# Assessing patient experience and attitude: BSC-PATIENT development, translation, and psychometric evaluation - a cross-sectional study

**Faten Amer** ( amer.faten@etk.pte.hu )
University of Pécs
**Sahar Hammoud**
University of Pécs
**David Onchonga**
University of Pécs
**Abdulsalam Alkaiyat**
An Najah National University
**Abdulnaser Nour**
An Najah National University
**Imre Boncz**
University of Pécs
**Dóra Endrei**
University of Pécs

**Research Article**

**Keywords:** Balanced scorecard, patient engagement, satisfaction, hospital, performance evaluation, quality

**Posted Date:** January 10th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1143609/v2



**Version of Record:** A version of this preprint was published at International Journal of Environmental Research and Public Health on June 10th, 2022. See the published version at https://doi.org/10.3390/ijerph19127149.

## Abstract

# Background

Improving the healthcare sector in Palestinian territories is challenging. Organizations frequently use the balanced scorecard (BSC) for performance evaluation (PE) worldwide. BSC includes an evaluation of five perspectives: financial, customer, internal process, knowledge and innovation, and environmental. Most HealthCare Organizations (HCO) evaluated patient satisfaction in BSC, but none considered engaging patients in the evaluation process. This paper aims to develop an instrument to engage patients in assessing BSC perspectives (BSC-PATIENT) and customize it for Palestinian hospitals.

# Methods

This is a cross-sectional study. Two panels of experts participated in the item generation of BSC-PATIENT. Forward and backward translation processes from English to Arabic and vice versa were performed. Pretesting was performed for 30 patients at one hospital. Then, 1000 patients were recruited at 14 hospitals between January and October 2021. Construct validity was tested through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Additionally, the composite reliability (CR), interitem correlation (IIC), and corrected item total correlation (CITC) were assessed to find redundant and low correlated items.

# Results

The scales had high adequate model fit in EFA and CFA. All items included had loadings higher than 0.50. Although the convergent validity was less than 0.50 for some constructs, their discriminant validity, CR, IIC, and CITC showed satisfactory results except the services experience (SERV EXR) construct, which had a CR of 0.50. The final best fit model in CFA comprised ten constructs with 34 items.

# Conclusion

BSC-PATIENT is the first self-administered questionnaire specially developed to engage patients in evaluating their experiences and attitudes based on the BSC perspectives. The questionnaire scales were carried out following relevant guidelines and regulations. The results showed adequacy in the psychometric properties of this instrument. Three-point Likert scales are recommended to assess patient experiences and attitudes since they have higher response rates. Future research can be conducted using this instrument to evaluate the impact of patient experience on attitudes toward BSC perspectives and compare the differences based on patient and hospital characteristics.


## 1. Introduction

In 1992, Norton and Kaplan proposed the initial design of the balanced scorecard (BSC), which incorporated four perspectives steered by the organizational strategy: financial, customer, internal process, and knowledge and innovation [1]. The environmental/social perspective was added later as the fifth pillar of BSC [2]. Unlike the other performance evaluation (PE) tools, which mainly focus on internal perspective assessment, the BSC is considered a holistic approach for PE [3]. BSC implementation proved to positively impact financial performance in the healthcare sector [4]. BSC also increased the patient satisfaction rate. Additionally, it influenced the HealthCare Workers (HCW) satisfaction rate, but to a lesser extent [4].

The performance of healthcare services is adversely affected by long waiting times, inefficiency, low productivity, burned-out medical staff, and dissatisfactory patients [5]. In addition to these universal challenges, the healthcare system in Palestinian territories has also been slapped by political and economic conflicts. Therefore, it is described to be incoherent and inadequate [6, 7]. The 87 hospitals in Palestinian territories have five major types based on the leadership style: 28 public, 39 nongovernmental organizations (NGOs), 17 private, two military, and one United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) [8]. Military hospitals are not operating yet in West Bank. The bed percentage per leadership style is approximately 59% public, 26% NGO, 14% private, and 1% UNRWA [9]. These hospitals are distributed as seven in eastern Jerusalem, 53 in West Bank, and 30 in Gaza [10]. The geographic separation with the disrupted mobility between these territories, added to the blockade of Gaza strip, the checkpoints in West Bank and Jerusalem, the separate de facto government health systems in Gaza and West Bank, the heavy reliance on external health financing, and the dependence on direct household expenditures imposed further challenges on improving the Palestinian healthcare system [6, 11–13]. The spread of coronavirus-19 (COVID-19) added an extra layer of challenge. A recent study [14] referred to the COVID-19 era in conjunction with political conflict to have a double epidemic effect on Palestinian territories, which eventually impacted the Palestinian health system and HealthCare Organizations (HCO) performance during the pandemic.

Patients represent the hospitals' end customers since they are the healthcare services receivers. Researchers have pointed to the engagement of patients (EoP) importance in the process of health policy planning, evaluation, and delivery improvement [15, 16]. Additionally, patient feedback was proven to positively impact the performance improvement in HCO [14]. Strategies to support EoP include patient needs assessment, communication skills improvement, managing patient conflicts and complaints, maintaining patient confidentiality, patient training, and asking patients to review outputs by assessing their perceptions and experiences [16, 17]. It is not sufficient to perform the PE of HCO based on manager and hospital records only; a focus on EoP among the selection of the key performance indicators (KPIs) at HCO was recommended [15]. However, the lack of patient and family member involvement in the evaluation process of BSC was perceived [18, 19].

The first aim of this research is to develop a comprehensive instrument (BSC-PATIENT) that is able to assess 1. patient experiences in light of BSC perspectives, 2. patient PI regarding BSC perspectives, and 3. patient satisfaction and loyalty attitude. The second aim of this research is to customize the

developed instrument at Palestinian hospitals, translate it into Arabic, and validate it.

## 2. The Conceptual Framework

Experiences and perceptions enable people to act in a particular behavior and develop an image, satisfaction, or loyalty attitudes [20]. Figure (1) represents our conceptual model.

### 2.1 The experience

Experience is defined as an event that was lived through [20]. Patient experiences at HCO are formed upon receiving the healthcare service or treatment. Becoming aware of the events, objects, or relationships utilizing senses or observation results in experience perceptions [20].

### 2.2 The attitudes

Attitudes form directly as a result of experiences. There are three types of attitudes, which are sometimes referred to as ABCs of attitude. First, the affective component is how the object, person, issue, or event makes someone feel. The behavioral component is how attitude influences someone's behavior. The cognitive component is someone's thoughts and beliefs about the subject. An example of attitude is image perception, satisfaction, and loyalty. Such evaluations are often positive or negative, but they can sometimes also be uncertain [21].

#### 2.2.1 Patient satisfaction attitude

Satisfaction is the most commonly used metric by managers to assess customer perceptions [22]. Satisfaction does not always lead to loyalty. However, loyalty often begins with a sense of satisfaction [23]. Studies found that patient satisfaction either played a direct impact on their loyalty attitude or played a moderating variable between service quality and loyalty attitude [24].

#### 2.2.2 Brand preference attitude

Brand preference is the degree to which consumers prefer a specific brand relative to competing alternatives. It is considered an essential component of customer loyalty [22].

#### 2.2.3 Perceived quality (PQ) attitude

Studies have proven that PQ exerts an indirect influence on patient loyalty. A rival hypothesis referred to satisfaction as a mediator between PQ and loyalty [24].

#### 2.2.4 Perceived image (PI) attitude

A hospital PI was defined as the sum of beliefs, ideas, and impressions that a patient holds toward a particular hospital [25]. Patients usually form a PI of a hospital from their own past treatment experiences relative to the PIs of competing hospitals [26]. A positive PI of a bank was found to significantly improve the PQ. Therefore, in healthcare, a positive hospital PI may positively influence PQ. However, a recent review showed that this has not yet been studied [26].

#### 2.2.5 Loyalty attitude

A loyalty attitude is a behavioral intention that reflects faithfulness and allegiance to something [20]. In the marketing management field, Kotler and Keller (2015) defined loyalty as a deeply held commitment to rebuy or repatronize a preferred product or service in the future despite influences to cause switching behavior [27]. A study revealed a need to use multiple indicators to predict customer loyalty behavior, such as customer satisfaction, brand preference against competitors, intention to return or repurchase, and willingness to recommend [22]. Moreover, customer behavior trends in the past were a good predictor of future customer behavior. It is important to emphasize that loyalty refers to customers' actual conduct, regardless of their attitudes or preferences. However, assessing customer loyalty attitudes can help predict their loyalty behavior in the future [28].

##### 2.2.5.1 Repurchase intention attitude

Researchers have used repurchase intentions to help predict future purchasing behavioral intentions and loyalty [22]. On the other hand, customer retention behavior is defined as customers stating the actual continuation of a relationship with the organization. It is well known in marketing that past customer behavior tends to be a relatively good predictor of future customer behavior. However, most researchers focus on assessing the repurchase intention attitude and neglect assessing actual customer retention behavior [22].

##### 2.2.5.2 Willingness to recommend an attitude

Word-of-mouth intention has been of importance to researchers in the past 30 years. Thus far, there is very little scientific research relating the intention of the recommendation to the actual recommendations [22].

## 3. Methods

### 3.1 Research design

This is part of a broad project that aims to strategically develop Palestinian hospitals using BSC. This research is a cross-sectional study. The questionnaire was created and validated based on the key authors Kaplan and Norton theortical framework [1] and the best practices for developing and validating health and behavioral scales [29].

