Kano Model applied to the evaluation a collaborative board game for teaching in environmental education

Marcus Parreiras[1], **Pedro Marques**[2], **Geraldo Xexéo**[1]

[1] LUDES - Programa de Engenharia de Sistemas e Computação
COPPE - Universidade Federal do Rio de Janeiro
Rua Avenida Horácio Macedo 2030, CT, Bloco H, sala 319
Rio de Janeiro, RJ - Brasil

[2]Instituto de Computação
Universidade Federal do Rio de Janeiro
Avenida Athos da Silveira Ramos, 274, CCMN
Rio de Janeiro, RJ - Brasil

{mparreiras,xexeo,pedromn}@cos.ufrj.br

Abstract. *This article discusses the application of the Kano Model to a collaborative educational board game focused on environmental sustainability. A satisfaction survey was conducted according to the model. The results point to a general satisfaction of the players towards the game's attributes. Analyzing the results, the Kano Model offers an excellent option to analyze the attributes of a game from the perspective of the players as a product, and can be a useful tool to understand players and position efforts for better results in future projects.*
Keywords. *Kano Model, Educational Game, Board Game*

1. Introduction

This article describes the use of the Kano Model to evaluate Salve à Terra!, an educational Collaborative Board Game about sustainability and environment, aimed at the undergraduate and high school student. Salve à Terra! does not favor competition between players: everyone pursues the same goal, and will win or fail together.

Generally speaking, similar works evaluate games using the Likert Scale. However, this method is more appropriate to gather subjective data, such as the opinion of players with each characteristic of the game, than to quantitatively test the satisfaction or dissatisfaction of players with these same characteristics. Therefore, we did two different surveys to test the game: one using the Likert Scale, to survey and analyze players' opinion about how much fun they had or think they learned from the game; and another using the Kano Model, to test the game attributes and mathematically measure the satisfaction and dissatisfaction of the players with the design decisions taken.

This article has been divided into the following sections. The second section introduce important concepts related to this paper, along with a brief list of similar works. In the third section, the methodology is exposed. The results and discussion of the work are presented in the fourth section. Finally, the last section shows the conclusions reached with the development of the research.

2. Related Concepts

This section introduces three main concepts related to this paper: Likert Scales and Validation Criteria, Kano Model and Collaborative Board Games.

2.1. Likert Scale and Validation Criteria

The Likert Scale is a symmetric scale widely used to check opinion with general content in games. The tool measures each attribute on a scale of 1 to 5, with 1 being "slightly relevant" and 5 being "very relevant". The answers obtained using the questionnaire in the Likert model were treated with the Kappa Index (K) and Content Validity Index (CVI) content validation criteria, in order to ensure the reliability and data consistency.

The Kappa Index ranges from -1 to +1, so it is more consistent the closer it gets to 1. A negative value indicates a poorly consistent criterion – that is, respondents disagree too much in their answers. The index suggests that a value greater than 0.6 is considered sufficiently consistent. The Content Validity Index is the percentage of agreement among respondents for each attribute, and, for it, a minimum of 60% agreement was considered acceptable.

2.2. Kano Model

The Kano Model is a powerful tool that classifies the attributes of a given product into categories according to their ability to generate satisfaction or dissatisfaction to the users [Asian et al. 2019]. For example, the model might point out that a customers of a 60-inch Smart TV gives much more importance to the finishing than to the image quality itself. Therefore if a TV brand wants to increase the satisfaction with the least amount of investment, it should raise the quality of the finishing.

[Paraschivescu and COTÎRLEŢ 2012] suggests that the modelling starts with the creation of a survey about the attributes of the product or service. For each attribute, two responses are collected: the Functional response, which measures the level of customer satisfaction in having that attribute in the product; and the Dysfunctional response, which measures the level of dissatisfaction for not having that attribute in the product. Figure 2 exemplifies these questions.

Figure 1. Example of Functional and Dysfunctional questions from the Kano Model. Source: adapted from BERGER, 1993

Then, using Table 1 as reference, for a particular costumer, each attribute should be classified as:

- **Attractive (A):** The higher the performance, the greater the satisfaction, however, it does not bring dissatisfaction if absent. Example: longer manufacturer warranty period than conventional.

