

Recherche et Applications en Marketing

<http://ram.sagepub.com/>

SERVQUAL: Une échelle multi-items de mesure des perceptions de la qualité de service par les consommateurs

A. Parasuraman, Valarie A. Zeithaml and Leonard L. Berry

Recherche et Applications en Marketing 1990 5: 19

DOI: 10.1177/076737019000500102

The online version of this article can be found at:

<http://ram.sagepub.com/content/5/1/19>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association Française du Marketing](#)

Additional services and information for *Recherche et Applications en Marketing* can be found at:

Email Alerts: <http://ram.sagepub.com/cgi/alerts>

Subscriptions: <http://ram.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://ram.sagepub.com/content/5/1/19.refs.html>

>> [Version of Record](#) - Mar 1, 1990

[What is This?](#)

SERVQUAL : une échelle multi-items de mesure des perceptions de la qualité de service par les consommateurs*

A. PARASURAMAN

*Professeur de distribution et d'études marketing
Université A & M du Texas
College Station, Texas*

Valarie A. ZEITHAML

*Professeur associé de marketing
Université Duke
Raleigh Durham, Caroline du Nord*

Leonard L. BERRY

*Professeur de distribution et d'études marketing
Université A & M du Texas
College Station, Texas*

RÉSUMÉ. — Cet article décrit le développement d'un instrument composé de 22 items (appelé SERVQUAL) destiné à évaluer les perceptions des consommateurs concernant la qualité de service offerte par les entreprises de service et de commerce de détail. Après avoir présenté la conceptualisation et l'opérationnalisation du concept de qualité de service, les auteurs décrivent les procédures utilisées pour construire et mettre au point une échelle multi-items de mesure de ce construit. À partir d'analyses de données faites sur quatre échantillons indépendants, l'article décrit ensuite les propriétés de l'échelle en termes de structure factorielle, de fiabilité et de validité. Les auteurs concluent par une présentation des applications potentielles de l'instrument.

L'intensification de la concurrence et la rapide dérégulation de certains marchés ont conduit de nombreux domaines des services et du commerce de détail à rechercher des voies efficaces de différenciation. L'une des stratégies associées au succès de ces secteurs consiste à offrir une haute qualité de service (Rudie et Wansley, 1985; Thompson, DeSouza et Gale, 1985). Proposer une qualité de service supérieure à celle de ses concurrents

* Traduit de A. Parasuraman, Valarie A. Zeithaml and Leonard L. Berry (1988), SERVQUAL : A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality, *Journal of Retailing*, 64 : 1, p. 12-40.

