

PostQual: a dedicated scale for postal perceived service quality in Algeria. Methodology and reliability results.

Bouabdallah Salah¹

¹ University Mohamed Boudiaf, M'sila, Faculty of economics, commerce and management, Laboratory of Economic Strategies and Policies for Algeria (LSPEA), Algeria, salah.bouabdallah@univ-msila.dz

Received: 06/11/2021

Accepted: 11/12/2021

Published: 31/12/2021

Abstract:

Algerian mail company is the most important public service provider in terms of the number of customers, the amount of funds detained, and the importance of its sales force. This paper reports the qualitative and quantitative steps followed to develop a dedicated scale to understand and measure how customers perceive postal service quality as well as the reliability results. The questionnaire has been administered to a large sample (n=557) in 17 post offices dispatched over five districts of northeast Algeria. The main result is a three-dimensional scale that contains 24 items. Those dimensions are: reception, outcome, and complaints handling. The paper discusses the results in light of the literature and gives some insights about managerial implications.

Keywords: Algerian mail company (Algérie Poste), perceived service quality (PSQ). Factor Analysis, Principal Components Analysis.

Jel Classification Codes: M21, L87, C38

1. INTRODUCTION

With more than 3000 post offices, visited by nearly a million citizens per day, “Algérie Poste” is the leading company in the postal industry and the number-one public service provider in Algeria (Algérie Poste, 2013). Algerian mail company is also number one in the country in terms of the

¹ **Corresponding author**

amount of money it holds and manages. Understanding how the customer evaluates the service of this company is crucial for the citizens, the company, and the national economy. In order for the company to be able to detect and address the right priorities of quality, the company has engaged a quality policy (Normalisation and quality department., 2017). This policy has to be informed and based on data facts about what matters most to the customer. This implicates measuring the customer perception of service quality thoroughly and on a regular basis, using a scientifically developed measurement tool. To our knowledge, no such dedicated scale has been presented in the Algerian postal context. This paper addresses this gap by suggesting a measurement tool designed for the postal service quality (SQ), specifically in the Algerian context.

In the postal industry, it has been proven that SQ has a strong influence on customer satisfaction (Siti Rapidah, Ali, & Siti Naemah, 2020). SQ is a key managerial concept, it has been involved in almost all models that researchers made to give a comprehensive view on customer's attitude and its influence on the firm's performance (Brady, et al., 2005) and (Furrer & Sollberger, 2007). The issue of measuring SQ has long been one of the main topics that captured the interest of marketing and management researchers, leading to the development of several universal/global measures; particularly, Servqual and Servperf. However, the weaknesses of these two universal measures highlight the need to develop context-specific and industry-specific SQ measures (Karatepe, Babakus, & Yavas, 2005). Some empirical studies revealed significant differences in the way SQ is evaluated according to the dominant culture, even within the same continent or a country, as reported by Ueltschy (2004) and (2009). For these reasons, many researchers have questioned the generalizability of the universal measures over countries and cultures, and claim the need to take into account the impact of the local culture on the manner the customer evaluates the service and expresses his feeling about it (Carrillat, Jaramillo, & Mulki, 2007).

This study aims to develop a dedicated scale that measures postal SQ in the Algerian context. We first draw on the theoretical background of the research, making a broad review of the relevant literature about measuring

SQ. We then relate the research methodology and procedures. Lastly, we report the reliability results and discuss the scope and limits of the dimensionality established as well as their scientific and managerial implications.

2. THEORETICAL BACKGROUND

2.1 Construct definition and operationalization

Service quality is defined as “the customer’s assessment of the overall excellence or superiority of the service” (Zeithaml, 1988). In the last three decades, a huge research effort has been made to define and delimit this indistinct construct. Once the bases of this conceptualization have been established, researchers have moved towards operationalizing perceived service quality (PSQ) by developing generic and dedicated scales to measure it. Two schools of thought have emerged and shaped the literature about PSQ; the North European school (Norwegian), and the American one. The European school of thought suggests that PSQ has three dimensions: functional quality, technical quality and corporate image (Grönroos, 1984); the functional, and the technical quality cover respectively the response of how the service is provided, and what is actually being delivered; while corporate image relates to the insinuation customers get from the brand, the company, and/or the service. In the same line of thought, Lehtinen and Lehtinen (1982) suggest a close dimensionality: interactive quality, physical quality, and corporate quality.

