

Crowd Voting: Impacts on Product Sales in Marketplaces

Milton Ruiz Rodrigues Alves¹

¹ School of Management, Pontifícia Universidade Católica, São Paulo, Brazil

Correspondence: School of Management, Pontifícia Universidade Católica, São Paulo, Brazil. Tel: 55-11-98315-4980.

Received: March 30, 2019

Accepted: April 15, 2019

Online Published: April 17, 2019

doi:10.5430/ijba.v10n3p90

URL: <https://doi.org/10.5430/ijba.v10n3p90>

Abstract

In the current time of information, we live in times that the expectations of thinkers as Lévy (1998) have become a reality. The profile of the consumer has changed as he connected with other consumers and begun to demand that there is a dialogue between the company and society. This new scenario required that technology was further developed and that the process of communication was made faster by means of online platforms that make the exchange of information viable and encourages it. The crowdsourcing arose with the availability of the crowd's intellectual capital that has changed the way companies interact and organize themselves as they come to realize the strategic importance of this information. In order to fully understand this new scenario, we have created three hypothesis that have served as guide for this paper, which are: H1 – Crowdsourcing affects products' performance; H2 – negative reviews from crowd voting provide the same effects that positive reviews; and H3- Crowd voting provides information about consumer's perception of the product being offered. Therefore, this paper aims to understand if the crowdsourcing affects product's performance of the companies that use crowdsourcing by analyzing the theoretical basis of crowdsourcing and also by reviewing a data-base with reviews made by consumers during a four-year period. Based on the review of the correlations between the variables obtained in Amazon's data base based on the reviews offered by the consumers, this work suggests that the crowdsourcing does affect the product's performance.

Keywords: crowd sourcing, crowd voting, open innovation

1. Introduction

1.1 The Problem

As a result of the globalization of the economy, there has been a significant intensification of competition between companies in all markets, forcing them to adapt more and more rapidly to demands, accelerating their innovation cycle and optimizing their operating costs to survive in this new economic environment. Add to this the occurrence of several financial crises around the world, to the example of the financial crisis of 2008 that increases uncertainties and commercial disputes.

In this scenario, a new variable, which represents the acceleration of the cycle of innovation and development of companies, together with the greater proximity of its target public, thus allowing a greater understanding of their wants/needs becomes not only desirable, but also represents a new priority to be assimilated by companies, so that with the knowledge of their process, can determine if this represents a competitive advantage.

Prahalad and Ramaswamy (2004, p. 1) described this shift in the role of the consumer by highlighting its antagonistic character: (...) from isolated to connected, from unaware to informed, from passive to active ". Howe (2008) called this change in the rebirth of amateurism, characterized by the more active role of the consumer/individual in society, in which he uses his free time not only in non-economic distractions but also in economically productive activities and content creation.

In this way, it is possible to affirm that civilization moves towards the formation of a collective memory/intelligence, as suggested by Levi (1998). The composition of these multidisciplinary groups is observed informally inside and outside organizations, and on several occasions has its potential underutilized. The correct use of this synergy can configure a competitive advantage that may allow the improvement of the process of questioning and search for solutions, so vital for the innovation and competitiveness of companies.

At this point, it is important to note that most of today's organizational structures are organized into business units that specialize in taking responsibility for solving specific problems. In the training of these teams, it is sought to train people in certain areas of knowledge to act as specialists. Often, therefore, specialized professionals have their visions limited not only by their organizational context but also by their life experience, which, while providing a unique and unique view, is often incomplete and does not meet the perspectives of its audience target.

It is often the case that solving complex problems is not limited to a specific field of human knowledge as it demands a broader analysis by multidisciplinary teams, which by their nature heterogeneity, self-regulate to avoid simplistic or biased solutions. This approach allows better use of the company's intellectual capital, which can be characterized, although intangible in its nature, as one of the vast wealth of the organization (Sveiby, 1998).

There is, therefore, a need for a more complete approach to systemic problems, which more accurately reflects the scenario in which organizations are situated today, a globalized world in which information is shifted and made available quickly to a great mass of people, of distinct formations and arranged to compose a collaborative community.

On this new context Brabham (2008) defined crowdsourcing as a multicultural group of individuals who share their time and knowledge in solving a problem posed by a group that provides data and guides the process, so that through knowledge of the crowd, to obtain answers and points of view that would not be found only within the corporate environment. Such a competitive advantage can be both to allow the development of a new business strategy and to enable the company to understand the market strategies of its competitors better and to position itself efficiently.

Thus, the importance of the strategic planning of the intellectual capital contained inside and outside the organizations is evident, in which its measurement by means of indicators, and the creation of an environment in which there is the tool and the cultural and social attitude in as well as the interaction between clients, suppliers and the body of the organizations point to a differential that provides a tremendous competitive advantage (Edvissen and Malone, 1998).

In this scenario, crowdsourcing presents itself as a useful tool to face the challenges that an organization experiences in the formulation and application of its production strategies or the practice of a low-cost marketing plan that is effective and presents different results. In turn, it is observed that this structure also presents certain disadvantages not identified by groups of current specialists and this scenario requires a more detailed analysis of the pros and cons offered in the use of crowdsourcing.

1.2 Relevant Scholarship

Through the study of organizations and the various proposals on organizational structures (Daft, 2003), companies seek to organize themselves to obtain the best response to the problems they must solve and the processes they must undertake (Salerno, 1999). Within these concepts, companies typically formulate business units in which groups of specialists address specific problems that negatively impact or hamper their performance.

Such groups, generally homogeneous, have a very close academic formation, if not the same and, thus, end up developing similar solutions. This situation is portrayed by Page (2009) that studied problems generated by organizations that contain structures formed by experts in a globalized world. The issues faced by large companies today tend to function as a complex system that is listed by many variables, which are often best handled by multidisciplinary groups. In this scenario, the use of collaborative internal and external groups represents a competitive advantage in solving problems.