### 3.2 Item generation

The first panel consists of five authors in this research. Two researchers in health management (first and fourth), two hospital managers who are also expert researchers in health management (sixth and seventh), and one expert in the BSC tool (fifth) provided expert input on all stages of instrument development. First, we performed a systematic review [3], in which 797 KPIs were extracted from 36 BSC implementations at HCO worldwide. Then, categorization and regrouping of these KPIs resulted in 46 subdimensions and 13 major dimensions that are frequently used by healthcare managers and are important for PE and the strategic development of HCO [3]. Next, this panel performed a four-round Delphi method [30]. In the first round, the panel prepared a survey for hospitals' top managers to rate the resulting 46 subdimensions on a 10-point semantic scale based on their importance for the strategic development of their hospitals. A description for each subdimension using the shortlisted KPIs was included in the manager survey. In the second round, the panelists reviewed the item face validity per subdimension [31]. Next, the first author asked a second panel consisting of 13 top hospital managers from 4 Palestinian hospitals to answer this survey individually. Additionally, hospital managers were asked to mention whether they considered any other subdimension or KPI that was not listed as essential. The subdimensions with an average score above 0.7 were chosen for the next step based on their ratings. In the third round, the first panel reviewed the resulting important subdimensions at the previous step and decided which subdimensions the patients could be engaged in its evaluation. As a result, 24 subdimensions resulted. In the fourth round, the panelists revised each item wording and clarity to patients. As a result, 52 items remained. In the fourth round, the panelists rated the relevance and importance for each remaining item based on four- and three-point ordinal scales, respectively [32]. Next, the first author calculated the content validity ratio (CVR), the item content validity index (I-CVI), the scale content validity index (S-CVI), and universal agreement among experts for the content validity index (CVI-UA) to assess the content validity per item and scale [32]. Only the items rated 0.99 or above in CVR were included as per Lawshe guidelines [33]. However, dimensions that scored 0.80-0.99 indicated the need to be revised. For the CVI, items that scored less than 0.60 were eliminated. Items that scored 0.6-0.79 were revised [32]. See Figure (2).

The panelists suggested using a three-point Likert scale: yes, neutral (I do not know), and no. This choice was due to the high number of the remaining items, the evidence of a high nonresponse rate of patients to the five-point Likert scale-validated tools [34–37], and the possibility for assessing item availability using yes/no questions. Additionally, this was found to lead to a faster and better item response, specifically considering the pandemic load on hospitals. All authors were asked to revise the instrument, and the final modifications were made accordingly.

### 3.3 Linguistic validation and translation

Since the resulting dimensions at the systematic review were in English, the questionnaire items were initially developed in English. Then, it was translated to Arabic. All translations were prepared as per the translation and validation guidelines [38]. Forward translation (English to Arabic) was performed by the first author, who has great knowledge about health management terminology and is bilingual (English and Arabic). The second forward translation was performed by a professional in linguistics who is bilingual and a certified English-Arabic translator. Both translations were reconciled into a single version. Then, it was back-translated by another translator into English. The first author performed a final review to produce the final corrected translation. An expert checked the final form in the BSC, and minor modifications were recommended.

### 3.4 Pretest and the internal consistency

The first version of the questionnaire was piloted in one NGO hospital in the south of West Bank. For that, 30 patients were asked to answer the first version of the questionnaire. They were asked to write their comments regarding language simplicity. The time needed to fill the questionnaire was also recorded. Items were coded before performing the analysis by IBM SPSS statistics 21 software. Then, Cronbach's alpha was calculated for each perspective to evaluate the internal consistency [39], and values above 0.6 were considered acceptable. Based on the results, some items were modified or deleted.

### 3.5 Sampling procedure and power calculation

Institutional Review Board (IRB) approval for this research was received on 31 May 2020. All methods described in this study were approved by the Research and Ethics Committee at the Faculty of Medicine and Health Sciences at An Najah National University with the reference code number (Mas, May/20/16). Afterward, requests at 15 hospitals in West Bank and three hospitals in Jerusalem were applied between June and December 2020. The hospitals were selected using a convenience sample. However, the total number of beds per administrative type and governorate was considered for choosing the participants (HCO and patients). Public hospital approval was first applied to the Palestinian Ministry of Health. Then, the request was applied to each hospital individually for all hospital types. The final form of the questionnaire was distributed between January and October 2021. The sample size was calculated according to Steven K. Thompson sample size equation [40];

$$n = \frac{N \times p(1-p)}{[N - 1 \times (d^2 \div z^2)] + p(1 - p)}$$

where n is the sample size, N is the population size, p is the estimated variability in the population (0.5), d is the margin of error (0.05), and the z score is at the 95% confidence interval (1.96). In our study, N was the population volume in the Palestinian territories [8]. Therefore, the needed sample size was found to be = 385 patients. Additionally, studies considered 300 participants as a good sample size to successfully run each exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) or 5 respondents per parameter [41–43]. Splitting the sample to perform EFA and CFA is

recommended to perform construct validity [44]. Therefore, a total of 1000 questionnaires were distributed, considering the lower response rate during the pandemic.

### 3.6 Data collection and participants

The first author and four medical students at An-Najah University performed the data collection. Each medical student received three hours of training on BSC and the data collection steps and ethics by the first author. Tasks and hospitals were delegated to them according to their living area: eastern Jerusalem and north, middle, and south of West Bank. Gaza Strip was excluded due to the political situation and accessibility obstacles during the study. Moreover, five hospitals were excluded: two military hospitals that were not operating yet, one psychiatric hospital, and two rehabilitation hospitals. We sought variation in our sample regarding hospital size, area, and leadership style. For that, the maximum variation sampling strategy was used. The number of hospitals and the number of beds per leadership style were considered upon recruiting the sample [8]. The patients were conveniently chosen based on their willingness to participate in this research.

Printed questionnaires were distributed to respondents instead of sending the questionnaires via email to reduce nonresponse bias [45]. Additionally, all participants were asked to agree on participation in a consent form that is coherent with the declaration of Helsinki ethical principles [46]. Patients were informed that participation was confidential. Additionally, all patients were informed that participation was voluntary, so they could refuse participation in the study or withdraw at any time. To reduce the response bias [45], the “I don’t know (neutral)” answer was added as an option, since experiences and attitudes can sometimes be uncertain [21]. Second, the data collectors ensured that the number of missing answers was minimized by checking the questionnaires upon retrieval. In case of missing parts, they drew the participant attention to answer them. When entering data, if any questions were found to be still missing, they were entered as I don’t know.

The inclusion and exclusion criteria were set to be a Palestinian patient above 15 years old of any gender. Outpatients should have finished receiving medical care at the assessed hospital or had received medical care at least once previously and returned to the same hospital. Inpatients should have been admitted for at least one day. The following departments were included: emergency room, internal medicine, surgery, gynecology, and pediatrics. In the emergency department, the questionnaires were completed by the patient companions. Additionally, in the pediatric department, the questionnaires were completed by one parent of the child. For the rest, questionnaires were completed by patients themselves; unless they were unable to complete the questionnaire, the questionnaires were read to them by the data collector or a family member and completed according to patient answers. To distinguish, a question was added to ask the respondent if his responses were based on his own, family, or friends’ experiences.

### 3.7 Statistical Analysis

The normality was tested using the Shapiro–Wilk test. The frequencies were used to analyze patient sociodemographics and the participating HCO characteristics. Our sample was split based on admission status to assess construct validity using EFA and CFA. EFA was performed for the inpatient sample using principal axis factoring with the Promax rotation method [47] in IBM SPSS statistics 21 software. The Kaiser–Meyer–Olkin (KMO) and Bartlett’s sphericity tests were tested to determine the adequacy of the EFA [48]. The inclusion or exclusion of a component was determined by an eigenvalue ≥1 [49] and the visual assessment of Cattell’s scree plot [50]. Item inclusion or exclusion was determined by a factor loading ≥0.50 and factor loadings on the assigned construct higher than all cross-loading of other constructs [42].

Second, CFA was performed for the components that resulted in EFA using the outpatient sample. The maximum likelihood estimation method in IBM Amos 26 Graphics software was applied. The goodness of fit for the competing models was evaluated through the most commonly used fit indices. Minimum discrepancies were divided by degrees of freedom less than five and closer to zero, *P* value higher than 0.05, goodness-of-fit index (GFI), comparative fit index (CFI), Tucker–Lewis’s index (TLI), and cutoff values close to 0.95. Additionally, a root mean square error of approximation (RMSEA) <0.06 and standardized root mean square residual (SRMR) value <0.08 are needed before we can conclude that there is a relatively good fit between the hypothesized model and the observed data. [51, 52]. Item inclusion or exclusion in CFA was determined by a factor loading ≥0.50.

Third, the interitem correlation (IIC) and the corrected item-total correlation (CITC) were calculated [53]. In this study, items with a correlation higher than 0.9 were considered redundant and deleted [54]. A correlation of 0.3 was considered the lower limit. Additionally, the composite reliability (CR) per construct was evaluated after performing CFA. CR is preferred over Cronbach’s alpha, specifically in structural equation modeling [55]. In the current study, a CR ≥0.6 was considered sufficient [56, 57].