The work of this school of thought has been later overshadowed by the influence of the American school led by three American authors: Parasuraman, Zeithaml, and Berry (referred to by the abbreviation PZB). Those authors’ major contribution is their 22-item scale, called Servqual, based on their famous five dimensions: reliability, assurance, tangibles, empathy, and responsiveness. Parasuraman and al. (1985) suggest that this dimensionality, called RATER, is stable and universal, meaning that the measure built on it can be applied with little rewording to almost all service industries, regardless of the context.

2.2 Servqual vs Servperf

RATER dimensionality is also used by another famous measure: Servperf; but while Servqual measures the gap between perceived and expected quality, Cronin and Taylor's scale: Servperf, uses the same items to measure only perceived performance (Cronin & Taylor, 1992). Both measures have been applied and tested in a wide range of settings and compared on different aspects. Mostly, Servqual is set to provide more managerial insights: by comparing perceived performance to consumer expectations, it is possible to know for example, whether a client's disappointment about the service is due to lower service performance or rising expectations. The use of the gap scores however has received a lot of criticisms concerning both theoretic and practical concerns (Buttle, 1996); for instance, expectations are reported to be little variable due to a social tendency to show high expectations. The expectations term is by itself ambiguous, especially when there is no available alternative, as it is the case for some postal services that Algerie Poste hold the monopoly (mainly ordinary mail and Current Postal Account). According to the critics, performance-only-measure is more practical and provides sounder psychometric indices (Cronin & Taylor, 1992).

2.3 Universal vs dedicated scales

Despite a lot of criticism, the RATER's dimensionality through Servqual and Servperf has received huge success as a practical tool. The idea of the possibility of measuring all services with the same tool is quite appealing; it makes possible a huge economy of effort needed to develop new industry-specific measures each time we want to measure the perceived quality of a particular service. Using a common measure also makes it possible to aggregate resulting data and allows comparability between services, which is very useful for nationwide analysis. Lastly, the success of universal measures is also due to the good results in terms of reliability and validity in a good number of studies. Yet, the generalizability of RATER dimensionality is questionable because of the wide variety of the service sector industries (Babakus & Gregory, 1992), (Myerscough, 2002), (Jungki, 2005). Service industries are diverse in terms of the degree of their reliance on technology, the importance of personnel contact, and the extent

of customer involvement in the service process (Lovelok & al., 2008). Service industries diverge also in terms of their degrees of evaluation complexity; for some services, the evaluation is rather obvious and immediate (like hospitality service) while for others, the assessment can only be made after a certain period of time (such as insurance). The type of customers also differs from one service industry to another; some subsectors serve common people while others have a specific type of clients (eg. universities and schools).

These discrepancies suggest that the idea of a common dimensionality should not be accepted without caution. The literature demonstrated that the dimensionality of SQ is not stable; the researchers found different numbers of dimensions in different sectors. For example, Karatepe et al. (2005) in the banking sector found four dimensions, while in the healthcare sector, Carman (1990) found nine dimensions, in the tourism setting, Saleh and Rayan (1992) found five dimensions.

The generalizability of RATER across countries has also been challenged (Mattila, 1999), (Zisis & al., 2009). Originally developed in the United States, this dimensionality may not fit different cultural and economic contexts. Brady et al. (2005) mentioned that the national character, the type of culture, and the economic status are all variables that may influence service evaluation and should be investigated; he particularly admits that “consumers in developed countries may view the service differently from those in developing countries”. It is now established that universal customer is a myth (Horn & Shy, 1996). Culture, as the set of beliefs, values, and habits that characterize a society does influence responses on satisfaction and SQ; what is considered as good quality in one nation maybe not enough or just not important in another one (Ueltschy & al., 2009).

Because of these differences, rewording the items of the universal measures has often proven to be insufficient to make the measure capture the reality on the ground. In many cases, the researchers were obliged to drastically modify the global scale by removing dimensions and/or adding

new ones. Even with considerable modifications, the universal dimensions RATER may give poor fitness for the specificities of the service measured or the context in which it is provided. For these reasons, we assume that the dimensionality of the perceived quality of the services delivered by Algerian Mail Company is industry-specific and has to be discovered through an appropriate methodology.

3. METHODOLOGY AND PROCEDURES

3.1 Overview of the methodology

In developing PostQual scale, we have been largely inspired by the famous Churchill methodology (1979) which is widely used in developing marketing and management measures. It contains eight interactive steps with the possibility of moving back to former stages according to the results of reliability and validity tests (Figure 1).