(2009, p.156) points out a possible lead in the work performed by multidisciplinary teams as follows: "Two heads do not think better than one when together they form a single head." Already in a heterogeneous group - possible to obtain in an open community - corresponding profiles substantially improve the capacity of problem-solving. In this sense, Page (2009) concludes that the community collects the knowledge dispersed in each so that the crowd is wiser than the most intelligent individuals.

In the same sense, Howe (2008) argues that the availability of a broader force of knowledge provides companies with a universe of information and resources that far exceed their structure, both in numbers of people willing to focus on a problem and about the analysis of complex issues by a team of specialists.

Globalization and the dissemination of knowledge, favored by the internet and the current communication tools, reorganized the concept of community. According to Lévy (1998), this adaptation favors the formation of collective memories that can be analyzed through knowledge tree models. Within this universe, we have the creation of collective intelligence, which, when well used, can provide creative and innovative solutions that adapt to the reality

of companies in this dynamic and competitive market. In this way, knowing how to exploit the diversity of these communities can be the key to gaining a competitive edge in business, and this happens by allowing the creation of an environment in which suppliers, customers, and all internal and external staff can interact freely. With a greater approximation of its public, allowing to understand its needs better and to obtain its aid in the development of new strategies through groups willing to collaborate with the companies in exchange of status, sense of inclusion or financial return.

According to Lévy (1998), collective intelligence can be defined as distributed intelligence everywhere, incessantly valued, coordinated in real time, resulting in an effective mobilization of competencies. In a reflection on the theoretical framework of each, an intelligence acquired in the course of their lives in a unique way, each becomes the holder of a unique wisdom, a set of personal experiences that throw a unique vision on any problem that presents itself, there being no monopolizing center of knowledge, it is in its view erroneously repressed by systematic methods of evaluation of human intelligence, and coordinated to deal together with a specific problem, from the collaboration of individual knowledge, in order to a particular result to the issue presented, thus indicating that the sum of the parts becomes more significant than the whole.

Within this universe, we potentiate the exchange of knowledge by going beyond explicit knowledge and exploiting the tacit understanding of each that becomes a variable to collaborate in the search for the solution of a presented problem. According to Nonaka and Takeuchi (1997), we have as explicit knowledge that done by facts, rules and political relations that can be codified and transmitted on a large scale without the need for discussion. This thinking is also corroborated by Wyatt (2001). In turn, tacit knowledge is personal and complicated, originating from the experience of an individual within a specific context (Vitelli, 2003). In the latter, we have the capacities of each and a differentiated transmission, which depends on prolonged interactions, errors and correctness (Castanheira, Peixoto and Cardoso, 2003).

It is assumed that a broad group of people may, in certain circumstances, be more "intelligent" than any of its members, a conclusion that is based on the principles of the Gestalt theory, is confirmed by Wooley et al (2010), concluding that the collective intelligence of a group of people, while exploring the tacit knowledge of this group, extends beyond the individual abilities of each member of this group, and this is known as collective intelligence (Slawsby and Riviera, 2007).

In this context, several theorists have expressed their views on the possible typologies of crowdsourcing and its applications. The principal authors are Howe (2008), Brabham (2010) and Schenk and Guittard (2009, 2011).

1.2.1 Howe's Typology (2008)

Howe was the first author to coin the term crowdsourcing, when he did it generically, as expected of a new concept, adopting it as a term for a very varied group of approaches within the idea of something originating from the contribution of crowds. Within this concept, Howe (2008) describes four categories of crowdsourcing: crowdfunding, crowd voting, crowd creation, and crowd wisdom.

1.2.1.1 Crowd Funding

An approach in which an idea is presented to the masses so that they can finance it, this being able to be used for the most diverse purposes, either for help and support to natural disasters or financing for the launching of new products or promotional campaigns (Bons et al., 2010).

1.2.1.2 Crowd Voting

Typology where a mass of people is given the opportunity to express their opinion through a vote or classification of data. A strategy used in the acquisition of data aimed at the formulation of population preferences and targeted advertising, as well as the organization of large libraries, as is the case of Google that uses the volume of pages with links directed from other sites to categorize the relevancy of the pages in your search engine.

1.2.1.3 Crowd Creation

The third type cited is the crowd creation involved in the production of value from the innovation, or the participatory collaboration of the crowds in front of a task, cases like this can be observed in Wikipedia or YouTube that receives its content, which entails in the maintenance of the platform's customer flow.

1.2.1.4 Crowd Wisdom

In the fourth and last type of typology proposed by Howe (2008), we have the crowd wisdom which is the wisdom of the masses or collective intelligence itself, as pointed out by Lévy (1998). Within this sphere, Howe (2008), proposes

three types of exploitation of the generated content, being:

- Idea jam, which translates into a great brainstorming session where any participant puts into vogue any question/idea that comes to mind.
- Crowd casting, where a person/entity reports a problem previously identified in the mass in the hope that some of the members or their interaction can solve it.
- Markets of forecasts, where the wisdom of the masses is used, to predict events and results with greater foresight, in tune with crowd voting. Predictions using big-date, or number of postings on forums to set a possible winner for some prize is one of the potential applications.

1.2.2 Typology of Brabham (2010)

Brabham (2010) suggests four types of crowdsourcing: Knowledge discovery and management, search diffusion, creative production with vetoes, and task distributions that require human intelligence.

1.2.2.1 Knowledge Discovery and Management

In this approach, an online community is challenged to find some information and derive from it in knowledge, expanding the capacity of the discovery of an organization with few resources. It starts from the philosophy that the necessary expertise exists and is located outside the organization and that through a well-managed process of using the knowledge of the masses to acquire an understanding that is widespread in the network.