Finally, the Fornell-Lacker criterion was used to evaluate convergent and discriminant/divergent validities [58]. The average variance extracted (AVE) was considered adequate for convergent validity if it was higher than 0.5. However, if a value <0.5 with CR >0.6, the convergent validity of the construct was still considered adequate [58]. To establish discriminant validity, the square root of the AVE (SQRT) should have a greater value than the correlations with other latent constructs [56]. Additionally, the construct uniqueness was evaluated depending on the value of Spearman correlation (r) with other constructs at the same scale. Researchers recommended the separation of dependent and independent variables since the correlation between them can be misleading in assessing discriminant validity [59]. Therefore, we assessed r for the independent and dependent constructs separately. Then, r was described as negligible when r <0.2, low (r = 0.2-0.49), moderate (r = 0.5-0.69), high (r = 0.7-0.85), or very high (r = 0.86-1.00). In this study, the absence of high or very high r between the subscale constructs indicated discriminant validity [60].

## 4. Results

### 4.1 Item generation and scoring

The demographics and characteristics of the second-panel hospital managers are shown in Table (1). The content validity resulted in removing one item and indicated that a revision is needed for eight items. The revised items required either further clarification and rewording or modification for specific participants. For example, CVR results indicated that financial and price items should not be included for nonprofit hospitals. Additionally, CVI results showed that particular items are relevant only to inpatients. This step raised the S-CVI, CVI-UA, and CVR from 0.90, 0.63, and 0.95 to 0.95, 0.78, and 0.97, respectively.

Table (1) is to be inserted here.

### 4.2 The instrument's structure and items

The patient sociodemographics and hospital characteristics section included age, gender, scientific degree, working sector, insurance availability, and type. Moreover, the number of visits to the evaluated hospital compares the attitudes of the new and previous customers. The number of earlier visits is considered necessary in the analysis since past customer behavior tends to be a good predictor of future behavior [21]. Moreover, the information source on which the respondent evaluation was built was recorded since perceptions and attitudes may emerge from direct personal experience or from observing other people's experiences, such as family and friends' experiences [20]. The second section of the questionnaire was designed to measure patient experiences in light of BSC perspectives and their attitudes toward them, including patient satisfaction, PQ, PI, and loyalty.

#### 4.2.1 The financial perspective

It evaluated the health services and medication's price affordability. This section was answered only by patients who did not have insurance.

#### 4.2.2 The internal perspective

This perspective assessed safety, time, and service availability. On the other hand, the PI of the cure rate, accuracy, complications, and PQ of services and medication were measured in the attitude section.

#### 4.2.3 The knowledge and innovation perspective

Information and training provided to patients were assessed in the experience section. Additionally, we assessed the PI of hospital technology and employee competencies in the attitude section.

#### 4.2.4 The customer perspective

It assessed patient-centeredness and the HCW-patient communication experience. The attitude section assessed actual patient satisfaction and loyalty attitudes. In previous studies, validated items for loyalty measurement included satisfaction measurement and loyalty attitude measurement, specifically the recommendation and return intentions [22, 26]. Using a single item to directly assess actual patient satisfaction was suggested to be better than its assessment through multidimensional items [61].

#### 4.2.5 The environment perspective

It evaluated the hospital building environment and the hospital capacity, ease of access, and female concern experiences. On the other hand, a comparison with the other hospitals' medical and social PIs was included in the attitude section.

Finally, three items were reversed in the instrument, PIN9, which assessed the long waiting time. Additionally, PIN4, PIN5, and PIN6 assessed readmission, referral to other hospitals, and postoperative infection probability expectations, respectively.

### 4. 3 The pretest and the internal consistency

The pretest was performed at one NGO hospital in the south of West Bank. Patients found the length of the questionnaire appropriate. Additionally, the layout was well accepted and clear. They gave specific minor comments that were incorporated. These corresponded to the rewording of a few items. The time for completing the questionnaire was less than 10 minutes.

Consequently, few modifications were made after piloting. Cronbach's alpha was calculated per BSC perspective. All perspectives had a Cronbach's alpha above 0.7 at the pretest, except for the environmental perspective, which was 0.59. Hence, some of its items were moved to other perspectives, and five items were deleted. As a result, 52 and 50 items remained for inpatients or outpatients, respectively.

### 4.4 Linguistic validation and translation

The final English and Arabic questionnaire forms were ready for use.

### 4.5 Sample size and characteristics

Since the research coincided during the COVID-19 pandemic, hospital approvals took six to nine months until received. Only 15 hospitals out of 18 agreed on participation. The data collection was performed between January and September 2021. The data of the pretest at one hospital were excluded. Next, we distributed 1000 questionnaires at the remaining 14 hospitals. As a result, 740 were returned (response rate was 74%). The characteristics and sociodemographics of the respondents are shown in Tables 2 & 3.

Table (2) is to be inserted here.

Table (3) is to be inserted here.

### 4.6 Statistical analysis

The statistical analysis using the Shapiro–Wilk test showed that the data were not normally distributed, so nonparametric tests were decided to be used. Then, construct validation was assessed for the instrument.

#### 4.6.1 Construct validity in EFA

EFA resulted in 37 items with loadings higher than 0.50 for 12 components. Eigenvalues for all components were higher than one. The KMO was 0.901, reflecting very high sampling adequacy [48, 56], and Bartlett's test was also significant. The cumulative variance was 63.29%. See Table (4). The 12 components were patient attitude toward BSC perspectives (BSCP ATT), patient experience (PT EXR), service experience (SERV EXR), price experience (PR EXR), building experience (BUIL EXR), access experience (ACC EXR), complication perceived image (COMP IMAGE), technology experience (TECH EXR), information experience (INFO EXR), hospital social responsibility perceived image (HSRP IMAGE), and waiting time experience (WT EXR). One item (SAT2) loaded on the 12th component. However, this item had a higher loading on the BSCP ATT. None of the specific inpatient items had loadings higher than 0.50. Moreover, the scree plot showed the necessity to delete the last three components.

Table (4) is to be inserted here.

#### 4.6.2 Construct validity in CFA

The resulting nine components in EFA were tested in the Amos program. The model was edited based on the item loadings, model fit indices, and calculations in the convergent, discriminant, CR, IIC, and CITC at the next step until we arrived at the best model. First, adding two items that did not have loadings to the INFO EXR construct showed good loadings in CFA. The same was regarding BSCP ATT and TECH IMAGE constructs. Second, splitting the BUIL EXR component into two separate constructs, building environment experience (BUILENV EXR) and building capacity experience (BUILCAP EXR), improved the item loadings and the model fit. Third, PEN9 and PLE7 items were removed from the PT EXR construct because they have loadings lower than 0.50. On the other hand, PIN 14 and PIN 16 were added to the latter construct since both had loadings higher than 0.50 and improved the model fit. Moreover, merging the TECH IMAGE and COMP IMAGE items at the BSCP ATT construct resulted in loadings lower than 0.5 and IIC lower than 0.30. Hence, three separate constructs in the attitude section were decided. Finally, the modification indices in the Amos program were utilized to improve the model. The final model revealed that the CMIN/df, CFI, GFI, TLI, RMSEA, and SRMR indices in CFA were above or close to the cutoff points, reflecting a good fit model. Despite that, the *P* value was <0.001, which can be referred to as its sensitivity to normality. See Figure (3) and Table (5).

Figure (3) is to be inserted here.

Table (5) is to be inserted here.

#### 4.6.3 Composite reliability and interitem correlations

The composite reliabilities for all constructs were higher than 0.6 except the SERV EXR construct. However, this construct's IIC and CTIC were higher than 0.3. The other constructs also had IICs higher than 0.3, and their CITC ranged from 0.328-0.853, reflecting satisfactory IIC and CITC. See Table (6).

Table (6) is to be inserted here.

#### 4.6.4 Convergent and discriminant validity

Convergent validity was less than 0.5 for BSCP ATT, BUILENV EXR, PTCOMINF EXR, SERV EXR, and COMP_IMAGE. However, the CR, IIC, and CITC showed satisfactory results [58], except the SERV EXR, which had a CR equal to 0.50 but an IIC and CITC higher than 0.3. On the other hand, the square roots of the AVE were higher than the off-diagonal correlations between constructs. Additionally, the lower the correlation between constructs indicates each construct uniqueness. The correlations between the independent constructs were either negligible or low, except between two constructs; the PT EXR and INFO EXR, which was moderate. Merging the two constructs lowered the loadings and the model fit indices in CFA. The same was perceived regarding merging the BUILENV EXR and BUILCAP EXR constructs. Consequently, separate constructs were determined, as mentioned earlier. In regard to the independent constructs, negligible or low correlations existed among them. Neither high nor very high correlations existed between the independent constructs. Therefore, this establishes discriminant validity and the uniqueness of the independent constructs. The same holds true for the dependent constructs. In other words, convergent validity was met for all constructs except SERV EXR. In comparison, discriminant validity was met for all constructs, as shown in Tables 7 & 8.

Table (7) is to be inserted here.

Table (8) is to be inserted here.

## 5. Discussion

### 5.1 Discussion of the main results

In congruent with this paper aim, we developed, translated, and validated the BSC-PATIENT instrument to engage patients at the evaluation of hospitals by measuring their experiences and attitudes toward the hospital in light of the BSC perspectives: the financial, internal, knowledge and innovation, customer, and

environmental perspectives. Our findings showed that patient attitude toward all BSC perspectives and dimensions loaded on one construct, except the images of technology and complications, loaded separately.