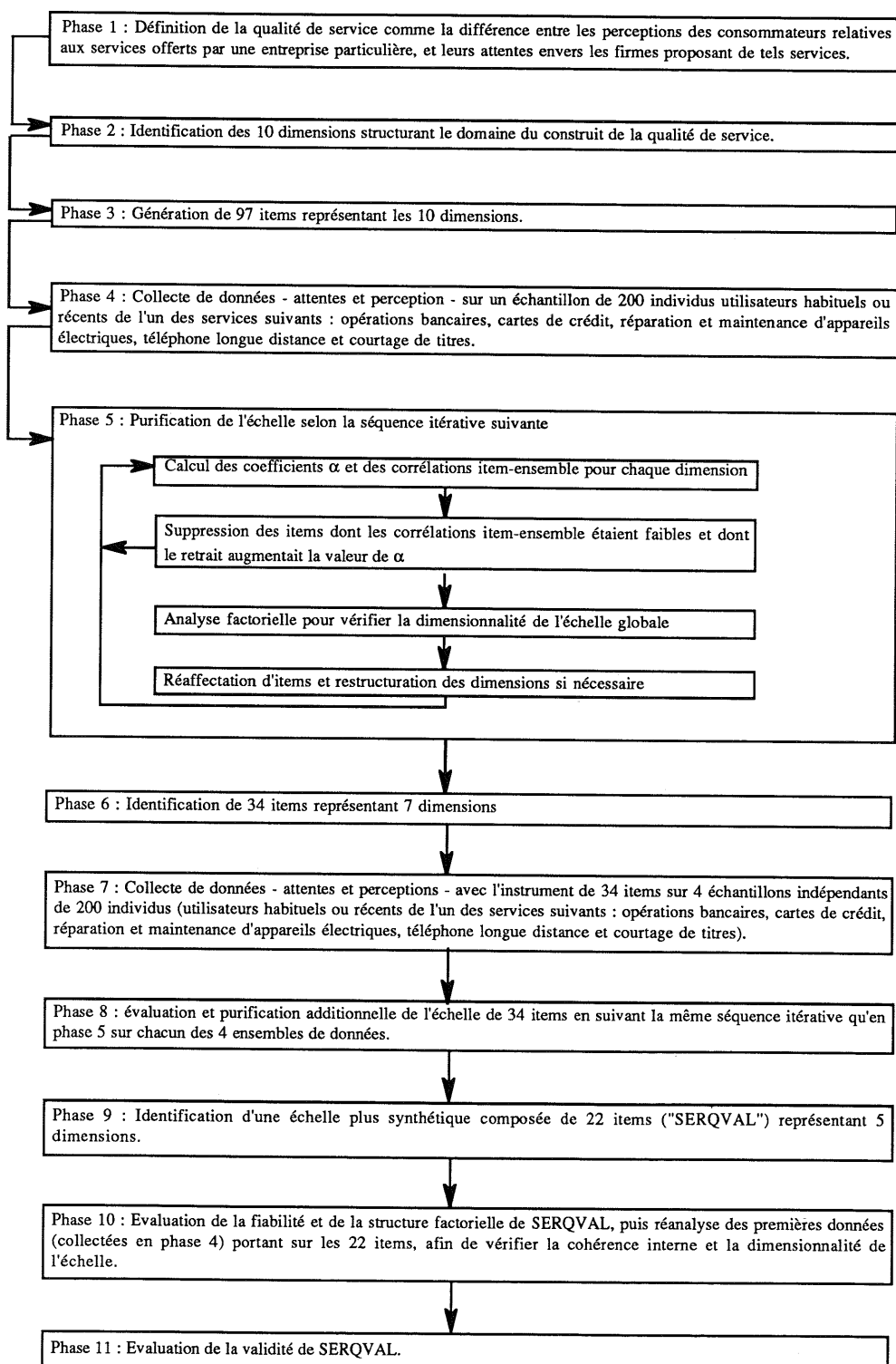


FIG. 1. — Résumé des étapes suivies pour développer l'échelle de mesure de la qualité de service

apparaît ainsi comme un facteur clé de succès, voire même de survie, dans ces domaines d'activité pour les années 1980 et suivantes.

Contrairement à la qualité des biens que l'on peut mesurer objectivement par des indicateurs comme leur durabilité ou le nombre de défauts de fabrication (Crosby, 1979; Garvin, 1983), la qualité de service représente un construit abstrait et diffus en raison des trois caractéristiques uniques associées aux services : intangibilité, hétérogénéité et inséparabilité de la production et de la consommation (Parasuraman, Zeithaml et Berry, 1985). En l'absence de mesures objectives, l'approche qu'il convient d'adopter pour évaluer la qualité des services offerts par une entreprise consiste à mesurer les perceptions de cette qualité par les consommateurs. Jusqu'à présent, néanmoins, on ne dispose d'aucune norme quantitative pour évaluer ces perceptions.

Cet article répond à un double objectif : 1) décrire le développement d'une échelle multi-items de mesure de la qualité de service (SERVQUAL), et 2) présenter ses propriétés ainsi que ses applications potentielles. Les étapes fondamentales suivies pour construire l'échelle sont très similaires aux procédures suggérées par le paradigme de Churchill (1979) en vue du développement de meilleures mesures des construits marketing. La figure 1 présente une vue d'ensemble de ces étapes.

Cet article se compose de cinq sections : la première délimite le domaine du construit de la qualité de service et décrit la génération des items de l'échelle (étapes 1, 2 et 3 de la figure 1). La deuxième section présente les procédures de collecte de données et de purification de l'échelle (étapes 4 à 9), tandis que la troisième apporte les éléments d'évaluation de la fiabilité et de la structure factorielle de l'échelle (étape 10). La section suivante concerne la validité de l'échelle (étape 11). La dernière partie enfin s'attache à proposer des applications potentielles de cet outil.

LE DOMAINE DU CONSTRUIT « QUALITÉ DE SERVICE »

Déplorant l'inadaptation des procédures de mesures utilisées dans la discipline du marketing, Jacoby (1978) déclarait :

« Nos mesures sont dans de nombreux cas développées à la fantaisie des chercheurs qui accordent un intérêt mineur à la question de savoir si oui ou non elles sont reliées à des fondements conceptuels explicites concernant les phénomènes ou la variable étudiés. La plupart du temps, nos concepts n'ont aucune identité hormis celle que leur confèrent l'instrument ou les procédures utilisées pour les mesurer » (p. 92).