Table 1. Classification of Attributes in the Kano Model. Source: adapted from ROOS, 2009

		Dysfunctional Response				
		I like	Must be	I'm neutral	I can bear	Don't Like
Functional Response	I like	Q	A	A	A	O
	Must be	R	I	I	I	M
	I'm Neutral	R	I	I	I	M
	I can bear	R	I	I	I	M
	Don't Like	R	R	R	R	O

- **Must-Be (M):** It does not bring satisfaction according to its positive performance, but its absence or insufficient performance brings dissatisfaction. Example: hygiene in a restaurant.
- **One-dimensional (O):** The higher the degree of performance, the higher the customer satisfaction and the lower the performance, the lower the customer satisfaction. Example: comfort in a car.
- **Indifferent (I):** They do not interfere with satisfaction regardless of performance. Example: product factory model of the product.
- **Reverse (R):** The lower the performance grade, the higher the customer satisfaction and the higher the performance, the lower the customer satisfaction. Example: price of home appliances.
- **Questionable (Q):** Represents an inconsistent answer, usually arising from the respondent's not understanding the question.

For each attribute, the percentage of each classification is calculated using data from all costumers. Through these percentages, subsequently, the Satisfaction Coefficient (CS) and the Dissatisfaction Coefficient (DS) are calculated [Bae and Shin 2019], according to the following equations:

$$CS = \frac{\%A + \%O}{\%A + \%O + \%M + \%I} \qquad DS = \frac{(\%O + \%M) \times -1}{\%A + \%O + \%M + \%I}$$

These coefficients are then used to create the Better-Worse Diagram, which positions each attribute between the Attractive, Must-Be, One-dimensional or Reverse quadrants. The Indifferent and Questioning classifications, as they do not represent an accentuation of satisfaction or dissatisfaction, are not part of this diagram.

2.3. Collaborative Board Games

In Collaborative Board Games (CBG) the players do not compete with each other in order to win the proposed challenge. Instead, they collaborate to jointly articulate their strategies and the specific mechanics available to achieve a common goal. Recent examples include Pandemic and Forbidden Island, created by Matt Leacock. [Linderoth 2011] comments that it is common that in these games the board offers the central obstacle, as it expands or changes, in the way that leads to the defeat conditions of the game. It is up to the players to create strategies to avoid this.

The main reason to choose this particular genre is to support the main argument of the game: environmental sustainability depends on individual actions in favor of the collective, and that there is no way to overcome the problem alone were the main reasons to build the game in this genre. In addition to that, there is a significant increase in interest for this type of game [Suominen et al. 2016].

As similar research, the card games by [Steinman and Blastos 2002] stand out, which aim to teach about immune systems, [Rowe 2001] about mathematics, [Odenweller et al. 1998] focused on natural sciences , and [Byrne and Fielding-Barnsley 1991] in the literacy process. Using card games as an educational tool, not just about sustainability, but about any topic, seems to be a promising way to attract attention and teach at the same time.

3. The Game

In this section we give a brief introduction to Salve à Terra! rules. The game was created to be strongly collaborative, and challenging.

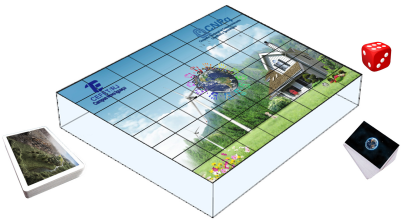

Figure 2. Salve à Terra! board

The game components consists of one board of 8x8 cells, one set of Energy Cards, one set of Junk Cards, one set of Pollution Cards, one set of Good Action Cards, stamped coins that symbolize players, stamped coins that symbolize markers, and some 6-sided dice.

At the beginning of the game, one of the players shuffle Junk and Energy Cards. This new set is called Field Cards. All cells of the board are filled with one face down Field Card, except the four corners. Players then distribute their tokens on those remaining corners.

In a clockwise direction, one by one player rolls the die, which will dictate the number of steps he will move their tokens. If it lands on a face-down card, the player flip that card. If the player lands on a Junk Card, it will be face up until removed. To remove it, the player must use a Good Action Card.

Good Action Cards can be drawn at the end of each player's round, and to use it, he will need to pay a number of collected Energy Cards. But, when trying drawn a Good Action Card, it is also possible that a drawn card is a Pollution Card. Each Pollution Card has a set of mechanics that hinder the game (putting more Junk Cards on the field, consuming Energy Cards, preventing token mobility, etc.). As a cooperative game, all players win together if they are able to remove all junk cards from the board, or lose together if they takes up all space.

Many important topics are explained to the players in the description of more the cards. Their strategic use is related to their concepts, so to be good at the game the player must have a basic understanding of the topics. For example, the Good Action Card "Reuse" talks about the concept of the sustainability of reusing the product in the same function or not, and allows the player to reuse an already discarded card; the Good Action Card "Rechargeable Batteries" explains the benefits of this type of pile and is able to transform Junk Cards into Energy Cards; the Pollution Card "Radiation" warns about the dangers of itself and the time it takes the environment to recover from radioactive waste and in the game prevents Junk Cards from being removed for several rounds.