The whole methodology process can be divided into two major stages; exploratory study and confirmatory study. The purpose of the exploratory study is to define the construct's span, formulate an extended scale, collect data and refine and downsize the scale according to its reliability (Roussel & Wâcheux, 2003). In the confirmatory stage, the refined scale is introduced to a larger sample. The collected data allow the researcher to validate the parsimonious scale developed through several metrics. In our research we followed the steps described below.

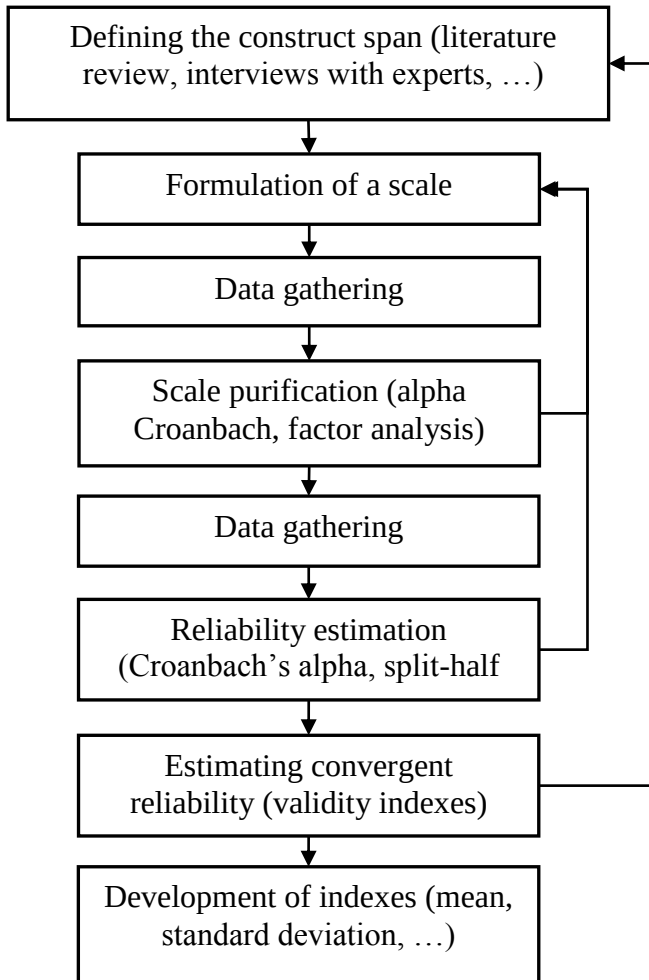
3.2 Defining the construct span

In this stage, we first made an extensive review of the available literature about PSQ, focusing more on those studies conducted in developing countries context. To have a deep understanding of the context of our case, we benefited from domestic newspapers' reports about quality problems encountered by the clients of the Algerian mail company. The company has also provided us with dozens of its official monthly magazine, baptized "Le facteur" (the postman), that informs the general public about the company's projects and policies.

3.3 Formulation of a scale and 1st data gathering

A primary extended scale has been built on the insights of the former work. The aim was to experiment RATER dimensionality through an adapted Sevrperf scale (Cronin & Taylor, 1992). Compared to the gap measure (Servqual), the performance-only measure, (Servperf), has proven to be more practical and has a greater explained variance. We formulated an extended 28-item scale based on the insights of consultations with ten professionals; three of them were management staff and seven were frontline workers. To refine this extended scale, we administered it to a primary sample of 20 participants in a post office. Following the insights of this pilot study, we deleted some items and changed the wording of others towards more clearness; we also decided to avoid negative formulation of items to prevent ambiguity as suggested by Kasper (1999) and Brian (2000).

Figure 1: The research steps



Source: adapted from: Roussel Patrice et Wacheux Frédéric, (2003). *Méthodes de recherche en sciences humaines et sociales, Management des Ressources Humaines*, UQAM – Education, France.

A refined questionnaire, containing 24 items, emerged from this step. This refined scale has then been administered to a convenient sample of 223 participants in three major post offices in M'sila, among them 214 have been exploitable. The majority of the respondents were male (62%), young (56% were between 18 and 30; and 31% were between 31 and 55), and had attended high school or university (63%).

