial distribution model for the awarding of the brickyards. Most teams also developed the cumulative binomial distribution model for selecting the winning brickyards (T2, T3, T4, T6, T8 & T10). As observed when using the *Trial-and-error approach with probability distributions*, the cumulative binomial distribution was one of the first options of the students, which facilitated the transition required to identify the winning brickyards, but only when they realized that with the inverse of the binomial distribution it was not easy to order the brickyards in ascending order because a tie between second and third place appeared. In this regard, the argumentation of T4 was the following:

> For example, Brickyard 3, here we use the cumulative binomial distribution, and already because if we do not use that distribution for example in brickyard number three and two gives us a loss of one brick, then they should be tied. But to make the tiebreaker, and good to have a reason for the tiebreaker so in decimals we use the cumulative, then the third place gives us a probability of 0.023 which indicates that the closer it is to zero the farther away from the 7% that is being asked, so that is why it is in third place, because it had 0.023. In second place we have the probability of 0.14, which is close to zero also, so it is in second place; and this is where I was telling you that if we used the inverse binomial distribution for second place and third place it gave us that they had two bricks of waste in favor. This was a tie, so that's how we broke the tiebreaker. [T4P]

As can be seen, the design of the MEA Brickyards allows teams to evaluate if they were the right solution or if they should look for another way, which is what they did when they saw that with the results of the inverse of the binomial distribution, Brickyards 2 and 3 were tied. As indicated by Lesh et al. (2000), the MEA should encourage students to constantly evaluate whether they are on the right path or must rectify to achieve the objective.

6. CONCLUSIONS

In the first attempt to solve the MEA Brickyards, the students used strategies based on ideas of proportionality and percentages. This became an obstacle, since in the MEA used, it is necessary to model a random phenomenon with the support of the binomial distribution to make appropriate decisions. Dooren et al. (2003) called this phenomenon the illusion of linearity, which many students use to solve probabilistic problems, and that this linear thinking is difficult to modify even after an intervention (Konold, 1995).

The MEA Brickyards was designed in such a way that the solution with percentage or proportionality was not a correct form of solution, so when someone adopts it, as happened in the first and second stages of this MEA, it becomes easier to convince the students to look for another alternative. In this sense, to move from the linear or proportional model to the probabilistic model, students must conceptualize several things: first, that the information comes from samples, and as such their results are variable. Second, that this variability can be modeled or predicted under certain assumptions, in the sense that the most probable values of the random variable (x) can be established as a function of the sample size and the average level of waste (7%). With this, an interval with a 95% probabilistic coverage can be established where the most probable values can fall, and thus there will be a decrease in waste when the observed value of the number of defective bricks is below this interval, which corresponds to the left tail of the binomial distribution. Third, it is possible to compare samples of different sizes with this distribution.

The MEA Brickyards allowed students to understand these concepts and led them to model the problem with the inverse of the binomial distribution to identify the disqualified brickyards and with the cumulative binomial distribution the awarded ones. Although the MEA Brickyards can be approached from a more sophisticated decision-making perspective, the first step is to understand that the problem is related to a random phenomenon that can be characterized by the binomial distribution. This understanding was reached by the students through the activity.

A conflict that appeared from the first stage of implementation of the MEA was related to the samples. Most teams proposed to equalize the sample size so that the contest would be fair, and then they used *RStudio* and plotted the distribution using the whole domain of the function (all possible values of x). From this it can be inferred that understanding about the relationships among population, sample and distribution, is a key but complex issue. Therefore, it is necessary to expose students to more situations in which they can confront their beliefs about the characterization of a random phenomenon through probability distributions.

Conceptualizing random phenomena is not simple, so attention should be paid to the way in which data are presented and analyzed, emphasizing their essence of variability, and exposing students to more situations that may be close to them, where such variability appears. In this regard, as future research, we envision using the MEA Brickyards to delve more deeply into students' thinking when they encounter problems involving random phenomena. Another possibility is to utilize the MEA Brickyards to explore the application of more sophisticated tools for decision-making under uncertainty, such as Bayesian statistical inference. We also contemplate the possibility of designing didactic sequences that include simulation activities of random processes to assist students in transitioning from the paradigm of proportional or linear reasoning to the paradigm of random phenomena (Budgett & Pfannkuch, 2018; Garfield et al., 2012; Noll et al., 2018), where the MEA Brickyards could be used to reinforce or assess progress on the transition.

REFERENCES

Batanero, C., Tauber, L., & Sánchez, V. (2001). Meaning and understanding of the normal distribution in an introductory data analysis course. *Quadrante*, *10*(1), 59–92. https://www.ugr.es/~batanero/pages/ARTICULOS/Quadrante.pdf

Ben-Zvi, D., & Makar, K. (2016). International perspectives on the teaching and learning of statistics. In D. Ben-Zvi & K. Makar (Eds.), *The teaching and learning of statistics* (pp. 1–10). Springer.

Bill, A., Watson, J., & Gayton P. (2009). Guessing answers to pass a 5-item True/False test: Solving a binomial problem three different ways. In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides.* Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia, Palmerston North (pp. 57–64). MERGA.