1.2.2.2 Search Broadcasting

It is a knowledge-seeking technique in which the answer is likely to exist already, but it is not known by the organization that seeks it, in this way the company aims to find the expert, or expert, in the middle of the mass so that it can guide it in the direction of the answer. Such a process is referred to by Brabham (2010, p.8) as finding "a needle in a haystack."

1.2.2.3 Creative Production With Vetoes

It is a way that through the participation of the masses in the creative process of a company, one can find among thousands of ideas an idea that surpasses the others, in this way the process of veto allows finding the best answers and closing the cone of the productive process immediately, thus shedding part of the market research that would be needed, where the company and the mass participate together in the process of co-creation.

1.2.2.4 Distributions of Tasks That Require Human Intelligence

It is a complex problem-solving platform where they are broken down into tasks with less complexity and divided among the individuals that make up the mass of people collaborating with the project so that human intelligence can be used to approach complex problems in large masses of data. In this process, although human knowledge is necessary, usually the methods after being divided and simplified require low levels of creativity and intellectuality, yet the components of the mass are rewarded for their collaboration (Brabham, 2010).

1.2.3 Typology of Schenk and Guittard (2009, 2011)

Schenk and Guittard (2009) define their typologies through two dimensions of the nature of the process: integrative or selective;

In procedural nature, the system becomes integrative when each collaboration is considered as adding value to the previous one and directed to an expected result. In contrast, it becomes selective when each partnership is compared one by one to choose the best contribution.

In this way, Schenk and Guittard (2009, 2011), propose one of the fundamental distinctions of crowdsourcing. This distinction is intrinsically related to the type of task to be performed by the crowds, which is, on the one hand, can be simple, such as the translation of words, can also present high complexity as the innovation of a product; or creative, such as the routing of advertising pieces.

1.2.3.1 Using Crowdsourcing for Simple Tasks

In this type of work, the added value is not derived from individual abilities, but it happens with the accomplishment of pure functions in large scale with low cost, although the financial incentive does not go beyond a small payment (Schenk and Guittard, 2009). In this way, people often act in crowdsourcing without realizing that they are doing it.

1.2.3.2 Crowdsourcing of Complex Tasks

Complex tasks can be defined as a task that can present several potential outcomes, various solutions and the

presence of uncertainty that require the understanding of the problem and specific knowledge and skills for its resolution (Campbell, 1988). At present, a company can decide whether its internal group of experts can benefit from a more significant number of opinions and different views on the answers to the problem to obtain a more satisfactory resolution to the complexity of the problem (Schenk and Guittard, 2011).

1.2.3.3 Crowdsourcing of Creative Tasks

The companies that opt for this model do not expect resolutions of issues, but to be able to make use of the creativity and the proximity to the public that the masses generate, and due to this characteristic the rewards to the group can be varied, not being bound exclusively to the financial return (Amabile et al., 2005).

You can see this concept in practice in the mobile app stores of the IOS and Android operating systems. The companies provide some tools and tutorials ("step by step") for the creation of applications, reducing the technical knowledge required for their development.

1.3 Hypotheses and Their Correspondence to Research Design

From the exploratory research on crowdsourcing, initially to revisit the theoretical framework and to test the advantage or disadvantage of crowdsourcing impacts (crowd voting within the sales process of a marketplace), the following hypotheses were proposed to understand the effects of this process better:

H1 - Crowdsourcing affects product performance.

H2 - Negative feedback from crowd voting has the same effects as positive feedback.

H3- Crowd voting provides data on what is being perceived by the consumer market versus the advertised product.

2. Method

2.1 Research Quali-Quantity

It is proposed a mixed qualitative and quantitative research, through the collection of secondary data obtained through the Internet, academic articles, case studies, theses and books on concepts related to crowdsourcing, directing the focus of studies to the construction of hypotheses. This approach is based on the reviews of Reto and Nunes (2001), by the theoretical robustness of the subject, the research presented will assume a more exploratory character and the hypotheses generated directed us to the research objectives to be achieved.

This was followed by a comprehensive research strategy, directing investigations through deductive logic, using existing theory to provide facts that would contribute to the proposed theoretical constructs, as well as an inductive reasoning, when events are used to construct a consistent approach with the same circumstances (Reto and Nunes, 2001), thus seeking to understand crowdsourcing in depth.

According to Godoy (1995), qualitative studies assume essentially three paths: Documentary research, case study, and ethnography. Since the objective in this work is a greater understanding of the concept of crowdsourcing and the benefits and harms of its practice by the companies, the case study may be too focused, not allowing a generalist view of the concept, in this way we choose a mixed path, conducting a qualitative and quantitative research, comprised of documentary analysis and direct observation of data through ethnography, since in qualitative research the natural environment is a direct source of data (Godoy, 1995).

For the quantitative analysis, we used the secondary data provided by the company Amazon, through Stanford University. It was understood that to test the proposed hypotheses a database was necessary to evaluate this phenomenon in action. In the search for this base, we contacted three companies with operations based on the Internet: Amazon, Submarino and Saraiva. Of these, only one returned the contacts made; the Amazon company responded to our communication and directed us to submit requests to Stanford University, which provided access to an eighteen-year database

2.2 Sample Size, Power and Precision

The database provided was from Amazon company and contained data to an eighteen-year database from June 1995 to March 2013 of Amazon product ratings including 34,686 .770 reviews by 6,643,669 users for 2,441,053 products.

The following data are available from the database: (i) the identification number (id) of the users registered on the Amazon platform; and (ii) the data provided was made public by its creators as they have been published so that anyone can access them from the Amazon product sales page.

Such data does not allow us to identify persons who have registered under the users' names, and the use of such data has been authorized by Amazon and Stanford University in academic papers, so there is no conflict or privacy

violation with the use of the database.