The instrument was customized to be compatible with Palestinian hospitals. Statistics revealed that out-of-pocket household payments constituted 39.8% of the Palestinian territories' total healthcare expenditures in 2018 [62]. This number is close to the results in our sample, which showed that 14.73% of patients did not have any insurance, and 19.32% had private insurance. Additionally, our analysis shows that another 35.41% or 1.49% of our sample had public or UNRWA insurance, respectively, but were receiving treatment at an NGO or private hospital at the study time. This situation indicates that the patients either made out-of-pocket payments or that the government paid a medical referral to private or NGO hospitals [8]. Therefore, incorporating the financial perspective consideration in this paper proved to be vital. Additionally, many BSC implementations in Afghanistan and Bangladesh revealed the need to consider the social and cultural perspective in evaluation, specifically, female attentiveness concerns [63–67]. The authors believed that this was also the case in Palestine, so the BSC-PATIENT included such items. However, in different cultures, this may not be such important. Hence, these items can be removed or replaced with other customized environment-related items. Finally, the technology perspective varies among Palestinian hospitals. Even though the Ministry of Health Hospitals and many other private hospitals have adopted the health medical information system for years, some hospitals still use the manual system for documentation. The authors also considered this perspective important in this evaluation.

The causal relationships between BSC dimensions that were described in BSC strategic maps may impose a challenge on producing a good fit model, specifically discriminant validity. Despite this challenge, our model proved satisfactory construct, convergent, and discriminant validity. The composite reliability was higher than 0.6 for all constructs except the SERV EXR construct. This may indicate that a separate evaluation for this construct item is needed. Moreover, the IIC and the CITC were satisfactory. In general, this questionnaire proved reliable and valid for engaging patients in hospital evaluations by measuring their experiences and attitudes toward Palestinian hospitals.

### 5.2 Comparison with other BSC studies

One of the main reasons for developing this questionnaire was the lack of patients involved in the evaluation process of BSC perspectives, which was addressed by all BSC reviews [3, 4, 19]. BSC studies utilized managerial observations or hospital records to evaluate financial, internal, knowledge and innovation and environmental perspectives. At the same time, patients were incorporated to assess patient satisfaction only instead of engaging them in BSC perspective evaluations.

### 5.3 Comparison with other validated instruments

#### 5.3.1 Service Quality scale (SERVQUAL)

One of the most popular models to measure service quality is the 44-question SERVQUAL instrument [68]. However, SERVQUAL has been criticized for encountering various shortcomings [69, 70]. First, numerous studies have questioned whether SERVQUAL is applicable as a generic scale for measuring service quality in all settings [69], as it was not initially designed for hospitals. In contrast, BSC-PATIENT was explicitly designed for hospitals. Second, the concept of “subtraction” in the SERVQUAL model is not equivalent to psychological function [69]. However, BSC-PATIENT was designed to be coherent with psychological definitions by distinguishing between experience observations and attitudes. Third, researchers uncovered some shortcomings of the discriminant validity at SERVQUAL [69]. They explained that reliability, responsiveness, assurance, and empathy dimensions were not distinct from each other and loaded into one factor in many studies due to the high degree of intercorrelation [69]. All BSC-PATIENT constructs passed discriminant validity. Fourth, SERVQUAL has been criticized for focusing on functional quality, not reputational quality [70]. This challenge was overcome in BSC-PATIENT through the separation of observations and attitudes.

#### 5.3.2. Press Ganey

Another commonly used instrument is Press Ganey [71], a 21-question instrument explicitly developed to measure hospital patient experience. However, Press Ganey also has a few shortcomings. Many studies using this instrument reported evidence of nonresponse bias [34, 35]. The response rate for BSC-PATIENT was 75% despite the COVID-19 situation. Many patients commented that the questionnaire was interesting to fill. This can also be referred to as the simplicity of the three-point scale, unlike the five- and seven-point Likert scales, which can contribute to greater respondent burden and fatigue and may lead to higher refusal rates [61]. Finally, building, services, technology, price experiences assessing items, and patient attitudes were not considered necessary in Press Ganey.

#### 5.3.3 Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)

The 29-question Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) [72] is widely used in the United States of America (USA) to evaluate patient experiences. It incorporates eight dimensions. However, the response rate for this instrument was found to be low [36, 37]. Additionally, accessibility, price, and technology experiences were neglected. Moreover, HCAHPS allows researchers to evaluate the overall patient satisfaction rate based on their subratings for different experience constructs, such as communication with HCW perception [36, 37, 73].
Although experience perceptions can predict patient attitudes, including satisfaction, a separate evaluation of experiences and satisfaction and a direct satisfaction assessment were recommended [61]. This point was taken into account when designing the BSC-PATIENT in the design of BSC-PATIENT.

### 5.4 Strengths and limitations

In general, this paper has several strengths. First, BSC-PATIENT is the first instrument that engages patients in BSC perspective assessment. Second, this instrument can determine patient attitudes, including PI toward BSC perspectives, PQ, and satisfaction and loyalty. Third, to our knowledge, this is the first

paper to distinguish between patient experiences and patient attitudes, which will allow us to study the relationship between patient experiences and attitudes in future studies. Fourth, this instrument was customized to be used for all insurance, leadership, and admission statuses. Fifth, this instrument was designed based on KPIs extracted from BSC implementations in primary, secondary, and tertiary healthcare settings in low-, middle-, and high-income countries worldwide. Hence, the implementation of BSC-PATIENT can be generalized to different healthcare settings and countries. However, the instrument may need some customization based on the healthcare setting strategy and the country's properties. For example, we customized the BSC-PATIENT at the environmental perspective based on Palestinian culture, the financial perspective based on leadership style, the innovation and technology perspective based on the health information system in Palestine, and the few items specific for inpatients based on admission status. Finally, this paper offers a comprehensive hospital assessment from patient perspectives during COVID-19. To date, no study has assessed Palestinian hospital performance during this era. However, this instrument has some limitations. Despite this instrument assessing items such as patient education on infection control measures, it lacks COVID-19-specific items, as this instrument was designed before the COVID pandemic, so COVID-19-related items can be considered in future versions of the BSC-PATIENT instrument. Second, patient literacy was not assessed. However, the academic qualifications were evaluated at the demographics to be considered in the analysis. Third, measuring patient experiences in the past may involve a bias of recall. Additionally, participant bias may have occurred since the sample was convenient and the included hospitals agreed on participation. However, the high percentage of the included hospitals (30%) from the total number of hospitals at West Bank and including all leadership style types from all regions may have reduced the selection bias. Another limitation is that we could not validate this instrument in English due to our inaccessibility to English-speaking patients. Future research needs to consider testing the psychometric properties of BSC-PATIENT in an English-speaking country.

## 6. Conclusion

The BSC-PATIENT instrument was developed to engage patients in the PE of hospitals. This instrument was validated in Arabic and customized for Palestinian hospitals. This is the first instrument to engage patients in evaluating their experiences and attitudes toward the BSC perspective. It consists of 19 items assessing patient experience observations and 15 items assessing patient attitudes. Both experiences and attitudes were designed based on BSC perspectives. The findings of this research showed adequacy in the psychometric properties of this instrument and suggest some recommendations for future research. First, we tested the psychometric properties of the BSC-PATIENT in English and other languages in different countries. Second, we consider developing instrumental BSC perspectives to engage other stakeholders in the PE of hospitals, such as doctors, nurses, and managers. Third, this instrument was used to assess the impact of patient experience on patient attitudes toward the hospital, specifically the PI, PQ, and satisfaction and loyalty. Fourth, managers must consider using a comprehensive approach for the PE of hospitals instead of limiting it to financial or internal indicators. Fifth, we compared the differences in patient experience and attitudes based on patient and hospital characteristics. Finally, enhancing patient engagement in the evaluation process instead of focusing on satisfaction alone must be considered in future BSC and PE implementations. Involving stakeholders in BSC's comprehensive evaluation will lead to a better and deeper understanding of hospital PE.

## 7. Declarations

**Ethics approval and consent to participate:**

The Institutional Review Board (IRB) was received from the Research and Ethics Committee at the Faculty of Medicine and Health Sciences at An-Najah National University with reference code number (Mas, May/20/16) on 31 May 2020. Additionally, approval was obtained from the Palestinian Ministry of Health and 15 hospitals individually. Informed consent was attached to the questionnaire for each patient, and they were asked to agree on participation. All methods were carried out in accordance with relevant guidelines and regulations.

**Consent for publication:**

Not Applicable.

**Availability of data and materials:**

The datasets generated and/or analyzed during the current study are not publicly available because the data are still not fully analyzed and the research is still in process but are available from the corresponding author (F. A) on reasonable request with the permission of the UNRWA, Palestinian Ministry of Health, and Al Makassed Hospital.

**Competing interests:**

The authors have declared that no competing interests exist.

**Funding:**

No funding was received for conducting this study. Data collection expenses were personally funded by F. A.

**Authors contributions**

This paper's conception and design were planned by F. A, A. A, A. N, I. B & D.E. Additionally, F. A & A. A were responsible for hospital approvals. F. A was responsible for the data acquisition, statistical analysis, interpretation of data, and writing the final draft. A. A, A. N, I. B & D. E supervised this research. F. A, S. H, D. O, A. A., A. N., I. B, D. E substantially revised the final manuscript draft. F. A, S. H, D. O, A. A., A. N., I. B, D. E approved the submitted version (and any

substantially modified version that involves the author's contribution to the study) and agreed to be personally accountable for the author's contributions and to ensure that the accuracy and integrity of any part of the work were appropriately investigated and resolved.

**Acknowledgments:**

We would like to thank Mr. Saad Abuzahra, Mr. Yazan Al-Habil, Mr. Haroun Nairoukh, and Ms. Mofeeda Afifi for their contribution to the data collection and entry at the north, middle, south of West Bank, and eastern Jerusalem, respectively. Additionally, we thank Ms. Duha al Shellah for her contributions in revising this paper.