Brady, C. (2018). Modelling and the representational imagination. *ZDM Mathematics Education*, *50*(1–2), 45–59. http://doi.org/10.1007/s11858-018-0926-4

Budgett, S., & Pfannkuch, M. (2018). Modeling and linking the Poisson and exponential distributions. *ZDM Mathematics Education*, *50*(7), 1281–1294. https://doi.org/10.1007/s11858-018-0957-x

Devore, J. L. (2011). *Probabilidad y Estadística para Ingeniería y Ciencias* [Probability and Statistics for Engineering and the Sciences] (8th ed.). Cengage Learning.

Dooren, W. V., De Bock, D., Depaepe, F., Janssens, D., & Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics*, *53*(2), 113–138.

Dos Santos, R., Yumi, V., & Karrer, M. (2014). Teaching probability with the support of the R statistical software. *Statistics Education Research Journal*, *13*(2), 132–147. https://doi.org/10.52041/serj.v13i2.286

Florensa, I., García, F., & Sala, G. (2020). Conditions for teaching mathematical modelling: Case studies at different educational levels. *AIEM - Advances in Mathematics Education Research*, (17), 21–37. https://doi.org/10.35763/aiem.v0i17

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K–12 curriculum framework*. ASA. https://www.amstat.org/asa/files/pdfs/gaise/gaiseprek-12_full.pdf

Garfield, J., del Mas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. *ZDM Mathematics Education, 44*(7), 883–898. https://doi.org/10.1007/s11858-012-0447-5

Gutiérrez, H. (2020). *Calidad y Productividad* [Quality and productivity] (5th ed.). McGraw-Hill.

Hines, W. W., Montgomery, D. C., Goldman, D. M., & Borror, C. M. (2008). *Probability and statistics in engineering*. John Wiley & Sons.

International Organization for Standardization. (2015). *Quality management systems: Requirements*. International Organization for Standardization (ISO 9001:2015). https://www.iso.org/obp/ui/es/#iso:std:iso:9000:ed-4:v1:en

Instituto Nacional de Ecología y Cambio Climático. (2016). *Análisis de mercado del sector de la construcción y proyecto piloto a nivel región, basado en un portafolio de políticas públicas con el objetivo de reducir los contaminantes climáticos de vida corta (ccvc) de ladrilleras artesanales en México. Informe final* [Market analysis of the construction sector and pilot project at the regional level, based on a portfolio of public policies with the objective of reducing short-lived climate pollutants from artisanal brick kilns in Mexico. Final report]. Instituto Nacional de Ecología y Cambio Climático.

Justice, N., Le, L., Sabbag, A., Fry, E., Ziegler, L., & Garfield, J. (2020). The CATALST curriculum: A story of change. *Journal of Statistics Education, 28*(2), 175–186. https://doi.org/10.1080/10691898.2020.1787115

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, *3*(1), 1–9. https://doi.org/10.1080/10691898.1995.11910479

Lesh, R. (2010). Tools, researchable issues and conjectures for investigating what it means to understand statistics (or other topics) meaningfully. *Journal of Mathematical Modelling and Application 1*(2), 16–48.

Lesh, R., & Doerr, H. (2003) Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh & H. Doerr (Eds.), *Beyond constructivism models and modeling perspectives on mathematics problem solving, learning, and teaching* (pp. 3–34). Lawrence Erlbaum Associates.

Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In A. Kelly & R. Lesh (Eds.), *Research design in mathematics and science education* (pp. 591–646). Lawrence Erlbaum Associates.

Noll, J., Clement, K., Dolor, J., Kirin, D., & Petersen, M. (2018). Students' use of narrative when constructing statistical models in TinkerPlots. *ZDM Mathematics Education*, *50*(7), 1267–1280. http://doi.org/10.1007/s11858-018-0981-x

Noll, J., & Kirin, D. (2017). TinkerPlots model construction approaches for comparing two groups: Student perspectives. *Statistics Education Research Journal*, *16*(2), 213–243. https://doi.org/10.52041/serj.v16i2.191

Pfannkuch, M., Ben-Zvi, D., & Budgett, S. (2018). Innovations in statistical modeling to connect data, chance and context. *ZDM Mathematics Education*, *50*(7), 1113–1123. http://doi.org/10.1007/s11858-018-0989-2

Sevinc, S., & Lesh, R. (2018). Training mathematics teachers for realistic math problems: A case of modeling-based teacher education courses. *ZDM Mathematics Education*, *50*(7), 301–314. https://doi.org/10.1007/s11858-017-0898-99

Stemock, B., & Kerns, L. (2019). Use of commercial and free software for teaching statistics. *Statistics Education Research Journal*, *18*(2), 54-67. https://doi.org/10.52041/serj.v18i2.140

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105–110. http://doi.org/10.1037/h0031322

Walpole, R., Myers, R., Myers, S., & Ye, K. (2012). *Probabilidad y estadística para ingeniería y ciencias* [Probability and statistics for engineering and science] (L. E. Pineda, Trans.; 12th ed.). Pearson Educación de México.

MARTHA ELENA AGUIAR BARRERA
Marcelino García Barragán #1421, esq Calzada Olímpica,
C.P. 44430, Guadalajara, Jalisco, México.

HUMBERTO GUTIÉRREZ PULIDO
Marcelino García Barragán #1421, esq Calzada Olímpica,
C.P. 44430, Guadalajara, Jalisco, México.

VERÓNICA VARGAS ALEJO
Marcelino García Barragán #1421, esq Calzada Olímpica,
C.P. 44430, Guadalajara, Jalisco, México.