D'autres chercheurs (par ex. Churchill, 1979; Peter, 1981) ont aussi souligné la nécessité de faire précéder le développement d'échelles par une spécification conceptuelle saine du construit à mesurer, le cadre conceptuel devant servir de point d'ancrage aux outils de mesure. En ce qui concerne l'échelle SERVQUAL, le cadre conceptuel est issu d'une part des travaux d'une poignée de chercheurs qui ont étudié la signification de la qualité de service

(Sasser, Olsen et Wickoff, 1978; Gronroos, 1982; Lehtinen et Lehtinen, 1982), d'autre part d'une recherche qualitative globale qui a permis de définir la qualité de service et de mettre en évidence les dimensions sur lesquelles les consommateurs la percevaient et l'évaluaient (Parasuraman, Zeithaml et Berry, 1985).

Conceptualisation de la qualité de service

Le construit de qualité tel qu'il est conceptualisé dans la littérature concernant les services et mesuré par SERVQUAL — l'échelle dont il est question dans cet article — concerne la qualité perçue. La qualité perçue représente le jugement du consommateur concernant le degré d'excellence ou de supériorité attribué à une entité (Zeithaml, 1987). Elle se démarque de la qualité objective (telle que l'ont définie par exemple Garvin, 1983, et Hjorth-Anderson, 1984), il s'agit d'une forme d'attitude, liée mais non équivalente à la satisfaction, qui résulte d'une comparaison entre des attentes et des perceptions relatives à la performance.

Qualité perçue versus qualité objective

Les chercheurs ont mis l'accent sur la différence entre qualité perçue et objective (Garvin, 1983; Dodds et Monroe, 1984; Holbrook et Corfman, 1985; Jacoby et Olson, 1985; Zeithaml, 1987). Par exemple, Holbrook et Corfman (1985) remarquent que les consommateurs n'emploient pas le terme de qualité dans le même sens que les chercheurs ou les praticiens du marketing qui la définissent de manière conceptuelle. La signification conceptuelle distingue la qualité mécaniste de la qualité humaniste : « [la qualité] mécaniste concerne un aspect ou une caractéristique d'ordre objectif d'une chose ou d'un événement; [la qualité] humaniste s'applique à la réaction subjective des individus face à des objets, et par là même représente un phénomène d'ordre relatif qui diffère en fonction des personnes émettant le jugement en question » (Holbrook et Corfman, 1985, p. 33). Garvin (1983) présente cinq approches de la définition de la qualité, parmi lesquelles deux se rapportent à la qualité objective (celles fondées sur le produit et sur la production) et une autre à la qualité perçue (celle fondée sur l'utilisateur).

Qualité en tant qu'attitude

Olshavsky (1985) considère la qualité comme une forme de l'évaluation globale d'un produit, semblable sous de nombreux aspects à l'attitude. Holbrook partage cette opinion quand il suggère que la qualité joue le rôle d'un jugement de valeur relativement global. La recherche exploratoire conduite par Parasuraman, Zeithaml et Berry (1985) confirme que la notion de qualité de service traduit une évaluation globale similaire à l'attitude;

ces chercheurs ont mené au total douze entretiens en focus groupes constitués de consommateurs habituels ou récents de quatre différents services — services bancaires, cartes de crédit, courtage de titres et réparation et maintenance de petits appareils électriques. Les discussions étaient centrées sur des questions de signification de la qualité dans ces contextes de services, de caractéristiques que le service et les prestataires devraient posséder pour diffuser une image de haute qualité, et des critères utilisés par les consommateurs pour évaluer la qualité de service. La comparaison des résultats entre les différents focus groupes a mis en évidence le fait que, quel que soit le type de service, les consommateurs utilisaient globalement les mêmes critères généraux pour formuler un jugement d'évaluation de la qualité de service.

Qualité versus satisfaction

Oliver (1981) résume les développements actuels concernant la satisfaction dans la définition suivante : « [la satisfaction traduit un] état psychologique synthétique qui se produit quand une émotion liée à des attentes non réalisées est associée avec les sentiments premiers du consommateur envers une expérience de consommation » (p. 27). Cette définition comme d'autres (par exemple celles de Howard et Sheth, 1969, et de Hunt, 1979), ainsi que la plupart des mesures de la satisfaction concernent une transaction spécifique. Oliver (1981) résume la nature de la satisfaction relative à une transaction spécifique, et la différencie de l'attitude de la façon suivante :

« L'attitude traduit l'orientation affective relativement stable d'un consommateur vis-à-vis d'un produit, d'un point de vente ou d'un processus (par exemple le service consommateur), tandis que la satisfaction traduit la réaction émotionnelle faisant suite à une expérience non confirmatoire qui agit au niveau de l'attitude de base et qui est spécifique à l'acte de consommation. C'est pourquoi l'attitude est mesurée en termes plus généraux relatifs au produit ou au point de vente, et est moins situationnelle » (p. 42).