3.4 Reliability estimation and convergent and discriminant validity

The analysis of the results demonstrated a good level of reliability for the five RATER dimensions (Cronbach alpha for each dimension above 0.8). The factor analysis (extraction method: principal component analysis) showed acceptable convergent validity but a poor discriminant one; in general, items loaded well on their intended dimensions but loaded on other dimensions too. As a result, and in line with Churchill's recommendation, we decided to reject the measure and went back to the first stage.

3.5 Redefining the construct area

we first remade a broader reexamination of the literature, especially concerning postal service in the developing word. Second, we conducted seven in-depth interviews with executives from different levels of the company's management: two Regional Directors; one manager of a Mail-shipping Center and four Postmasters. The discussions took about one hour each, and the debate included the following issues:

- What are the quality attributes you think to be critical to quality from customers' standpoint?
- What are the problems involved in providing SQ?
- What steps the company is taking to face those problems and to control and improve SQ?

3.6 Rebuilding a new scale

The customers concerns emerged from the content analysis of these interviews fall into three dimensions: **reception (or functional quality), outcome (or technical quality), and complaints treatment**. Based on the outcome of the interviews, we rebuilt an extended scale that contains 32 items; 18 for the 1st dimension, 10 for the second, and 4 for the 3rd one.

We presented the questionnaire for a consultation to three academicians (assistant professors in quality and marketing). Their feedback contained observations and suggestions about wording, possible ambiguity and survey form. We incorporated their contribution in the final survey.

3.7 Second data gathering

The extended scale had to be refined and downsized; therefore, we administered it, as a pilot study, to a convenient sample inside two major post offices (in Setif and Al Elma cities, $N = 140$). The questionnaires with similar responses or more than five missing data were omitted. The total number of valid surveys collected was 133, which we considered sufficient as a primary study aiming to purify the scale based on its reliability (Karatepe, Babakus, & Yavas, 2005). The survey was administered in Arabic, but French copies were also made available, as some of the customers are more at ease with the French language. The participants were asked to give their opinion about their whole experience with the company rather than one service or one post-office. We used a seven-point scale, with the neutral point in the middle (1 = very bad, 2 = bad, 3 = weak, 4 = do not know, 5 = acceptable, 6 = good, and 7 = excellent). Ascendant and descendant scaled questionnaires were alternatively administered to ensure that the responses would not be affected by the ordering of the scale points; later on, a statistic test showed no difference between the two alternative ordering ways. We used a neutral formulation of the items; e.g. “The assiduity of the servers”. Using positively-worded items would not be optimal since it would hold the risk of encouraging respondents to give high or low marks. Mixing positive and negatively-worded items won’t be a solution since it has proven to be problematic too; the two sets of items tend to give different means and correlate with different groups of variables (Buttle, 1996). To ensure that the sample adequately represents the typical population of postal services customers, the questionnaire concluded with demographic questions concerning gender, age, education level, and activity.

3.8 Purification of the extended scale through reliability criteria

According to Churchill (1979) reliability analysis is to be performed before testing the scale for convergent and discriminant validity by factor analysis, this is to withdraw items that are not consistent with their initial components so that they do not influence the result of factor analysis by rising artificial components. This is recommended when the targeted construct has got a well-established theoretic framework, which is the case

here; in fact, SQ is among the most studied concepts in marketing and management literature.

In the purification process, we searched for items that showed greater Cronbach's alpha for the scale if deleted or showed weak correlation with the total of the other items, less than 0.2 (Churchill, 1979). We also took into account the number of missing values and their singularity. As a precautionary measure, we considered the mark: 4 (do not know) as a missing value to be sure that only responses based on the experience of the service are accounted for. This precaution measure has augmented the number of missing values, making them essential in the purification process.

Both three components showed good reliability, with Cronbach's alpha exceeding 0.8 (reception: 0.818, outcome: 0.907, complaints handling: 0.922). The correlation between each item and the total of other items of their respective groups met the cutting point of 0.2, and most of them exceeded 0.3 (Churchill, 1979); they were between 0.2 and 0.6 for reception; between 0.4 and 0.76 for outcome; and between 0.76 and 0.89 for complaints handling. The process of purification led to removing a total of 8 items, among them, three (one from reception and 2 from outcome) were dropped because of their high number of missing values (29 to 65 among 133) and five for practical reasons, mainly the need to obtain a more parsimonious scale, rather than reliability issues. We also decided that several items in this dimension needed to be clarified, which we did by adding notes to them.