Data were made available following the storage structure below:

- product/product Id: asin, e.g. amazon.com/dp/B00006HAXW
- product/title: title of the product
- product/price: price of the product
- Review/ userId: id of the user, eg A1RSDE90N6RSZF
- review / profileName: name of the user
- review/helpfulness: fraction of users who found the review helpful
- review/score: rating of the product
- review/time: time of the review (Unix time)
- review / summary: review summary
- review/text: text of the review

Once the available database presents user interaction by evaluating products and documenting their positive/negative critiques of these, we are now focusing our focus on crowdsourcing for crowd voting, Howe's typology (2008), which includes the evaluation of scores awarded to products by users of the Amazon sales platform.

On this basis, to allow manipulation of the database, due to processing limitations, it was chosen to extract a sample of small size, but that was complete enough to reflect the reality of the original database. The final example is made up of the watch products segment, which initially comprised 68 356 ratings. From this sample, we chose to work with the last four years of available data (2010 to 2013) and to present the completed score field.

This final sample resulted in 1 750 evaluations, which were organized into a segregated data file so that it could have its data characterized and imported into statistical-oriented programs.

2.3 Pearson's Correlation Test

It was decided to use two software, Microsoft Excel for data manipulation and data formatting for import in statistical software and Eviews version 7, to carry out the statistical studies of the correlation between observed variables.

To test the proposed hypothesis, we selected five variables, these being:

- Product Score: Product rating on a scale from 1 to 5 provided by the user.
- Polarity of review: From the score assigned to the product by the reviewer was defined as grades 1 to 3 as negative ratings and grades 3 to 5 as positive evaluations.
- Number of characters in the review: the number of characters in the review, already discounted the spaces.
- Review help rating: Fraction of users who will understand the review helpful.
- Product price: Advertised selling price for the evaluated product.

For these variables, the correlation of the variables with each other was analyzed using Pearson's linear correlation test. This test evaluates the correlation by measuring the value of the X and Y variables at the interval or ratio level. From this analysis, the test provides the Pearson - r Coefficient - which can vary from -1 to +1, and the closer to these values, the stronger the association of the variables under analysis. The zero scores of this coefficient indicate no correlation between the variables. In this way, we perform the following analyzes:

1- Analyze whether products with higher scores (REVIEW_SCORE) or lower have a more significant number of characters in the review (N °CHAR), which would suggest a greater dedication of the users to the products that cause them some frustration.

2- Correlation evaluation between the number of characters in the review (N °CHAR) and the assessment of help provided by the review (REVIEW_HELPFULNESS_NUMBER) to understand the behavior of the mass of users against reviews of short and extensive lengths.

3- Correlation of the polarity of the review, if positive, neutral or negative (POLARITY), with the score of the product (REVIEW_SCORE).

4- Correlation of the number of characters (N °CHAR) and polarity (POLARITY) of the review.

5- Evaluation of the correlation of the review help (REVIEW_HELP_NUMBER) with the final score of the product.

6- Evaluation of the correlation of the product score (review_score) with the price of the product (PRODUCT_PRICE).

7- Evaluation of the correlation of the number of characters of the review (N °CHAR) with the price of the product (PRODUCT_PRICE).

These analyzes initially approach the entire sample worked in this dissertation, but for a second analysis with the objective of complementing the vision of the behavior of the user mass that have their time to evaluate the products, we divide the sample according to the perception that the user had the product, its polarity, positive or negative, in order to compare the variables against well evaluated and poorly evaluated products. The purpose of this separation is to test the guiding hypothesis of this article, that the interaction of the users and their evaluations end up impacting the performance of the products evaluated by them.

2.4 Hypothesis Test

In the research strategy adopted, we tried to formulate three hypotheses that could act as guidelines of the work, aiming to emphasize the importance of crowdsourcing, with focus on the crowd voting, based on the typology of Howe (2008), firstly to use search engines for the purpose of to obtain various articles on the topic crowdsourcing and correlated topics such as knowledge trees, collective intelligence, and from the volume found and the slope of the journals, evaluate the authors' / researchers' perception of the benefits and harms of crowdsourcing, and their impacts in the sales strategies of the companies, since in order to obtain a better focus on the research we limit the influence of crowdsourcing on the products offered by the companies.

Consequently, after extensive documentary research on crowdsourcing and its underlying concepts, we sought to build the theoretical panorama of crowdsourcing and the applications of crowd voting so that we could find foundations that objectively support this study and allow to form a paradigmatic vision strong enough to suggest new paths of research and a valid guidance for the quantitative deepening of the theme.

From this survey, we try to respond with the current maturity of crowdsourcing, if there are indications strong enough to suggest the application of the crowdsourcing strategy, focusing on the performance of the product according to the influence of crowd voting, Howe (2008).

Confronting then the hypotheses originally constituted against the data obtained from the correlation of the chosen variables in the methodology and in the extraction of content that can be used by the companies to map a consumer profile of their products.

3. Results

Through the database granted by Stanford University about the operations of the last 4 years of watch sales through the Amazon site, Person's correlations were analyzed among the variables highlighted in the sample, in order to study the influence of variables, and to map the impacts of crowd voting on product performance, since the score obtained by the products, through the evaluation of the users, influences their disposition in the search results and the directed advertisements of Amazon to the users of the electronic commerce.

For this analysis, to understand the behavior of the user mass that evaluates the products, we performed the correlations in the sample of 1 750 Amazon Clock Assessment Records, and later we verified the correlation between the variables from a filter based in the polarity of the comment, whether positive or negative.