La distinction entre qualité de service et satisfaction est cohérente avec celle qui existe entre attitude et satisfaction : la qualité de service perçue représente un jugement global, ou attitude, qui concerne la supériorité du service, tandis que la satisfaction est associée à une transaction spécifique. En effet, les répondants des douze focus groupes de l'étude exploratoire de Parasuraman, Zeithaml et Berry (1985) ont donné plusieurs descriptions de situations où ils étaient satisfaits d'un service particulier, mais ne percevaient pas le service de l'entreprise comme étant de haute qualité. De cette manière, les deux construits sont reliés au sens où des événements agissant sur la satisfaction se traduisent avec le temps en perception au plan de la qualité de service. En reprenant les termes d'Oliver (1981) : « La satisfaction se fonde rapidement dans l'attitude globale d'un individu vis-à-vis de l'achat d'un produit. »

Attentes comparées aux perceptions

Les contributions de Sasser, Olsen et Wickoff (1978), Gronroos (1982), Lehtinen et Lehtinen (1982) et les entretiens de groupe de Parasuraman, Zeithaml et Berry (1985) montrent de façon très claire le fait que la qualité de service, telle qu'elle est perçue par les consommateurs, résulte d'une comparaison entre ce qu'ils considèrent devoir être le service offert par la firme (*i.e.* leurs attentes) et leurs perceptions de la performance des prestataires de ce service. Ainsi, la qualité de service perçue est considérée comme la différence entre attentes et perceptions des consommateurs vue sous l'angle de son amplitude et de sa direction.

Le terme d'« attentes » utilisé dans la littérature sur la qualité de service diffère de son emploi dans la littérature sur la satisfaction du consommateur. Plus précisément, la littérature sur la satisfaction considère les attentes comme des *prédictions* émises par les consommateurs au sujet de ce qui est susceptible de se produire lors d'une transaction ou d'un échange imminents. Par exemple, selon Oliver (1981), « on s'accorde généralement sur le fait que les attentes sont des probabilités d'occurrence d'événements positifs ou négatifs définies par le consommateur, si celui-ci adopte un comportement donné » (p. 33). Par contraste, la littérature sur la qualité de service considère les attentes comme des désirs ou des besoins des consommateurs, *i.e.* ce qu'ils ressentent comme *devant être* fourni par un prestataire plutôt que ce qui *pourrait l'être*.

Dimensions de la qualité de service

La recherche exploratoire de Parasuraman, Zeithaml et Berry (1985) a montré que les critères utilisés par les consommateurs lors de l'évaluation de la qualité de service s'articulaient autour de 10 dimensions potentielles se recoupant. Ces dimensions sont les suivantes : éléments tangibles, fiabilité, serviabilité, communication, crédibilité, sécurité, compétence, courtoisie, compréhension/connaissance du consommateur et accessibilité (le lecteur trouvera une description de ces dimensions à la page 47 de l'article en question). Ces 10 dimensions et leurs descriptions ont servi à la structuration fondamentale du domaine de la « qualité de service » à partir duquel les items de l'échelle SERVQUAL ont été développés.

La génération des items de l'échelle

Les items représentant les différentes facettes des 10 dimensions de la qualité de service ont été générés afin de constituer un premier jeu d'items de l'instrument SERVQUAL. Ce processus a conduit à la génération de 97 items (approximativement 10 items par dimension). Chaque item a été refondu en deux énoncés — l'un mesurant les attentes au sujet des entreprises en général au sein de la catégorie de services étudiée, l'autre mesurant les perceptions au sujet d'entreprises particulières dont la qualité de service était évaluée.