We also checked for test-retest reliability, by introducing the questionnaire twice to a sample of 20 master students of marketing (their names transcript on the questionnaire) with an interval of one month between the two polls.

3.9 Final data gathering

We administered the purified scale (24 items) to a final sample of 667 participants, in 17 post offices, dispatched over five districts (Wilayas) of the Eastern Territorial Postal Administration, which accepted to participate in the study. The number of questionnaires distributed in each district is

proportional to its urban population. Data gathering was in different periods of the working day and the month. The whole gathering data operation was accomplished over five months. 557 usable questionnaires were returned.

4. RESULTS

4.1 Sample characteristics

For representation purposes, we made the number of questionnaires administered in each Wilaya practically proportional to its urban population: 32.09% in Setif, 28.36 in Msila, 15.82 in Bourdj Bouarriridj, 13.13 in Mila, and 10.60% in Jijel. We also tried to represent different categories of the clients of the company in terms of gender (60% men and 40% women); age (30% were under the age of 25, 49% were between 25 and 35, 21.1 were over 36 years old) and activity (58% civil servant, and 25% students, the rest are retirees, free profession or jobless), as well as education (61% of the respondents had some university studies, 25% had secondary school level and 9.3% had mean school level).

4.2 Test-retest reliability

The whole scale showed excellent results in reliability tests. For test-retest reliability, that checks for consistency of the responses against measurement error, we introduced the questionnaire twice to a sample of 20 students with a sufficient interval of four weeks between the two polls. A T-test for paired samples showed a good and significant correlation between the totals of the two samples ($r = 0.5$, $t = 2.43$, $df = 18$, $p = 0.013$). We also tested the score difference between the two paired-samples means (79.15 and 80.05) and found no significance difference ($t = 0.39$, $df = 19$, $p = 0.698$).

4.3 Cronbach's alpha reliability

Cronbach's alpha is a famous measure of consistency of scales, it compares the covariance of the scale items to their total variance.

Table 1. Reliability of the subscales

Subscale	code	item	corrected item-to-total correlation	Expected alpha Cronbach if item deleted
Reception N= 424(*), $\alpha = 0.883$	R1	Work schedule	0.406	0.882
	R2	Employee assiduity	0.606	0.872
	R3	Service reliability in post offices	0.494	0.878
	R4	The proximity of the post offices	0.386	0.883
	R5	Employees' appearance	0.541	0.876
	R6	Waiting hall	0.591	0.873
	R7	The organization of waiting	0.611	0.872
	R8	Employees' Responsiveness	0.639	0.870
	R9	Security in post offices	0.525	0.877
	R10	Speed of service	0.619	0.872
	R11	The appearance of post offices	0.589	0.873
	R12	Employess' empathy	0.704	0.867
	R13	Equality	0.665	0.869
Outcome N= 102(*), $\alpha = 0.945$	O1	Speed of ordinary mail	0.822	0.936
	O2	Reliability of registred mail	0.813	0.936
	O3	Speed of express mail	0.820	0.936
	O4	Reliability of package service	0.851	0.934
	O5	Rapid delivery of parcels	0.840	0.935
	O6	Speed mandates	0.826	0.935
	O7	Reliability of service mandates	0.810	0.937
	O8	Reliability of ATMs(**)	0.623	0.949
Complaints treatment N= 91(*), $\alpha = 0.918$	C1	Ease of presenting complaints	0.852	0.867
	C2	Speed in treating complaints	0.824	0.891
	C3	Justice of responses to complaints.	0.827	0.888

(*) Number of respondents on all items of the subscale. We introduced before the last subscale (Complaint treatment) a pivot question to take only the marks of those who had real experience complaints.

(**) ATMs: Automatic Teller Machines

Source: SPSS output

The reliability of the reception subscale (13 items) was quite good. Cronbach's alpha measured 0.883, which is well above the cutting point of 0.7 (Nunnally, 1978), indicating very good reliability. The estimated

Cronbach's alpha if item deleted were all below the subscale's alpha, meaning that the items are contributing positively to its homogeneity. Correlations between items scores and the total of the scale are between 0.386 and 0.705; which indicates that no items needed to be removed.

The outcome subscale (8 items) showed also excellent reliability, with Cronbach's alpha of 0.945, and its estimated value did not rise if any item were deleted. Correlations between items scores and the total of the subscale were all strong, above 0.6.