4. Methodology

This research was divided in two different surveys: one to measure the game's success, as well as to test the hypothesis of it being able to transmit knowledge about the theme, using the Likert Scale, and statistical methods of validation of data consistency Kappa Index (K) and Validity Index of Content (ICV); and another to analyze the degree of satisfaction and dissatisfaction with the game's attributes, using the Kano Model.

The evaluation took place at CEFET/RJ, during the EXPOTEC/EXPOSUP, in Introduction to Engineering and Sustainability classes. After having played the game at least once, all respondents received the forms by email. The participants are students from higher education and high school.

It is noteworthy too that there was no coercion for students to respond to the questionnaire, in order to reduce as much as possible any bias in the responses obtained. To endorse this, the questionnaires were anonymous, collecting only the respondent's educational level. Not all players responded to the online surveys, so respondents are not necessarily the same for both surveys.

4.1. Survey according to Likert Scale

The sample size for this survey was 22 students: 9 undergraduate students and 13 high school students. The questionnaire measures the degree of satisfaction of players in the following aspects: game duration, card and text design, interactivity and ability to hold attention, and contribution to knowledge in sustainability. The questionnaire was also accompanied by a space for suggestions, criticisms and improvements to the research

4.2. Survey according to Kano Model

The sample size for this questionnaire was 21 students, being 10 undergraduate students and 11 high school students. The survey was carried out according to the Kano Model, which intended to measure the level of satisfaction and dissatisfaction with the presence or absence of each attribute listed in the project. The questionnaire consisted of ten questions, five attributes that were divided into their functional and dysfunctional form, these attributes being the following characteristics of the Salve à Terra!: being educational, serving as entertainment, being challenging, having a low cost, and having a sustainability thematic

5. Results and Discussion

Table 2 summarizes the CVI and Kappa validation results calculated for each of the four questions of the questionnaire according to the Likert Scale. Considering that no Kappa

Table 2. Consolidation of the validation of the Likert Scale results relating Kappa and CV

Question	Kappa	IVC	p-value	Confidence Interval(95%)	
1 - Duration	0,66	58%	<0,001	Superior	0,72
				Inferior	0,22
2 - Design	0,74	71%	<0,001	Superior	0,82
				Inferior	0,32
3 - Interactivity	0,77	72%	<0,001	Superior	0,81
				Inferior	0,11
4 - Learning	0,98	99%	<0,001	Superior	1,00
				Inferior	0,4

Table 3. Consolidation of respondents' responses from high school regarding the Likert Scale

Question	Excellent	Good	Neutral	Base	Terrible
1 - Duration	10	2	1	0	0
2 - Design	9	4	0	0	0
3 - Interactivity	10	1	2	0	0
4 - Learning	9	4	0	0	0
Total	38	11	3	0	0
%	73%	21%	6%	0,00%	0,00%

result was less than 0.61, it is shown that there is sufficient consistency to state that the agreement of the data collected classifies it as valid by this method of validation. There is a contradiction, however, in the result of the Content Validity Index of question 1, where ICV $< 60\%$. The phenomenon is explained by the wide dispersion of the result of the question about the duration of strategy and board games: some will give maximum marks even if games like War and Monopoly last for hours; many others, however, will hate this feature. This subjectivity is not expected in the other criteria. It is not expected, for example, that half of the players are like the design used, and the other half doesn't like it.

Table 4. Consolidation of respondents' responses from higher education regarding the Likert Scale

Question	Excellent	Good	Neutral	Base	Terrible
1 - Duration	7	1	1	0	0
2 - Design	7	2	0	0	0
3 - Interactivity	7	0	2	0	0
4 - Learning	7	2	0	0	0
Total	28	5	3	0	0
%	77,78%	13,89%	8,33%	0,00%	0,00%

Figures 3 and 4 summarize the responses to the first questionnaire regarding how much they liked and how much they enjoyed the game. In general terms, there is a consistency and uniqueness between both tables, with higher education having answered 78% "great" while high school answered 73%, and the answers as "neutral" are lower in high

Table 5. Consolidation of respondents' responses regarding the attributes of the Kano Model