## 8. References

1. Kaplan R, Norton D. The balanced scorecard–measures that drive performance. Harv Bus Rev. 1992;70:71–9.
2. Kalender ZT, Vayvay Ö. The Fifth Pillar of the Balanced Scorecard: Sustainability. Procedia - Soc Behav Sci. 2016;235 June 2017:76–83. doi:10.1016/j.sbspro.2016.11.027.
3. Amer F, Hammoud S, Khatatbeh H, Lohner S, Boncz I. A Systematic Review: The Dimensions and Indicators utilized in the Performance Evaluation of Health Care Organizations- An Implication during COVID-19 Pandemic. Preprint: http://dx.doi.org/10.1101/2021.06.26.21259568. 2021.
4. Amer F, Hammoud S, Khatatbeh H, Lohner S, Boncz I, Endrei D. Balanced Scorecard in Health Care Organizations: Is it Really Beneficial ? A Systematic Review. Preprint: 10.1101/2021.08.05.21261666. 2021.
5. Meena K, Thakkar J. Development of Balanced Scorecard for healthcare using Interpretive Structural Modeling and Analytic Network Process. J Adv Manag Res. 2014;11:232–56.
6. Giacaman R, Abdul-Rahim HF, Wick L. Health sector reform in the Occupied Palestinian Territories (OPT): Targeting the forest or the trees? Health Policy Plan. 2003;18:59–67.
7. Barghouthi, M. & Lennock J. Health in Palestine: Potential and Challenges. Jerusalem: Palestine Economic Policy Research Institute (MAS). 1997.
8. Ministry of Health. Health Annual Report, Palestine 2020. May, 2021.
9. Sabella AR, Kashou R, Omran O. Assessing quality of management practices in Palestinian hospitals. Int J Organ Anal. 2015;23:213–32. doi:10.1108/IJOA-03-2014-0747.
10. Palestinian News and Info Agency (WAFA). Hospitals in Palestine. 2020. https://info.wafa.ps/ar_page.aspx?id=14977. Accessed 10 Jul 2021.
11. Regional Health Systems Observatory (EMRO). Health Systems Profile: Palestine. Cairo, Egypt; 2006.
12. World Health Organisation (WHO). Vulnerability and the International Health Response in the West Bank and Gaza Strip,An analysis of health and the health sector. 2001; November.
13. World Health Organisation (WHO). Health conditions in the occupied Palestinian territory, including east Jerusalem, and in the occupied Syrian Golan Report. 2020.
14. Hammoudeh W, Kienzler H, Meagher K, Giacaman R. Social and political determinants of health in the Occupied Palestine Territory (OPT) during the COVID-19 pandemic: who is responsible? BMJ Glob Heal. 2020;5:e003683.
15. Gagliardi AR, Lemieux-Charles L, Brown AD, Sullivan T, Goel V. Barriers to patient involvement in health service planning and evaluation: An exploratory study. Patient Educ Couns. 2008;70:234–41. doi:10.1016/j.pec.2007.09.009.
16. Anderson NN, Baker GR, Moody L, Scane K, Urquhart R, Wodchis WP, et al. Approaches to optimize patient and family engagement in hospital planning and improvement: Qualitative interviews. Heal Expect. 2021;24:967–77. doi:10.1111/hex.13239.
17. Bellows M, Kovacs Burns K, Jackson K, Surgeoner B, Gallivan J. Meaningful and effective patient engagement: What matters most to stakeholders. Patient Exp J. 2015;2:18–28.
18. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: A systematic review. BMC Health Serv Res. 2014;14:1–9.
19. Bohm V, Lacaille D, Spencer N, Barber CE. Scoping review of balanced scorecards for use in healthcare settings: development and implementation. BMJ open Qual. 2021;10:1–10. doi:10.1136/bmjoq-2020-001293.
20. American Psychological Association. APA Dictionary of Psychology. https://dictionary.apa.org/. Accessed 25 Jul 2021.
21. David Susman. Attitudes and Behavior in Psychology. 2021. https://www.verywellmind.com/. Accessed 25 Jul 2021.
22. Keiningham TL, Cooil B, Aksoy L, Andreassen TW, Weiner J. The value of different customer satisfaction and loyalty metrics in predicting customer retention, recommendation, and share-of-wallet. Manag Serv Qual An Int J. 2007;17:361–84.
23. Genoveva G. Analyzing of Customer Satisfaction and Customer Loyalty Based on Brand Image and Perceived Service Quality. J US-China Public Adm. 2015;12.
24. Zhou W-J, Wan Q-Q, Liu C-Y, Feng X-L, Shang S-M. Determinants of patient loyalty to healthcare providers: An integrative review. Int J Qual Heal Care. 2017;29:442–9. doi:10.1093/intqhc/mzx058.
25. Kotler P, RN. Clarke. Marketing for healthcare organizations. 1st edition. Englewood Cliffs, NJ: Prentice Hall; 1987.
26. Taneja U. Brand image to loyalty through perceived service quality and patient satisfaction: A conceptual framework. Heal Serv Manag Res. 2020;:095148482096230. doi:10.1177/0951484820962303.
27. Kotler P, Keller K. Marketing Management. 15th edition. Saddle River: Prentice Hall; 2014.

28. Don Peppers. Customer Loyalty: Is It an Attitude? Or a Behavior? Peppers & Rogers Group. 2012.
29. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Front Public Heal. 2018;6 June:1–18.
30. Dalkey N, Helmer O. An experimental apllication of Deplhi method to use of experts. Management Science. 1963;9:458–67.
31. Simbar M, Rahmanian F, Nazarpour S, Ramezankhani A, Eskandari N, Zayeri F. Design and psychometric properties of a questionnaire to assess gender sensitivity of perinatal care services: A sequential exploratory study. BMC Public Health. 2020;20:1–13.
32. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J Caring Sci. 2015;4:165–78. doi:10.15171/jcs.2015.017.
33. Lawshe CH. A QUANTITATIVE APPROACH TO CONTENT VALIDITY. Pers Psychol. 1975;28:563–75. doi:10.1111/j.1744-6570.1975.tb01393.x.
34. Tyser AR, Abtahi AM, McFadden M, Presson AP. Evidence of non-response bias in the Press-Ganey patient satisfaction survey. BMC Health Serv Res. 2016;16:350.
35. Patel AB, LaCouture T, Hunter K, Tartaglia A, Kubicek GJ. Pitfalls and Predictors of Poor Press Ganey Patient Satisfaction Scores. Int J Radiat Oncol. 2014;90:S719–20. doi:10.1016/j.ijrobp.2014.05.2101.
36. Mann RK, Siddiqui Z, Kurbanova N, Qayyum R. Effect of HCAHPS reporting on patient satisfaction with physician communication. J Hosp Med. 2016;11:105–10.
37. Siddiqui ZK, Wu AW, Kurbanova N, Qayyum R. Comparison of hospital consumer assessment of healthcare providers and systems patient satisfaction scores for specialty hospitals and general medical hospitals: Confounding effect of survey response rate. J Hosp Med. 2014;9:590–3.
38. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract. 2011;17:268–74. doi:10.1111/j.1365-2753.2010.01434.x.
39. Cho E, Kim S. Cronbach's Coefficient Alpha: Well Known but Poorly Understood. Organ Res Methods. 2015;18:207–30. doi:10.1177/1094428114555994.
40. Thompson SK. Sampling. Third Edition. Hoboken, New Jersey: A JOHN WILEY & SONS, INC.; 2012.
41. Comrey AL, Lee HB. A First Course in Factor Analysis. Psychology Press; 2013.
42. Williams B, Onsman A, Brown T. Exploratory factor analysis: A five-step guide for novices. Australas J Paramed. 2010;8.
43. BENTLER PM, CHOU C-P. Practical Issues in Structural Modeling. Sociol Methods Res. 1987;16:78–117. doi:10.1177/0049124187016001004.
44. Knafl GJ, Grey M. Factor analysis model evaluation through likelihood cross-validation. Stat Methods Med Res. 2007;16:77–102.
45. Sedgwick P. Non-response bias versus response bias. BMJ. 2014;348 April:9–10.
46. World Medical Association Declaration of Helsinki. JAMA. 2013;310:2191.
47. Dean J. Choosing the Right Type of Rotation in PCA and EFA. Shiken JALT Test Eval SIG Newsl. 2009;13 November:20–5.
48. Kaiser HF, Rice J. Little Jiffy, Mark Iv. Educ Psychol Meas. 1974;34:111–7.
49. Larsen R, Warne RT. Estimating confidence intervals for eigenvalues in exploratory factor analysis. Behav Res Methods. 2010;42:871–6.
50. Cattell RB. The Scree Test For The Number Of Factors. Multivariate Behav Res. 1966;1:245–76.
51. Hooper D, Coughlan J, Mullen MR. Structural equation modelling: Guidelines for determining model fit. Electron J Bus Res Methods. 2008;6:53–60.
52. Shi D, Lee T, Maydeu-Olivares A. Understanding the Model Size Effect on SEM Fit Indices. Educ Psychol Meas. 2019;79:310–34.
53. Cohen J. Set Correlation and Contingency Tables. Appl Psychol Meas. 1988;12:425–34.
54. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53–5.
55. Peterson RA, Kim Y. On the relationship between coefficient alpha and composite reliability. J Appl Psychol. 2013;98:194–8. doi:10.1037/a0030767.
56. Ab Hamid MR, Sami W, Mohmad Sidek MH. Discriminant Validity Assessment: Use of Fornell & Larcker criterion versus HTMT Criterion. J Phys Conf Ser. 2017;890.
57. Hair J, Hult GTM, Ringle C SM. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Los Angeles: SAGE Publications, Incorporated; 2014.
58. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. J Mark Res. 1981;18:39–50.
59. Hair J, Wolfinbarger M, Money AH, Samouel P, Page MJ. Essentials of Business Research Methods. Routledge; 2015.
60. Bookter Al. Convergent and Divergent Validity of the Learning Transfer Questionnaire. 1999.
61. Willits FK, Theodori GL, A. E. Luloff. Another look at likert scales. J Rural Soc Sci. 2016;31:126.
62. Palestinian Central Bureau of Statistics (PCBS). Percentage Distribution of Current Expenditure on Health in Palestine by Financing Schemes, 2017-2018. 2020. https://www.pcbs.gov.ps/. Accessed 27 Nov 2021.
63. Rowe JS, Natiq K, Alonge O, Gupta S, Agarwal A, Peters DH. Evaluating the use of locally-based health facility assessments in Afghanistan: a pilot study of a novel research method. Confl Health. 2014;8:1–9.
64. Edward A, Osei-bonsu K, Branchini C, Yarghal T, Arwal SH, Naeem AJ. Enhancing governance and health system accountability for people centered healthcare: an exploratory study of community scorecards in Afghanistan. BMC Health Serv Res. 2015;15:1–15.
65. Khan MM, Hotchkiss RD, Dmytraczenko T, Zunaid Ahsan K. Use of a Balanced Scorecard in strengthening health systems in developing countries: An analysis based on nationally representative Bangladesh Health Facility Survey. Int J Health Plann Manage. 2013;28:202–15.
66. Peters DH, Noor AA, Singh LP, Kakar FK, Hansen PM, Burnhama G. A balanced scorecard for health services in Afghanistan. Bull World Health Organ. 2007;85:146–51.