The complaint's handling (3 items) also showed high reliability. Cronbach's alpha measured 0.918, and this value diminished if any item were deleted; meaning a positive contribution to all items in the subscale homogeneity. The correlations between each item and the total of the subscale were all strong, above 0.86. We introduced before this subscale a pivot question to take only the marks of those who had real experience with complaints.

4.4 Factor extraction and rotation

The three-dimensional structure of PSQ has been obtained using PCA (Principal Components Analysis). This method explores the correlation matrix between the items of the scale (24 items) to reveal independent groups of variables that correlate to each other constituting factors. The method enables the selection of a small number of these factors that explain an acceptable percentage of the variance existing in the data. A rotation is then applied to the factors selected to ameliorate their interpretation. In the following, we give a summary of the extraction and the rotation process, and their main output.

We first checked whether the correlation matrix is suitable for factor analysis. As expected, we found that the matrix contains some strong correlations (more than 0.8). Also, the variances of the variables are not too different, so the solution will not be influenced by some too dispersed variables. The KMO (Kaizer-Meyer-Olkin) statistic scored high ($0.852 > 0.7$), meaning an important proportion of covariance that could be explained by indirect components. Finally, Bartlett's test of sphericity is significant ($p < 0.001$); this test against the hypothesis of noncorrelated

variables confirms that factor analysis is suitable for the data. The selection of the number of principal components is made according to Kaiser's criterion and the inflection point of the scree plot. The scree plot shows a clear and unique inflection point after the third component and its value is above the cutting value of 1.

The total variance explained by the three principal components is 56.18%, which exceeds the 50% limit. We then used varimax rotation to ameliorate the interpretation of these retained factors (Table 2).

Table 2. Variance explained before and after varimax rotation.

Components	Eigen value			Principal components after rotation		
	Total	% of variance	% cumulative	Total	% of variance	% cumulative
1	9.71	40.467	40.467	5.338	22.243	22.243
2	2.11	8.779	49.246	4.883	20.347	42.590
3	1.66	6.935	56.181	3.262	13.591	56.181
4	0.983	4.097	60.278			
...	...	...	...			
24	0.069	.2900	100.000			

Source: SPSS output

4.5 Level of perceived service quality by dimension and category

The overall level of quality expressed by the respondent was slightly below the cutting point of 4 out of 7 (3.95), but not significantly ($p = 0.325 > 0.05$). This bad performance is mainly due to the low level of the Reception dimension (Mean = 3.81, $p = 0.000$). This situation was somehow expected regarding the long waiting time in post offices in payment days of civil servants, students, and retirees. In fact, the items about waiting time, wait organization, and hall equipment were the worst scored, especially "waiting time" (2.52 out of 7). As t-test requires normality, we checked this hypothesis using K-S (Kolmogorov-Smirnov) statistic, and the result didn't show a significant difference from normality ($p = 0.185$ for the scale mean and $p = 0.586$ for reception). "ATM: Automatic Teller Machine reliability"

also scored quite low (3.45 out of 7). The best-scored items are “order reliability” and “order rapidity” (4.76 and 4.74 out of 7).

Table 3. T- test of PSQ mean

	Mean	Reference level: 4 out of 7		
		t	df	P (2-tailed)
Mean of the overall scale	3.95	-.984	556	0.325
Reception (Functional Quality)	3.81	-3.711	556	0.000
Outcome (Technical Quality)	4.31	5.271	538	0.000
Complaints treatment	4.028	.182	143	0.856

Source: SPSS output

We also tested the difference between mean levels of PSQ of different demographic categories, using n-way ANOVA and post-hoc tests (Shefé test). The main result is a significant difference (at 0.05 level) between categories of education level; respondents who have university-level expressed a lower mean of PSQ. This is probably due to their experience with the long waiting line. No evidence of difference emerged due to gender, age, activity, or region (urban vs rural region). As homogeneity is required in ANOVA tests, we checked this condition and found no evidence of non-homogeneity (Levene’s test gives $p = 0.577$).

5. DISCUSSION

The PostQual structure diverges from the famous “RATER” dimensionality (Reliability, Tangibles, Empathy, Assurance, and Responsiveness). With its three-dimensional structure, our measure rather confirms the North European school of thought that is based on the How and What paradigm. The items that relate to Reception deal with the question of how the service is provided. This dimension is sometimes called Functional quality or Interaction quality or Service delivery (Rust & Oliver, 1994). The Outcome items answer the question of what is actually delivered, which is called in the literature Technical quality, or sometimes Core service or product-service (Brady & Cronin, 2001). Complaint’s

treatment however is less mentioned in the literature, even though it can emerge as a distinct dimension in many services that deal with the public.