Pour suivre les recommandations de Churchill (1979) au sujet des procédures de construction d'échelles, environ la moitié des paires d'énoncés a été formulée positivement et l'autre moitié négativement. Chaque énoncé a été associé avec une échelle à 7 points allant de « tout à fait d'accord » (7) à « tout à fait en désaccord » (1), sans attribuer d'intitulé verbal pour les points intermédiaires de 2 à 6 (les valeurs de l'échelle ont été inversées avant l'analyse de données pour les énoncés à formulation négative). Les énoncés concernant les attentes ont été regroupés ensemble pour former la première moitié de l'instrument, les énoncés de perceptions correspondants formant la seconde partie. Une version abrégée de l'instrument contenant un ensemble d'énoncés relatifs aux attentes (nommé E), un ensemble correspondant d'énoncés relatifs aux perceptions (nommé P), ainsi que les instructions pour y répondre, se trouve en annexe à cet article. Les énoncés formulés négativement sont identifiés par un signe moins entre parenthèses.

COLLECTES DE DONNÉES ET PURIFICATION DE L'ÉCHELLE

L'instrument composé des 97 items a fait l'objet de deux étapes de collecte de données et de purification. La première d'entre elles avait pour objet principal : 1) de condenser l'instrument en ne retenant que les items discriminant correctement les répondants ayant des perceptions différentes de la qualité au sujet d'entreprises de plusieurs catégories de services, et 2) d'examiner le caractère dimensionnel de l'échelle et d'établir les fiabilités respectives de ses composantes. La seconde étape, de nature principalement confirmatoire, avait pour but de réévaluer la dimensionalité de l'échelle condensée ainsi que sa fiabilité en analysant des données nouvelles à partir de quatre échantillons indépendants. Des améliorations supplémentaires ont été effectuées lors de cette phase.

Collecte de données, première étape

Les données utilisées pour la purification initiale de l'instrument composé de 97 items ont été collectées sur un échantillon de 200 adultes (âgés de 25 ans au moins) choisis selon la méthode des quotas et sélectionnés par une entreprise d'études marketing dans un espace commerçant d'une grande zone métropolitaine du sud-ouest des Etats-Unis. La taille de l'échantillon (200) a été choisie parce que d'autres constructeurs d'échelles dans le domaine du marketing ont utilisé des tailles d'échantillon similaires lors de l'étape initiale de purification d'instruments contenant un nombre identique d'items à celui de notre échelle (par ex. Churchill, Ford et Walker, 1974; Saxe et Weitz, 1982). L'échantillon était composé globalement d'autant d'hommes que de femmes. De plus, les répondants ont été répartis en cinq catégories de services différents : réparation et maintenance de petits appareils électriques, services bancaires, téléphone longue distance,

courtage de titres et cartes de crédit. Ces catégories ont été choisies pour représenter une large part des services variant selon les dimensions clés utilisées par Lovelock (1980, 1983) dans sa classification des services. Pour chaque catégorie de services, on a établi un quota de 40 utilisateurs récents. Le critère filtre pour participer à l'enquête était d'avoir utilisé le service en question au cours des trois derniers mois.

Les individus sélectionnés ont répondu eux-mêmes au questionnaire en deux parties consistant en 97 énoncés relatifs à leurs attentes suivis de 97 énoncés relatifs à leurs perceptions. Dans la première partie, les répondants devaient indiquer le niveau de service qui devrait être proposé par les entreprises de la catégorie en question. Dans la seconde partie, on leur demandait d'abord de nommer une entreprise (de la catégorie de service) à laquelle ils avaient eu recours et qui leur était la plus familière. Ensuite, les individus devaient exprimer leurs perceptions de ce prestataire.

Purification de l'échelle, première étape

L'instrument composé des 97 items a été affiné par analyse sur les données regroupées (*i.e.* sur les données des cinq catégories de services considérées ensemble). Le regroupement des données correspond à un choix délibéré et approprié parce que l'objectif essentiel de cette étape de la recherche était de développer une version plus concise de l'instrument qui soit fiable et valide pour l'évaluation de la qualité d'une variété de secteurs des services. Autrement dit, le but était d'obtenir une échelle pouvant être appliquée de manière générale.

La purification de l'instrument commence par le calcul du coefficient alpha de Cronbach (Cronbach, 1951) ce qui concorde avec les recommandations de Churchill (1979). En raison de la multidimensionalité du construit « qualité de service », le coefficient alpha a été calculé séparément pour les dix dimensions afin de déterminer dans quelle mesure les items composant les dimensions partageaient une notion commune. Les données brutes utilisées pour le calcul du coefficient alpha (et pour les analyses suivantes) ont été mises sous la forme de différences de scores. Plus précisément, on a calculé pour chaque item un score Q (représentant la qualité perçue sur cet item) défini par $Q = P - E$, où P et E sont les réponses respectives aux énoncés correspondants de perception et d'attentes. L'idée d'utiliser des différences de scores pour purifier une échelle n'est pas récente. En effet, cette approche a été employée dans le développement d'échelles de mesure de construits comme le conflit de rôles (Ford, Walker et Churchill, 1975).