3.1 Full Sample

Pearson's correlation between the number of characters used in the review and the score obtained in the user's evaluation:

Table 1. Pearson's correlation N °char x review score

	N °CHAR	REVIEW_SCORE
N °CHAR	1.000000	-0.060642
REVIEW_SCORE	-0.060642	1.000000

Source: Own elaboration.

The correlation test indicates, through its negative r, that there is a low and negative correlation between these variables, r of -0.06. Below is the detailed table with the information resulting from the test.

Pearson's correlation between the number of characters used in the review and the utility rating score provider's evaluation:

Table 2. Pearson's correlation N °char x review helpfulness

	N °CHAR	REVIEW_HELP_NUMBE
N °CHAR	1.000000	0.269291
REVIEW_HELP_NUMBE	0.269291	1.000000

Source: Own elaboration.

A positive correlation of 26.9% between the size of the comments provided and the evaluation of its usefulness is observed, suggesting that once the consumer chooses to use comments made on a product, the amount of time spent collaborates for a perception of its value in consumer assessment.

Pearson's correlation between the score obtained by the product and the polarity of the evaluation:

Table 3. Pearson's correlation polarity x review score

	POLARITY	REVIEW_SCORE
POLARITY	1.000000	0.936781
REVIEW_SCORE	0.936781	1.000000

Source: Own elaboration.

In this case, a high correlation is observed, presenting a positive correlation of 93.67%. This indicates that the product scores accompany the user's perception of the product, also influenced by the fact that the polarity of the commentary is presented as a variable dependent on the score variable.

Pearson's correlation between the positive, neutral or negative perception of the product, polarity, against the number of characters of the evaluations:

Table 4. Pearson's correlation N °char x polarity

	N °CHAR	POLARITY
N °CHAR	1.000000	-0.057724
POLARITY	-0.057724	1.000000

Source: Own elaboration.

In this evaluation, we observed a negative but low correlation, which indicates that users tend to write more about products that caused them a negative perception, yet with a correlation of only 5.77%.

Pearson's correlation test between the utility of the Review and the Score of the product:

Table 5. Pearson's correlation review helpfulness x review score

	REVIEW_HELP_NUMBE	REVIEW_SCORE
REVIEW_HELP_NUMBE	1.000000	-0.056979
REVIEW_SCORE	-0.056979	1.000000

Source: Own elaboration.

The correlation between these variables showed a negative and low correlation, of 5.69% among the observed variables. This suggests that the evaluations of the comments provided are more focused on products with lower scores, still showing little correlation with each other.

Pearson's correlation between product price and product score:

Table 6. Pearson's correlation - product price x review score

	PRODUCT_PRICE	REVIEW_SCORE
PRODUCT_PRICE	1.000000	0.040435
REVIEW_SCORE	0.040435	1.000000

Source: Own elaboration.

As a result of this evaluation, a positive but low correlation of 4% between the price of the product and the review obtained by it is observed. Pointing out that the value of the product presents a low correlation with the evaluation, but positive, showing that the price influences the perception of the consumer in the value of the product.

Pearson's correlation between the price of the product and the size of the comments provided:

Table 7. Pearson's correlation - product price x N °char

	PRODUCT_PRICE	NUMERO_DE_CARACTERES
POLARIDADE	1.000000	0.0157761
REVIEW_SCORE	0.157761	1.000000

Source: Own elaboration.

The test indicates a positive correlation of 15%, noting that the higher the price of a product on analysis, the greater the time spent by the user in the evaluation of the same, indicating that the price is a factor of commitment of the consumer against the item to be acquired.

Analyzing the correlations tests, it is observed that the information, although composing the scenario of the evaluation of the product, alone do not present great correlations with the score of the product. The correlation with the score of the product exhibited a value close to 5%, which even though it shows a degree of influence, is still low, suggesting that the weight of the evaluation introduces a personal character, varying from user to user, which totalizes this information and according to their perception of the product acquired expresses their understanding.

To broaden this perception and seek a more unobstructed view of crowd voting behavior to the mass of data, we also compare the variables according to the users' positive or negative perception of the product. In this division we obtained the following correlations:

3.2 Sample With Negative Polarity (Ratings 1 to 2)

Tabela 8. Pearson's correlation – sample with negative polarity

	N °CHAR	PRODUCT_PRICE	REVIEW_HELP_NUMBER	REVIEW_SCORE
N °CHAR	1.000000	0.112126	0.140710	0.166989
PRODUCT_PRICE	0.112126	1.000000	-0.029278	-0.012072
REVIEW_HELP_NUMBER	0.140710	-0.029278	1.000000	0.026750
REVIEW_SCORE	0.1669890	-0.012072	0.026750	1.000000

Source: Own elaboration.

By separating the sample according to the product perception, a different view of the sample can be obtained,

without the nullities obtained from the comparison of the positive and negative samples.

There was a considerable increase in the correlation between the number of characters in the Review (N °CHAR) with the grade obtained in the evaluation (REVIEW_SCORE); in the sample without the filter we had a positive correlation of 6.06% between these variables, and after the screen, we obtained a positive correlation of 16.69%. Noting that in poorly rated products, users tend to focus more on reviews made by other users than to make their own decisions about what product is being offered. This trend corroborates the initial hypothesis that crowdsourcing affects product performance.

When analyzing the correlation between the number of characters of the review (N °CHAR) and the usefulness of the evaluation made available (REVIEW_HELP_NUMBER), with the sample without filters, we obtained a positive correlation of 26.92%, whereas in the sample focused on products with negative perception, we got a positive correlation of 14.07%. That indicates the tendency of a comment to be considered useful according to the time dedicated in its formulation, although in a smaller number.

As for the correlation test between the number of characters of the review (N °CHAR) and the price of the product (PRODUCT_PRICE), in the sample without filters we had obtained a positive correlation of 15.77%, while in the sample focused on products with a low score, we got a positive relationship of 11.21%. This value reflects the tendency of the price to influence the dedication of the user when elaborating the comment about the product.