6. CONCLUSION

Measuring perceived service quality (PSQ) is quite challenging; as service is inherently intangible, perishable, and of volatile quality. The literature provides us with some scales that are considered “universal”, but their generalizability to the context of the developing countries is questionable, as well as to specific service industries. This study aimed to develop a dedicated scale to measure PSQ in postal setting in Algeria, following Churchill’s (1979) methodology.

Three components of perceived postal SQ in the Algerian context have emerged from the qualitative then the quantitative study; **Reception (or Functional quality), Outcome (or Technical quality), and Complaints treatment.**

The level of quality expressed by the respondents in the three dimensions is medium to low. The items that scored the lowest are related to Reception, especially those items related to the waiting time in the post offices. The analysis of variance revealed that students and respondents who have university-level gave a lower score to the service, probably because of their experience with the waiting line. Technical quality and Complaint’s treatment are scored slightly above the 4 out of 7, and the best-scored items are related to order reliability and rapidity.

6.1 Managerial implications

The Reception dimension is mainly related to the conditions of waiting in post offices. Following the remarkable growth in demand for financial services, long waiting lines have become frequent in post offices, especially in peak days of payment of salaries and pensions. The company has benefited from the government support to tackle this problem by several measures: expanding its network, including multiplying small post-offices, extending the network of ATMs (Automated Teller Machines), expanding the computer network, and providing alternative ways to check the account balance online and by phone. Nevertheless, the level expressed by the

respondents in this study demonstrates that the success of these policies has been limited. The cultural factor is perhaps to be considered; many people are reluctant to change their habits and tend to get their services the old way. As the problem of waiting remains, the presence of the client in post offices is having great importance; the waiting halls need to be vast, well-equipped, and well-maintained. Problems related to the interaction between the customer and the staff have also been mentioned by the respondents; tellers under the pressure of customers affluence in peak days are less able to show perfect empathy and kindness.

The second dimension revealed by this study is related to the core of the service itself. Problems like ATMs down, post offices lacking liquidity, not issued payment card, mail returned to the expenditure, delayed or open or damaged packages, not-paid orders or paid at the office instead of at home; all are quality issues that affect the service outcome, i.e., what the customer gets from the company. Although Technical quality has been scored better than Functional quality, problems related to mail service risk to be more and more frequent in the future if the low activity in this domain lead to less investment and thus a lack of human and material resources.

The third type of problems is related to complaints handling. Complaints are a common business in both post offices and higher-level management at local. They need specific quality attributes: such as efficient communication, transparency, fairness, and rapidity. A company can be good or bad in dealing with complaints independently from its ability in the other dimensions of the service quality.

PostQual can serve as a diagnostic tool that will allow managers to define areas that are weak and need attention; or that are power points and have to be strengthened. Dividing SQ into independent components as we did in this study allows the manager to focus temporarily on one or more type of problems since the requirements to ameliorate each component differ from the other ones, and it would be difficult to deal with all of them at the same time. Algerie Poste can use this questionnaire in its regular investigations about SQ. Response bias has to be dealt with by not neglecting customers who cannot read and write.

The evaluation process of each of the three dimensions has to be adapted in time, place, subject, and the clients concerned. For example, Functional quality is better measured during the service, while Technical Quality and Complaints handling can be measured during or after the service; also, all clients are concerned with the measurement of Reception and Outcome, but only clients who had to present complaints can be asked to respond to the items of the third dimension.

6.2 Perspectives

The measurement tool suggested in this study has been developed following a well-recognized methodology, and showed a good level of reliability; however, more tests need to be done to fully validate this scale especially convergent, discriminant and criterion validity. The construct of PSQ in Postal industry needs to be investigated in a broader perspective that puts it in a comprehensive model including its neighboring constructs such as satisfaction and loyalty. We also recommend retesting the questionnaire using other means than self-administration in post offices, for example via phone calls or emails.