67. Hansen PM, Peters DH, Niayesh H, Singh LP, Dwivedi V, Burnham G. Measuring and managing progress in the establishment of basic health services: the Afghanistan Health Sector Balanced Scorecard. Int J Health Plann Manage. 2008;23:107–17.
68. Parasuraman A, Zeithaml VA, Berry LL. A Conceptual Model of Service Quality and Its Implications for Future Research. J Mark. 1985;49:41. doi:10.2307/1251430.
69. Ladhari R. A review of twenty years of SERVQUAL research. Int J Qual Serv Sci. 2009;1:172–98.
70. Buttle F. SERVQUAL: review, critique, research agenda. Eur J Mark. 1996;30:8–32. doi:10.1108/03090569610105762.
71. Press Ganey Corporation. Press Ganey Associates, Inc. https://www.pressganey.com/products/patient-experience. Accessed 20 Nov 2021.
72. Centers for Medicare & Medicaid. Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS). Agency for Healthcare Research and Quality. https://www.ahrq.gov/cahps/surveys-guidance/hospital/about/adult_hp_survey.html. Accessed 27 Nov 2021.
73. Bin Traiki TA, AlShammari SA, AlAli MN, Aljomah NA, Alhassan NS, Alkhayal KA, et al. Impact of COVID-19 pandemic on patient satisfaction and surgical outcomes: A retrospective and cross sectional study. Ann Med Surg. 2020;58:14–9. doi:10.1016/j.amsu.2020.08.020.

## Tables

Table 1. Sociodemographic and characteristics of the second panel (executive managers)

| Sociodemographic characteristic | Panelists N | % | Sociodemographic characteristic | Panelists N | % |
|---|---|---|---|---|---|
| **Age** | | | **Position** | | |
| 30-39 years | 4 | 30.7 | CMO | 3 | 23.1 |
| 40-49 years | 7 | 53.8 | CFO | 3 | 23.1 |
| 60-69 years | 2 | 15.4 | CEO | 3 | 23.1 |
| **Gender** | | | Managing director | 3 | 23.1 |
| Male | 7 | 53.8 | Operation manager | 1 | 7.7 |
| Female | 6 | 46.2 | **Highest degree** | | |
| **Academic background** | | | Bachelor degree | 8 | 61.5 |
| Medicine | 4 | 30.8 | Master's degree | 5 | 38.5 |
| Management | 4 | 30.8 | **Administrative type** | | |
| Accounting | 3 | 23.1 | Private | 4 | 30.8 |
| Accounting and management | 2 | 15.4 | NGO | 4 | 30.8 |
| **Years of experience** | | | Public | 5 | 38.5 |
| 5-10 years | 1 | 7.6 | | | |
| More than 10 years | 12 | 92.3 | | | |

CMO, Chief Medical Officer; CFO, Chief Financial Officer; CEO, Chief Executive Officer, NGO, Non-Governmental Organization.

Table 2. Charasterestics and sociodemographics of respondents (patients)

| | | Number of patients (N= 740) | % | | | Number of patients (N= 740) | % |
|---|---|---|---|---|---|---|---|
| **Age (years)** | | | | | | | |
| | Less than 20 | 63 | 8.5 | Income (NIS) | Less than 1000 | 195 | 26.4 |
| | 20-29 | 209 | 28.2 | | 1000-2000 | 98 | 13.2 |
| | 30-39 | 208 | 28.1 | | 2001-3000 | 152 | 20.5 |
| | 40-49 | 159 | 21.5 | | 3001-4000 | 140 | 18.9 |
| | 50-59 | 71 | 9.6 | | More than 4000 | 155 | 20.9 |
| | 60-69 | 24 | 3.2 | Insurance type # | | | |
| | More than 70 | 6 | 0.8 | | Public | 492 | 66.5 |
| **Gender** | | | | | Private | 143 | 19.3 |
| | Females | 325 | 43.9 | | UNRWA | 63 | 8.5 |
| | Males | 415 | 56.1 | | No insurance | 109 | 14.7 |
| **Highest degree** | | | | | | | |
| | Elementary | 85 | 11.5 | Number of the current visit | First | 227 | 30.7 |
| | Secondary | 217 | 29.3 | | Second | 187 | 25.3 |
| | Bachelor | 366 | 49.5 | | Third | 91 | 12.3 |
| | Masters | 63 | 8.5 | | Fourth | 54 | 7.3 |
| | PhD | 9 | 1.2 | | Fifth | 181 | 24.5 |
| **Working sector** | | | | Admission status | | | |
| | Public | 175 | 23.6 | | Inpatients | 350 | 47.3 |
| | Private | 183 | 24.7 | | Outpatients | 390 | 52.7 |
| | Free lancer | 156 | 21.1 | Respondent opinion is based on # | | | |
| | Retired | 17 | 2.3 | | Personal experience | 570 | 77 |
| | Unemployed | 209 | 28.2 | | Family experience | 306 | 41.4 |
| | | | | | Friends experience | 96 | 13 |

NIS, New Israeli shekel; UNRWA, The United Nations Relief and Works Agency for Palestine Refugees in the Near East; NGO, Non-Governmental Organization; #, multiple response question.

Table 3. No of patients and hospitals based on hospital charasterestics

| | | Number of patients (N= 740) | % | Number of hospitals (N= 14) | % |
|---|---|---|---|---|---|
| **Administrative Type** | Public | 252 | 34.1 | 5 | 36 |
| | NGO | 277 | 37.4 | 5 | 36 |
| | private | 159 | 21.5 | 3 | 21 |
| | UNRWA | 52 | 7 | 1 | 7 |
| **City** | | | | | |
| | Hebron | 150 | 20.3 | 3 | 21 |
| | Jerusalem | 86 | 11.6 | 1 | 7 |
| | Nablus | 249 | 33.6 | 5 | 36 |
| | Qalqelia | 52 | 7 | 1 | 7 |
| | Ramallah | 151 | 20.4 | 3 | 21 |
| | Tulkarm | 52 | 7 | 1 | 7 |
| **Area** | | | | | |
| | North | 353 | 47.7 | 7 | 50 |
| | Middle | 237 | 32 | 4 | 29 |
| | South | 150 | 20.3 | 3 | 21 |
| | | | | | |
| **Accredited hospital** | Yes | 185 | 25 | 3 | 21 |
| | No | 555 | 75 | 11 | 79 |
| **Size** | | | | | |
| | Small (No. Of beds <80) | 241 | 32.6 | 5 | 36 |
| | Medium (No. Of beds 80-160) | 261 | 35.3 | 5 | 36 |
| | Large (No. Of beds >160) | 238 | 32.2 | 4 | 29 |

UNRWA, The United Nations Relief and Works Agency for Palestine Refugees in the Near East; NGO, Non-Governmental Organization.