Pearson's correlation between the product score (REVIEW_SCORE) and the usefulness of the review (REVIEW_HELP_NUMBER) originally had a positive relationship of 5.69%, and now reflecting this trend, but with a slight reduction, we obtained a positive correlation of 2.67%.

In the evaluation of the correlation between the product score (REVIEW_SCORE) and the product price (PRODUCT_PRICE), we had initially obtained a positive relationship of 4.04%, showing the tendency, although low, of the product price to accompany the evaluation of the same , when the filter was run on the sample, a negative correlation of 1.2% was observed, which suggests that when a product does not satisfy the user's perspective, the movement is reversed, and in this way, when presenting a higher price ends up receive an even worse rating.

3.3 Sample With Neutral Polarity (Ratings Equal to 3)

Next, we chose to filter the sample from products that obtained only a score equal to 3, understanding that the average grade can be considered a neutral evaluation, and from this sample to analyze the behavior of the individuals that compose and the population participating of crowd voting on the products marketed by Amazon company for the period from 2010 to 2013 in the watch segment.

Table 9. Pearson's correlation – sample with neutral polarity

	N °CHAR	PRODUCT_PRICE	REVIEW_HELP_NUMBER
N °CHAR	1.000000	0.207689	0.368453
PRODUCT_PRICE	0.207689	1.000000	0.070264
REVIEW_HELP_NUMBER	0.368453	0.070264	1.000000

Source: Own elaboration.

It is observed that the correlations obtained follow the trend of the relationships acquired in the general sample, but with a considerable increase in the degree of relationship between the variable number of characters of the review (N °CHAR) and the price variable (PRODUCT_PRICE). The correlation obtained was 20.76% versus 11.21% of the original sample. This trend suggests that for products that are expected to be medium, the price is a fact that ends up weighing considerably on the choice, as can be observed in the word cloud formed by this sample.

It is also noted that the correlation between the number of characters (N °CHAR) and the review utility variable (REVIEW_HELP_NUMBER) showed a considerable increase compared to the original sample, where we previously had a 14.07% correlation between them, we now observe 36.84 %.

Another point that suggests that for median products, where at first, one can imagine that price exerts the main weight in the choice of product, in a competitive market it is noticed that the individuals endeavor to compare them and express their opinions in order to assist or direct future purchases in a line that they think is more beneficial, seeking the cost benefit. An indication that these comments may affect the demand for this product.

3.4 Sample With Positive Polarity (Ratings From 4 to 5)

Table 10. Pearson's correlation – sample with positive polarity

	N °CHAR	REVIEW_HELP_NUMB	REVIEW_SCORE	PRODUCT_PRICE
N °CHAR	1.000000	0.297720	-0.062909	0.166248
REVIEW_HELP_NUMB	0.297720	1.000000	0.036637	0.092636
REVIEW_SCORE	-0.062909	0.036637	1.000000	0.073424
PRODUCT_PRICE	0.166248	0.092636	0.073424	1.000000

Source: Own elaboration.

By adjusting the filter for positive product perception, positive polarity, user feedback, we obtained a complementary view of the negative polarity view, pointing out distinct behaviors to products well evaluated by users, and consequently, outlining community reactions that may affect the voting results obtained by crowd voting.

The correlation between the number of Review characters (N °CHAR) and the score obtained in the evaluation (REVIEW_SCORE) was reversed. In the sample without the filter, we had got a positive correlation of 6.06% between these variables, whereas, after the screen, we obtained a negative correlation of -6.29%. This contrast indicates that in well-evaluated products, users tend to be more objective in their comments, presenting a reduction in the size of these comments which is inversely proportional to the score obtained by the product.

This point is interesting for pointing out that the second hypothesis used as the guideline of this study (H2 - Negative comments originated by crowd voting have the same effects as positive comments), appears to be flawed since negative perceptions tend to attract more dedication comments.

As for the correlation between the number of characters of the review (N °CHAR) and the usefulness of the evaluation made available (REVIEW_HELP_NUMBER), with the sample without filters, we obtained a positive correlation of 26.92%. On the other hand, in the sample with the positive perception products filter, we observed a positive association of 29.77%. These results indicate the tendency of a comment to be considered useful according to the time spent in its formulation, following the trend of the original sample.

When performing the Pearson correlation test between the number of characters of the review and the price of the product (PRODUCT_PRICE), for the original sample, without filters, we had obtained a positive correlation of 15.77%, while in the sample with a focus on products with a low score, we got a positive correlation of 16.62%. These values reflect the price trend was influencing the dedication of the user when elaborating the comment about the product.

Pearson's correlation between the score of the product (REVIEW_SCORE) and the usefulness of the review (REVIEW_HELP_NUMBER) originally had a positive correlation of 5.69%, and the sample with positive filter reflects this tendency, but with a significant increase, we obtained a positive correlation of 36.63%.

It is observed an unusual behavior; the reviews elaborated by other users are used as a subsidy in the elaboration of positive comments. It should be noted that the correlation between the number of characters of the evaluation decreases as the grade obtained by the product goes up. Taken together, these data suggest that there is a tendency for products with an excellent score to be generally accepted by the community to receive positive comments from new users, even though the comments left are more succinct.

In the evaluation of the correlation between the product score (REVIEW_SCORE) and the product price (PRODUCT_PRICE), we had initially obtained a positive correlation of 4.04%, which indicates the tendency, although low, of the price of the product to accompany the evaluation the same. To the application of the filter in the sample, a positive correlation of 7.34% was observed, which accompanies the first trend.