Attributes	A	O	M	I	R	Q
1 - Being educational	6%	75%	13%	6%	0%	0%
2 - Serving as entertainment	0%	44%	50%	6%	0%	0%
3 - Being challenging	44%	13%	6%	31%	6%	0%
4 - Having low cost	13%	25%	37%	25%	0%	0%
5 - Being about sustainability	12%	44%	19%	25%	0%	0%

Table 6. Consolidation of respondents' responses regarding the attributes of the Kano Model

Attributes	CS	DS
1 - Being educational	0,81	-0,88
2 - Serving as entertainment	0,44	-0,94
3 - Being	0,60	-0,20
4 - Having Low Cost	0,38	-0,63
5 - being about sustainability	0,56	-0,63

school, with 6% against 8% in higher education. There were no "bad" or "very bad" answers, either in higher or secondary education, in any of the interviews, in the face-to-face or electronic questionnaire. It can be seen that despite a performance of more than 72% as excellent in any attribute, the attribute referring to question 1 – about the duration of the game – has a more equal division between "good" or "neutral" answers. Question number 3 – about interactivity – had 18% of "neutral" responses, thus representing a higher potential for improvement than the others. Questions 2 and 4, coincidentally, generated identical graphs, although the responses of the samples from secondary and higher education were different – given that the sum is equal.

Contrary to what might be imagined at first, the results on how much was learned from the game did not differ much regardless of the sample group: both high school and college level students responded that they were able to learn a lot from the game, with a variation of less than 10% between each group. This also implies that, as proposed, the project is capable of more than informing new things: it serves to fix ideas and give them a new look.

Table 5 refers to the results of the attributes of the game itself, and they classify each one according to the Kano Model. It summarizes the percentage of each rating of individual respondents' responses.

Table 6 shows the coefficients of satisfaction with the presence and dissatisfaction with the absence of each attribute. There is great satisfaction with being an educational game, as well as great dissatisfaction if the performance in this criterion is low.

The Figure 3 shows the Better-Worse Diagram, which uses the satisfaction and dissatisfaction coefficients to represent and classify each game user's need in the Attractive, One-Dimensional, Indifferent and Must-Be quadrants.

It can be said that, due to the validation criteria used and the confidence adopted, the results obtained by the Likert Scale questionnaire are acceptable, and they point to

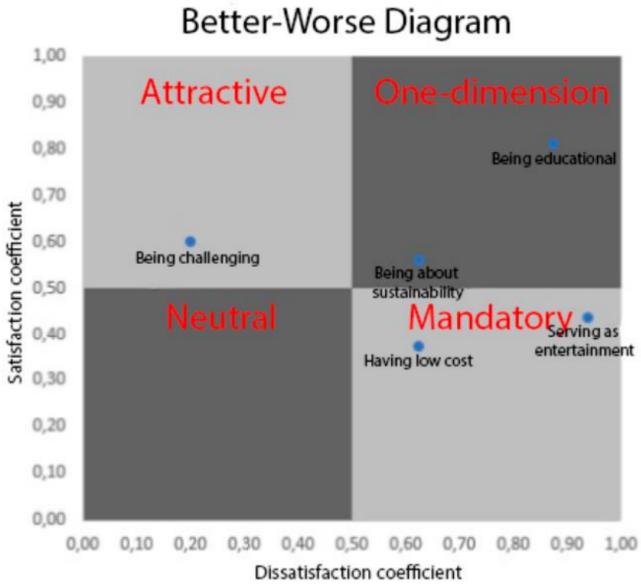

Figure 3. Better-Worse diagram, representing the satisfaction and dissatisfaction coefficients

an almost absolute success and acceptance by the players. In all data collection opportunities, no attribute was considered bad or terrible by any of the respondents. In fact, few considered any criterion as good, and even less as neutral: the overwhelming majority considered each of the criteria to be excellent, whose percentage share is greater than twice the sum of all other classifications, both for high school and for the higher. Based on this, a conclusive positive result is reached, where the tested hypothesis is confirmed: Salve à Terra! was able to serve as a learning tool, while entertaining and keeping players' attention.

6. Conclusion

There are two groups of results: the conclusions about the game success using Likert Scale and the conclusions about the application of Kano Model to evaluate games.

For the first group, the responses obtained suggest the overall success of the game's design decisions, interactivity and educational capability, and the data seem consistent considering both validation tools used. The most interesting results are the second group, which is also the main part of the survey.

In second group of results, derived from the application of the Kano model, it was noticeable which game attributes are more relevant with regard to customer satisfaction. This notion is especially important to encourage one or another investment in improving some characteristics.