Table 4. Exploratory Factor Analysis (EFA).

| Component | Item | Item code | Component/item loadings | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| BSCP ATT | I will recommend this hospital to my family and friends. | SAT3 | .894 | | | | | | | | | |
| | I believe I receive an accurate medical examination at this hospital. | PIN1 | .783 | | | | | | | | | |
| | I will choose this hospital again when I need a medical consultation. | PEN2 | .754 | | | | | | | | | |
| | I believe this hospital offers me better treatment than the other Palestinian hospitals. | PEN3 | .686 | | | | | | | | | |
| | My overall satisfaction with this hospital's performance is high. | SAT1 | .683 | | | | | | | | | |
| | I believe this hospital has a high cure rate. | PEN1 | .651 | | | | | | | | | |
| | I will choose this hospital again when I need a medical consultation. | SAT2 | .579 | | | | | | | | | |
| | I believe the staff at this hospital are competent, knowledgeable, updated, and skilled. | PLE1 | .537 | | | | | | | | | |
| PT EXR | This hospital distributes surveys to assess my satisfaction before discharge. | PCU4 | | .968 | | | | | | | | |
| | This hospital distributes surveys to assess my needs upon arrival to the hospital, admission, or during the stay. | PCU3 | | .755 | | | | | | | | |
| | Separate male/female waiting area are available at this hospital. | PEN9 | | .655 | | | | | | | | |
| | This hospital follows up with me after the discharge. | PLE11 | | .645 | | | | | | | | |
| | My complaints are taken seriously into consideration | PCU5 | | .601 | | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | and solved immediately at this hospital. | | | | | | |
| | I can book an online or a phone appointment at this hospital easily. | PLE7 | .586 | | | | |
| | Staff trained me on infection precaution measures such as hand hygiene, cough etiquette, isolation rational, personal protective equipment, etc. | PLE6 | .560 | | | | |
| SERV EXR | Female doctors are available at this hospital. | PEN8 | | .625 | | | |
| | There are a variety of departments at this hospital. | PIN12 | | .616 | | | |
| | Services at night, vacations, and weekends are available at this hospital. | PIN18 | | .556 | | | |
| | There are a variety of specialties at this hospital. | PIN15 | | .540 | | | |
| PR EXR | I pay a reasonable price for the other medical services (laboratory, radiology, etc.) at this hospital. | PFI2 | | | .959 | | |
| | I pay a reasonable price for the medications at this hospital. | PFI3 | | | .888 | | |
| | I pay a reasonable price for the medical consultation at this hospital. | PFI1 | | | .848 | | |
| BUIL EXR | There is a sufficient number of chairs in the waiting area. | PEN13 | | | | .639 | |
| | The hospital has clean departments, corridors, rooms, bathrooms. | PEN12 | | | | .585 | |
| | The capacity of departments at this hospital including (ER, ICU, waiting room, etc.) is sufficient enough. | PEN14 | | | | .562 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | This hospital has new building infrastructure (walls, ceiling, bathrooms, etc.). | PEN11 | | .519 | | | | | |
| ACC EXR | The accessibility to this hospital is easy by either public transportation or my car. | PEN4 | | | .910 | | | | |
| | The accessibility to this hospital in an emergency is easy. | PEN5 | | | .907 | | | | |
| COMP IMAGE | There is a probability for postoperative bacterial infection at this hospital | PIN6 | | | | .765 | | | |
| | There is a probability for case referral to another hospital | PIN5 | | | | .752 | | | |
| | There is a probability for case readmission at the same hospital | PIN4 | | | | .602 | | | |
| TECH IMAGE | This hospital use technology to link my prescriptions and tests with pharmacy and labs. | PLE9 | | | | | .842 | | |
| | This hospital use technology for saving my records. | PLE10 | | | | | .564 | | |
| INFO EXR | Information provided to me to be used after discharge is sufficient (medication and side effects, health condition, etc.). | PLE4 | | | | | | .708 | |
| HSRP IMAGE | I believe this hospital offers social and volunteering activities to the community. | PEN7 | | | | | | | .601 |
| | I believe this hospital offers exemptions for poor patients. | PEN6 | | | | | | | .566 |
| WT EXR | I wait for a long time before receiving the medical service at this hospital. | PIN9 | | | | | | | |
| No loadings | The services provided to me at this hospital have high quality.>> | PIN16 | | | | | | | |

| | |
|---|---|
| The medical staff at this hospital speaks a simple language with me. | PCU1 |
| The staff at this hospital protects and respect my privacy. | PEN10 |
| The food offered to you at this hospital has high quality. | PIN19 |
| The staff at this hospital are kind, deal with courtesy and respect, and have a good relationship with me and my family. | PCU2 |
| The hospital staff can respond to my inquiries rapidly. | PIN17 |
| I believe this hospital uses the newest technology and devices for diagnosing and treating patients. << | PLE8 |
| Patient counseling services are available at this hospital. | PLE5 |
| I believe the medications prescribed to me at this hospital have good quality and efficacy. | PIN14 |
| The doctors and nurses at this hospital spend sufficient time with me. | PIN8 |
| The hospital staff applies safety standards (gloves, masks, hygiene, etc.). | PIN7 |
| My room is calm and peaceful | PEN15 |
| The medications prescribed to me are available at the hospital's pharmacy. | PIN13 |
| Oral and written information | PLE3 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | provided to me or my family during my hospital experience is sufficient. # | | | | | | | | | | | |
| | Information and guidance provided at admission or the first visit are sufficient. # | PLE2 | | | | | | | | | | |
| Percentage of Variance (%) | | 27.46 | 5.81 | 5.02 | 3.71 | 3.40 | 3.24 | 2.79 | 2.70 | 2.48 | 2.37 | 2. |
| Total variance = 63.29% | | | | | | | | | | | | |
| Eigenvalues | | 14.28 | 3.02 | 2.61 | 1.93 | 1.78 | 1.69 | 1.45 | 1.40 | 1.29 | 1.23 | 1. |

BSCP ATT, patient attitude toward balanced scorecard perspectives; PT EXR, patient experience; SERV EXR, services experience; PR EXR, price experience; BUIL EXR, building experience; ACC EXR, access experience; COMP IMAGE, complications perceived image; TECH IMAGE, technology perceived image; INFO EXR, information experience; HSRP IMAGE, hospital social responsibility perceived image; WT EXR, waiting time experience; #, items were added to INFO EXR construct in CFA; >>, items added to BSCP ATT construct in CFA; <<, items were added to TECH IMAGE construct in CFA.

Table 5. Goodness of fit indices in EFA and CFA and results.

| **EFA** [43,50] | | **CFA** [74] | |
|---|---|---|---|
| **Criteria for good fit** [49,56] | **Measurements** | **Criteria for good fit** | **Measurements** |
| -KMO:<br>.6: low adequacy<br>.7: medium adequacy<br>.8: high adequacy<br>.9: very high adequacy<br>-Bartlett's test *P value* <.05<br>-Inclusion/exclusion criteria for the components:<br>1. Eigenvalues ≥1<br>2. Visual assessment of Catell's scree plot.<br>-Inclusion/exclusion criteria for the items:<br>1. The factor loading ≥.50.<br>2. Factor loadings on the assigned construct ≥all cross-loading of other constructs. | -KMO = .901 (Chi square = 9052.693, degrees of freedom = 1326)<br>-Bartlett's test *P value* <.001<br>-12 components which have Eigenvalues above 1<br>-Cumulative variance = 63.29% | • χ2/df <5 and closer to zero<br>• The *P value* >.05<br>• GFI<br>- CFI<br>- TLI<br>GFI, CFI, and TLI close to .95<br>-RMSEA <.06<br>-SRMR ≤.08 | χ2/df = 1.58<br>*P value* <0.001<br>GFI = .901<br>CFI = .953<br>TLI = .944<br>RMSEA = .039<br>SRMR = .0439 |

EFA, Exploratory Factor Analysis; CFA, Confirmatory Factor Analysis; KMO, Kaiser–Meyer–Olkin; χ2/df, minimum discrepancy divided by its degrees of freedom; GFI, the Goodness-of-Fit Index; CFI, Comparative Fit Index; TLI, Tucker–Lewis's Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual.

Table 6. Constructs IIC, CTIC, and CR.

| Construct | IIC (Min.-Max.) | CTIC (Min.-Max.) | CR | N of Items (Total = 34) |
|---|---|---|---|---|
| COMP IMAGE | .395-.411 | .474-.486 | .664 | 3 |
| TECH IMAGE | .390-.594 | .486-.642 | .794 | 3 |
| BSCP ATT | .328-.641 | .505-.735 | .861 | 9 |
| INFO EXR | .389-.531 | .501-.609 | .750 | 3 |
| PR EXR | .509-.725[>>] | .596-.760[>>] | .948 | 3 |
| PT EXR | .413-.678 | .552-.736 | .841 | 5 |
| ACC EXR | .853 | .853 | .906 | 2 |
| SERV EXR | .360 | .360 | .502 | 2 |
| BUILENV EXR | .412 | .412 | .643 | 2 |
| BUILCAP EXR | .527 | .527 | .721 | 2 |

COMP IMAGE, complications perceived image; TECH IMAGE, technology perceived image; BSCP ATT, patient attitude toward balanced scorecard perspectives; INFO EXR, information experience; PR EXR, price experience; PT EXR, patient experience;; ACC EXR, access experience; SERV EXR, services experience; BUILENV EXR, building environment experience; BUILCAP EXR, building capacity experience; IIC, Inter-Item Correlation; CITC, Corrected Item Total Correlation; CR, Composite reliability; [>>], was calculated only for patients who pay at the evaluated hospitals.