To visualize the data more clearly, the table below summarizes the comparisons between the original sample, and after the positive and negative perception filters of the products:

3.5 Comparative Summary Between the Original Sample and Polarized Samples

Table 11. Pearson's correlation – comparativo de amostras

Pearson's Correlation	Original Sample	Sample Polarity	– Positive	Sample Polarity	– Negative
N°CHAR X REVIEW_SCORE	6,06%	-6,29%		16,69%	
N°CHAR X REVIEW_HELP	26,92%	29,77%		14,07%	
N°CHAR X PRODUCT PRICE	15,77%	16,62%		11,21%	
REVIEW_SCORE	X	5,69%	36,63%		5,69%
REVIEW_HELP_NUMB					
REVIEW_SCORE X PRODUCT_PRICE	4,04%	7,34%		4,04	

Source: Own elaboration.

4. Discussion

After comparing the correlation between the sample data from Amazon's watch database and reading the available academic production on the subject, we obtained indications that suggest a response to the tests of the hypotheses H1 and H2 initially proposed. These being:

H1 - Crowdsourcing affects product performance.

For a company that uses as its primary means of delivering its products in the web environment, and that provides the tool for users to evaluate their products of interest and contribute their comments to the community, the correlations, although low, indicating that other market participants consult the inferences made products and converge for the evaluation of a product score.

This score influences how Amazon presents the information on your e-commerce site and presents itself as an independent opinion from that shown by the seller who feeds the buyer's decision-making on the site. In this way, the data and the way the tool is offered suggest, in the test of this hypothesis, that the assertion is true, and in this way, it is correct to affirm that crowdsourcing affects the performance of the linked products, at least with respect to the web environment in a company with the profile of Amazon, that is, an e-commerce site that provides the product evaluation tool and uses this data in the layout of the products when offering the customer.

H2 - Negative feedback from crowd voting has the same effects as positive feedback.

In the original sample, without the application of polarity filters of the reviews, this answer was not clear, not presenting a direction that could support the test of the same one. However, after dividing the original sample into well-evaluated products and poorly evaluated products, positive and negative polarity, it was possible to verify that negative evaluations tend to motivate users to devote more effort to their comments on the products. This is because the comparison of the Pearson correlation indices between number_characters and review_score on the negative polarity filter is 16.69% versus -6.06% of the positive polarity filter sample.

H3- Crowd voting provides data on what is being perceived by the consumer market versus the advertised product.

In the evaluation of the available data in the database of product reviews made available by Amazon, through the University of Standford, a series of information about the individual perceptions of each consumer that manifested itself against the possibility of evaluating a product offered by the platform of the company's sales.

This information presented a considerable challenge to be evaluated since relevant information is masked by a series of data that is lost in diffuse expressions, regional expressions, slang, and the need for computational processing to evaluate extensive databases.

Nevertheless, through the statistical analysis of the work, demonstrated by Pearson's correlation on the usefulness of comments compared to the number of characters in each comment, we have results that suggest that crowd voting does provide data on what is being perceived by the consumer market to announced products. These data provide information that can collaborate in sales strategies and marketing negotiation with this consumer market

After testing the hypotheses formulated to direct this dissertation, in front of the available literature on the subject and the tests carried out on the database analyzed, with 1 750 different user evaluations for products of the watch

segment offered between 2010 and 2013 by the trading platform Amazon electronic company, these hypotheses have shown to be coherent and thus suggesting that these are true. Notwithstanding this, a thorough understanding of the subject calls for further research on the topic, deepening the analysis and broadening the survey for more massive databases.

The study carried out limited the processing capacity of the database, since the initial base contained eighteen years of data, covering a period from June 1995 to March 2013, including the Amazon product evaluations, including Customer Reviews (0). There are no reviews. This base, although inviting, proved to be complicated to be manipulated without a data processing laboratory, requiring time and resources not available during the present work. This limitation resulted in the segregation of records relating to clocks, which for the 18 years of registration had 68 356 reviews.

In order to further reduce the size of the data mass, we have chosen to limit the analysis between the years 2010 to 2013 in order to provide a mass of data large enough to portray a large data data from external users of the corporation, formatting a scenario conducive to the analysis of crowdsourcing through crowd voting. From the obtained data, we discard incomplete data, resulting in a data mass of 1 750 reviews to be analyzed.

In this way, the results of this article point out that, for a company with operations based on the web environment, focused on e-commerce products and that provides a platform for the interaction of users among themselves, as is the case of the Amazon company, As a result of the influence of crowdsourcing, with a focus on Howe's crowd voting typology (2008), the study suggests that the opinion and interaction of the populations between themselves and with the company end up influencing the performance of the products instead.

Although the correlations found by the Pearson tests show a low correlation, the own benchmarking of this database with thousands of reviews, where the comment utility variable (Review_helpfulness), points to growing importance that the consumer; previously passive, from isolated to connected, as argued by Prahalad and Ramaswamy (2004). It is therefore seen that consumers have provided information and can be considered as suitable and disassociated sources to the seller of the product.

5. Conclusion

The crowdsourcing phenomenon is an essential component of strategy among technology companies, an environment in which crowdsourcing, and its various typologies emerge and operate with greater acceptance. This movement is due to the strong dependence of technological tools that provide a platform where the masses can converge and that the generated information can be registered and treated.

Due to the large volume of information generated by the reach of the internet and the globalization of markets, creating movements of a mass that interacts with each other and dialogues with companies and markets, in a formation of a collective consciousness, as explained by Lévy (1998), the most significant challenge is to separate the useful content business strategies from random data masses, a problem that also imposed on the work, acting as part of the limitation imposed on this article. The processing capacity available for the research conducted here, as well as cost and processing time, limited the scope of the database analyzes provided by Amazon.

The motivation to research with greater depth is analyzed, analyzing the comments and working correctly the linguistic barriers that impose, seeking a greater understanding of the participating users.