Half of the respondents considered, through the combination of their functional and dysfunctional responses, "serving as entertainment" as a MANDATORY requirement for the game. This implies that, necessarily, the game needs to be fun, even if all respondents are aware that it is a game with a primarily educational purpose, and are academically active people.

As the previous attribute, “Having a low cost” was considered MANDATORY, so it did not bring satisfaction for being present, but their absence would lead to dissatisfaction. It is also clear that having a low cost is less important in terms of obligation than being fun, for respondents. That is, people are more willing to pay more than they would like for the game than they would be willing to play a game with a better price but less fun. This is an interesting conclusion especially considering that it is an educational game.

“Being about sustainability” attribute is very close to the center of the graph, resulting in a slightly positive proportion, being the most neutral criterion for most, quite neutral in relation to the theme. This could mean an initial awkwardness with the topic, or simply a lack of immediate interest in playing an educational board game. In this case, it reinforces the idea that games of this type in education are a potential field and little explored. This may also suggest that most people were not particularly excited about the game because of this theme, which could be seen as something popular, not only given the extensive debate and importance of the topic, but also because most respondents take related courses to the environment. This conclusion also implies that, probably, other themes could have equally good results, so that this game could be reproduced again for another subject.

The only attribute with a reverse percentage is “being challenging”: 6.25% of respondents are less satisfied if the game is more difficult. This can be explained with individual player type preferences. More casual players may prefer games that are shorter and easier to win, while other players may prefer games that are as challenging as possible. Given that the game design chose to create a challenging game in order to allow players to articulate strategies by studying the cards (favoring card reading, fixing sustainability concepts and cooperation), it was reasonable that some players didn’t like it.

Despite this, it could be seen that the attribute was considered ATTRACTIVE, so that his presence increased players’ satisfaction, but his absence did not lead to dissatisfaction. This conclusion is interesting because it demonstrates that there is a need to include challenges even in educational games, but that these must be measured so as not to create a situation that diverts attention from the main focus.

The most expressive attribute is “being educational”, with 75% being classified as ONE-DIMENSIONAL. It means that 3 out of 4 players will be more satisfied if the game can be more educational, and will be more dissatisfied if it is less. This conclusion matters to decide where to invest the resources in a safe way. As this attribute was quite expressively one-dimensional, every extra effort could be invested here, avoiding bad ratings (which, as one-dimensional, would bring dissatisfaction) and maximizing the attribute (which would necessarily bring more satisfaction). For example, if the game is finished with a considerable time slack, a safe decision would be to use that time to maximize this attribute.

A limitation of the study is that it does not consider the current performance of each attribute listed for the Kano Model in relation to other existing products of the same type. The model does not exogenously predict the market or the scenario where the product is located, so it only indicates which criteria should be encouraged, and does not measure their current individual efficiency.

Another limiting factor was the finite number of samples, which are natural to a

new game and which lasts more than an hour on average each game, limiting to 1 to 4 respondents for every 1 or 2 hours of questioning. Analyzing the results, the Kano Model offers an excellent option to analyze the attributes of a game from the perspective of the players as a product. Although little used for this, the model mathematically considers player responses and can be a useful tool to understand players and position efforts for better results in future projects.

References

Asian, S., Pool, J. K., Nazarpour, A., and Tabaeeian, R. A. (2019). On the importance of service performance and customer satisfaction in third-party logistics selection: An application of kano model. *Benchmarking: An International Journal*.

Bae, J.-H. and Shin, H.-Y. (2019). A study on the factor of satisfaction or dissatisfaction of e-learning using kano model and timko's customer satisfaction coefficients. *Journal of the Korea Convergence Society*, 10(7):325–333.

Byrne, B. and Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational psychology*, 83(4):451.

Linderoth, J. (2011). Exploring anonymity in cooperative board games. In *DiGRA Conference*. Citeseer.

Odenweller, C. M., Hsu, C. T., and DiCarlo, S. E. (1998). Educational card games for understanding gastrointestinal physiology. *Advances in Physiology Education*, 275(6):S78.

Paraschivescu, A. O. and COTÎRLEŢ, A. (2012). Kano model. *Economy Transdisciplinarity Cognition*, 15(2).

Rowe, J. (2001). An experiment in the use of games in the teaching of mental arithmetic. *Philosophy of Mathematics Education*, 14:1–23.

Steinman, R. A. and Blastos, M. T. (2002). A trading-card game teaching about host defence. *Medical education*, 36(12):1201–1208.

Suominen, A., Rilla, N., Oksanen, J., and Still, K. (2016). Insights from social network analysis–case board interlocks in finnish game industry. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 4515–4524. IEEE.