6. Bibliography List

- *Algérie Poste*. (2013, march 3). Retrieved from Algérie Poste: <http://www.poste.dz>
- Babakus, E., & Gregory, W. B. (1992). An empirical assessment of the SERVQUAL scale. *Journal of Business Research*(24), 253-268.
- Brady, M. A., Gary, K. A., Cronin, J. J., Tomas, G., Hult, M., & Bruce, D. K. (2005). Removing the contextual lens: A multinational, multi-setting comparison of service evaluation models. *Journal of Retailing*, 81(3), 215-230.
- Brady, M. K., & Cronin, J. J. (2001). Some new thoughts on conceptualizing perceived service quality: a hierarchical approach. *Journal of Marketing*, 65, 34-49.
- Brian, T. E., Workman, L., & Singh, M. (2000). Ensuring Service Quality for Campus Career Services Centres: A Modified SERVQUAL Scale. *Journal of Marketing Education*, 22(3), 236-245.

- Buttle, F. (1996). SERVQUAL: review, critique, and research agenda. *European Journal of Marketing*, 30(1), 8-32.
- Carman, J. M. (1990). Consumer perceptions of service quality: an assessment of SERVQUAL dimensions. *Journal of Retailing*(66), 33-55.
- Carrillat, F. A., Jaramillo, F., & Mulki, J. P. (2007). The validity of the SERVQUAL and SERVPERF scales A meta-analytic view of 17 years of research across five continents. *International Journal of Service Industry Management*, 18(5), 472-490.
- Churchill, G. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 1(6), 64-73. Retrieved 8 11, 2012, from <http://www.scribd.com/doc/79117067/Churchill-G-a-1979>
- Cronin, J. J., & Taylor, S. (1992). Measuring service quality: a re-examination and extension. *Journal of Marketing*, 56(7), 55-68.
- Furrer, O., & Sollberger, P. (2007). The Dynamics and Evolution of the Service Marketing Literature: 1993-2003. *Service Business*(1), 93-117.
- Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, 18(4), 36-44.
- Horn, H., & Shy, O. (1996). Bundling and international market segmentation. *Int. Economic Review*(37), 51-69.
- Jungki, L. (2005). Measuring service quality in a medical setting in a developing country: The applicability of SERVQUAL. *Services Marketing Quarterly*, 27(2), 1-14.
- Karatepe, O. M., Babakus, E., & Yavas, U. (2005). Measuring service quality of banks, Scale development and validation. *Journal of Retailing and Consumer Services*(12), 373-383.
- Kasper, H., Van, P. H., & de Vries, W. J. (1999). *Services Marketing Management. An International Perspective*. Chichester: John Wiley and Sons.
- Lehtinen, J., & Lehtinen, O. (1982). Service quality: a study of quality dimensions, Working Paper. Helsinki: Service Management Institute.
- Lovelock, & al. (2008). *Marketing des services* (5th ed.). Paris: Pearson Education.
- Mattila, S. A. (1999). The role of culture in the service evaluation process. *Journal of Service Research*, 1(3), 250-261.

- Myerscough, M. A. (2002). Concerns about Servqual's underlying dimensions. Illinois: Illinois State University, IACIS.
- Normalisation and quality department. (2017). *uploads/quality-policy*. Retrieved November 5, 2021, from <https://www.poste.dz/uploads/quality-policy.pdf>
- Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Parasuraman, A., Zeithaml, V., & Berry, L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*(49), 41-50.
- Roussel, P., & Wâcheux, F. (2003). *Méthodes de recherche en sciences humaines et sociales, Management des ressources humaines*. UQAM – Education, France.
- Rust, R., & Oliver, R. (1994). *Service Quality, New directions in theory and practice*. Thousand Oaks, CA : Sage Publications.
- Saleh, F., & Ryan, C. (1992). Analysing service quality in the hospitality industry using the SERVQUAL model. *Services Industries Journal*, 11(3), 324-343.
- Siti Rapidah, O., Ali, & Siti Naemah, A. H. (2020, Aug). Service Quality and Customer Satisfaction: Experience of Customers' in Postal Service. *Jurnal Intelek*, 15(2), 67-75.
- Ueltschy, L. C., & al. (2004). Cross-cultural Invariance of Measures of Satisfaction and Service Quality. *Journal of Business Research*(57), 901-912.
- Ueltschy, L. C., & al. (2009). Is there really an Asian connection? Professional service quality perceptions and customer satisfaction. *Journal of Business Research*(62), 972–979.
- Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, 52(7), 2–22.
- Zisis, P., & al. (2009). The Application of Performance Measurement in the Service Quality Concept: The Case of a Greek Service Organisation. *Journal of Money, Investment and Banking*(9), 27-41.