Nevertheless, despite the limitations, the results obtained by the study suggest that crowdsourcing influences the performance of the products to be marketed, and thus, indirectly, affects the performance of companies. The affirmative answer as to the integrity of the hypotheses proposed as guidelines of this article emphasizes the importance of the theme, which can result in strategic gains for companies that are in the web environment and face a globalized and highly competitive environment.

Thus, in order to contribute to the complement of the theme, the study still suggests that the periods of analysis and the size of the data mass be expanded, so that a temporal line of behavior of the populations can be formed over time, in this way the degree of involvement of the network in front of the intersection of the concept of crowdsourcing, since the database available starts before the theme comes into vogue and to have its platforms modernized so that they could support and handle the new volume of data.

References

- Amabile, T., Barsade, S., Mueller, J., & Staw, B. (2005). Affect and Creativity at Work. *Admin. Sci. Quart.*, 50, 367-403. <https://doi.org/10.2189/asqu.2005.50.3.367>
- Bloem, J., Doorn, M. V., & Ommeren, E. (2007). *Open for business - Open Source Inspired Innovation*. Soggetto/Vitt.

- Bons, E., Daams, M., Neijnens, E., Ottenhey, D., Segeren, M., & Van Der Sommen, G. (2010). Open innovation: The benefits of crowdsourcing. *Working paper*, Tilburg University, Netherlands.
- Brabham, D. C. (2008, June). Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First Monday*, 13(6). <https://doi.org/10.5210/fm.v13i6.2159>
- Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*. <https://doi.org/10.1177/1473095209104824>
- Brabham, D. C. (2010). Crowdsourcing as a model for problem solving leveraging the collective intelligence of online communities for public good. 2010. *PHD Thesis*, Utah University. Retrieved March, 31, 2019 from <http://cdmbuntu.lib.utah.edu/utils/getfile/collection/etd2/id/1190/filename/419.pdf>
- Brabham, D. C. (2010, August). Moving the crowd at threadless. *Information, Communication & Society*, 13(8), 1122-1145. <https://doi.org/10.1080/13691181003624090>
- Campbell, D. J. (1988). Task complexity: A review and analysis. *The Academy of Management Review*, 13(1), 40-52. <https://doi.org/10.5465/amr.1988.4306775>
- Castanheira, M., Peixoto, J., & Fachada, T. (2003). *Conhecimento, reflexidade e desempenho: Anotações para uma discussão*. Rio de Janeiro: Papel Virtual Editora.
- Daft, R. L. (2003). *Organizações: Teorias e Projetos*. São Paulo: Pioneira Thomson Learning.
- Doan, A., Ramakrishnan, R. E., & Halevy, A. Y. (2011). Crowdsourcing systems on the World-Wide Web. *Communications of the ACM*, 54, 86. <https://doi.org/10.1145/1924421.1924442>
- Edvinsson, L., & Malone, M. S. (1998). *Capital Intelectual*. São Paulo: MakronBooks.
- Geiger, D., Rosemann, M., & Fiebt, E. (2011). *Crowdsourcing information systems – A systems theory perspective*. Australasian Conference on Information Systems (ACIS), Sydney, Australia.
- Godoy, A. S. (1995). Pesquisa qualitativa: Tipos fundamentais. *Revista de Administração de Empresas*, 35(3), 20-29. <https://doi.org/10.1590/S0034-75901995000300004>
- Howe, J. O. (2008). *Poder das Multidões*. São Paulo: Elsevier Técnico.
- Levy, P. (1998). *A inteligência Coletiva, por uma antropologia do ciberespaço*. São Paulo: Edições Loyola.
- Nonaka, I., & Takeuchi, H. (1997). *Criação de conhecimento na empresa*. Rio de Janeiro: Campus.
- Noubel, J. F. (2004). *Collective Intelligence: The Invisible Revolution*. Retrieved October 25, 2014, from <http://publishing.yudu.com/Library/Arswi/CollectiveIntelligen/resources/index.htm>
- Page, S. E. (2009). *The Difference - How the Power of diversity creates better groups, firms, schools and societies*. New Jersey: Princeton University Press. <https://doi.org/10.1515/9781400830282>
- Prahalad, C. K., & Ramaswamy, V. (2004, June). Co-creation experiences: The next practice in value creation. *Journal of interactive marketing*, 18(3). <https://doi.org/10.1002/dir.20015>
- Reto, L. E., & Nunes, F. (2001). *Normas de elaboração de teses de mestrado*. INDEG/ISCTE, Lisboa.
- Salerno, S. M. (1999). *Projeto de Organizações integradas e flexíveis: Processos, Grupos e gestão democrática via espaços de comunicação - negociação*. São Paulo: Atlas.
- Schenk, E., & Guittard, C. (2009). Crowdsourcing: What can be outsourced to the crowd, and why?. *Working paper of Graduate School of Science and Technology*. Retrieved Setember 2, 2014, from <https://halshs.archives-ouvertes.fr/halshs-00439256v1>
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics*, 7(1), 93-107. <https://doi.org/10.3917/jie.007.0093>
- Slawsby, A., & Rivera, C. (2007). Collective Innovation. Massachusetts Institute of Technology. Retrieved Setember 2, 2014, from <http://dspace.mit.edu/handle/1721.1/39518>
- Sveiby, K. E. (1998). *A Nova Riqueza das Organizações*. Rio de Janeiro: Campus.
- Wolley, A. W. et al. (2010). Evidence for a Collective Intelligence Factor in the Performance of Human Groups. *Science*, 330(6004), 686-688. <https://doi.org/10.1126/science.1193147>
- Wyatt, J. (2001). Management of Explicit and Tacit Knowledge. *Journal of the Royal Society of Medicine*, 94, 6-9. <https://doi.org/10.1177/014107680109400102>