Table 7. Convergent, discriminant, and divergent validity for the independent constructs

| Construct | AVE | INFO EXR | PR EXR | PT EXR | ACC EXR | SERV EXR | BUILENV EXR | BUILCAP EXR |
|---|---|---|---|---|---|---|---|---|
| INFO EXR | .501 | **.708** | | | | | | |
| PR EXR | .858 | *.084*[*] | **.926** | | | | | |
| PT EXR | .515 | *.507*[**] | *.095*[*] | **.718** | | | | |
| ACC EXR | .828 | *.121*[**] | *-.005* | *.053* | **.910** | | | |
| SERV EXR | .337 | *.341*[**] | *.002* | *.242*[**] | *.164*[**] | **.581** | | |
| BUILENV EXR | .477 | *.302*[**] | *-.006* | *.336*[**] | *.110*[**] | *.209*[**] | **.691** | |
| BUILCAP EXR | .564 | *.288*[**] | *.016* | *.366*[**] | *.164*[**] | *.238*[**] | *.394*[**] | **.751** |

PT EXR, patient experience; INFO EXR, information experience; PR EXR, price experience; COMM EXR, communication experience; ACC EXR, access experience; BUILCAP EXR, building capacity experience; TECH EXR, technology experience; DEPV EXR, departments variety experience, SERV EXR, services; WT EXR, waiting time experience; BUILENV EXR, building environment experience; [*], $P < 0.05$; [**], $P < 0.01$; AVE, Average Variance Extracted calculated by the average square of loadings at each construct and used to evaluate the convergent validity; **Bold**, Square Roots of the Average Variance Extracted; *Italic*, Spearman correlations between independent constructs, both are used to evaluate discriminant validity; [*], $P < 0.05$; [**], $P < 0.01$.

Table 8. Convergent, discriminant, and divergent validity for the dependent constructs

| Construct | AVE | BSCP ATT | TECH IMAGE | COMP IMAGE |
|---|---|---|---|---|
| BSCP ATT | .413 | **.643** | | |
| TECH IMAGE | .564 | *.397*[**] | **.751** | |
| COMP IMAGE | .400 | *.216*[**] | *.156*[**] | **.633** |

BSCP PI, BSC perspective perceived image; HOSP PMEDI, hospital perceived medical image, HOSP PSOCIALI, hospital perceived social image; [*], $P < 0.05$; [**], $P < 0.01$; AVE, Average Variance Extracted calculated by the average square of loadings at each construct and used to evaluate the convergent validity; **Bold**, Square Roots of the Average Variance Extracted; *Italic*, Spearman correlations between independent constructs, both are used to evaluate discriminant validity; [*], $P < 0.05$; [**], $P < 0.01$.

## Figures

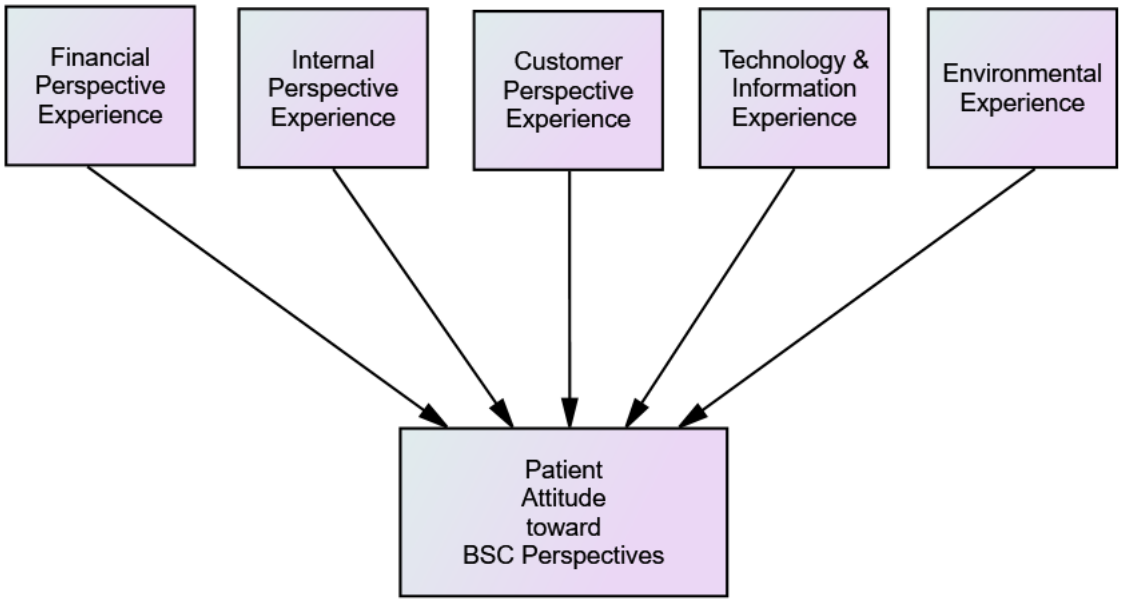


**Figure 1**

BSC-PATIENT conceptual model.

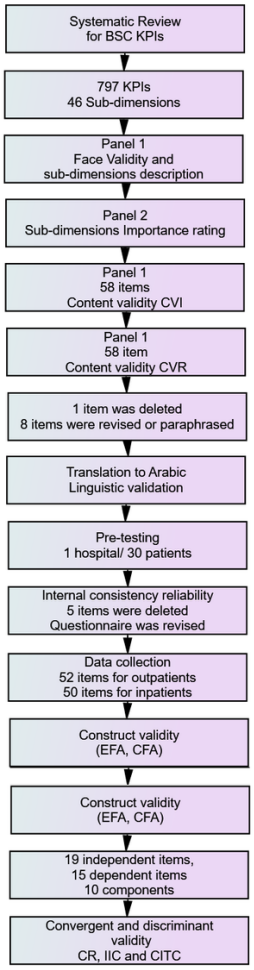


**Figure 2**

Flow chart for BSC-PATIENT development and psychometric validation.

BSC KPI, Balanced Scorecard Key Performance Indicators; CVI, Content Validity Index; CVR, Content Validity Ratio; CR, Composite Reliability; IIC, Inter-Item Correlation; CITC, Corrected Item-Total Correlation.

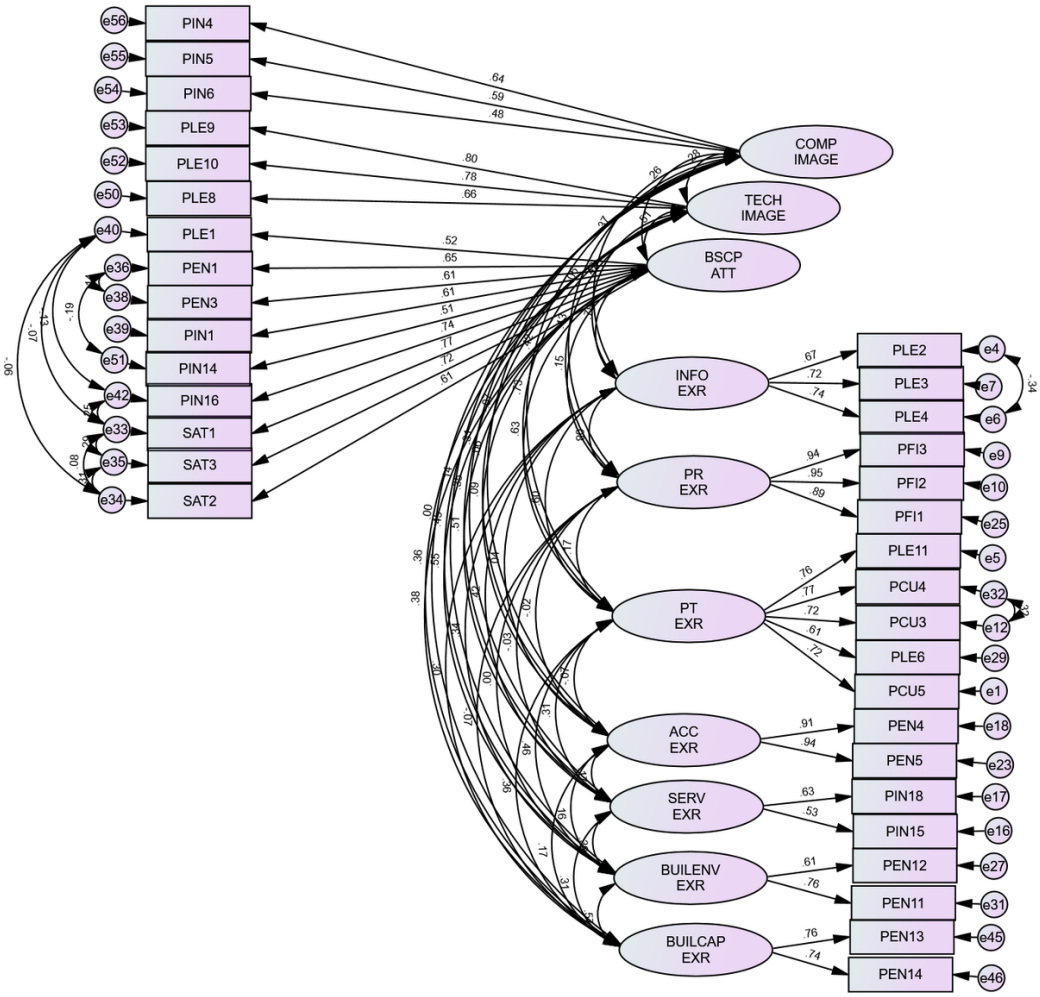


**Figure 3**

Confirmatory Factor Analysis (CFA). Independent items on the right side, and dependent items on the left side.

COMP IMAGE, complications perceived image; TECH IMAGE, technology perceived image; BSCP ATT, patient attitude toward balanced scorecard perspectives; INFO EXR, information experience; PR EXR, price experience; PT EXR, patient experience; ACC EXR, access experience; SERV EXR, services experience; BUILENV EXR, building environment experience; BUILCAP EXR, building capacity experience.