Les valeurs du coefficient alpha étaient comprises entre 0,55 et 0,78 pour l'ensemble des dimensions, ce qui suggérait de supprimer certains items de chaque dimension pour améliorer la valeur des coefficients. Le critère utilisé pour décider de la suppression d'un item est la corrélation corrigée item-ensemble, soit la corrélation entre le score sur l'item et la somme des scores sur *tous les autres* items composant la dimension de l'item en question. Les corrélations corrigées item-ensemble ont été représentées graphiquement par ordre décroissant pour chaque dimension. Les items possédant de faibles

corrélations et/ou ceux dont les corrélations conduisaient à une brusque chute dans la forme de la représentation ont été supprimés. Le calcul de nouvelles valeurs de alpha sur l'ensemble réduit d'énoncés et l'examen des nouvelles corrélations corrigées item-ensemble ont conduit à de nouvelles suppressions d'items permettant d'améliorer les valeurs de alpha correspondantes. Cette séquence itérative, d'une part de calcul des coefficients alpha et des corrélations corrigées, d'autre part d'élimination d'items, a été répétée plusieurs fois pour aboutir à un ensemble de 54 items dont les valeurs de alpha étaient comprises entre 0,72 et 0,83 sur l'ensemble des 10 dimensions.

La tâche suivante de cette étape de purification de l'échelle consistait à étudier la dimensionalité de l'échelle composée des 54 items. Ceci a été fait par analyse factorielle des scores de différence sur les 54 items. On a utilisé la procédure factorielle en composantes principales (Harman, 1967) et 10 facteurs ont été fixés *a priori*. La rotation orthogonale pour les 10 facteurs n'a pas permis de faire émerger de structure factorielle claire. De nombreux items présentaient de forts *loadings* sur plusieurs facteurs, ce qui implique que les facteurs étaient susceptibles de ne pas être indépendants les uns des autres. De plus on pouvait anticiper un certain degré de chevauchement entre les 10 dimensions initialement identifiées et interprétées (Parasuraman, Zeithaml et Berry, 1985). C'est pourquoi la solution *a priori* des 10 facteurs a été soumise à une rotation oblique (en utilisant la procédure Oblimin de SPSS-X) afin d'autoriser les intercorrélations entre dimensions et d'induire une interprétation simple.

La rotation oblique a produit une matrice des *loadings* factoriels aisément interprétable. Cependant, plusieurs items possédaient encore de forts *loadings* sur plus d'un facteur. Quand on supprimait de tels items de la matrice des *loadings* factoriels, plusieurs facteurs n'avaient aucun sens puisqu'ils avaient des corrélations avec les autres items proches de zéro, ce qui suggérerait une réduction du caractère dimensionnel supposé du domaine de la qualité de service. De plus, les plus forts *loadings* des quelques items restants concernaient des facteurs auxquels ils n'avaient pas été assignés à l'origine. En d'autres termes, les *loadings* factoriels suggéraient une réaffectation de certains items.

La suppression de certains items (et la réduction conséquente du nombre total de facteurs ou de groupes d'items), ainsi que la réaffectation de certains autres, ont conduit d'une part à calculer à nouveau les coefficients alpha et les corrélations item-ensemble, d'autre part à réexaminer la structure factorielle sur l'ensemble réduit d'items. Cette séquence itérative (étape 5 de la figure 1) a été répétée plusieurs fois et a permis d'obtenir un groupe final de 34 items représentant sept dimensions distinctes. Les valeurs de alpha et les *loadings* factoriels concernant l'instrument à 34 items sont présentés dans le tableau 1.

Comme on peut le constater, cinq des 10 dimensions originelles — éléments tangibles, fiabilité, serviabilité, compréhension/connaissance des consommateurs et accessibilité — sont restées distinctes. Les cinq autres dimensions — communication, crédibilité, sécurité, compétence et courtoisie — se sont fondues en deux dimensions (D₄ et D₅), chacune consistant

TABLEAU 1. — *Synthèse des résultats de la première étape de purification de l'échelle*

Dimensions	Intitulé	Coeffi- cients de fiabilité (Alphas)	Nombre d'items	Loadings factoriels des items sur les dimensions auxquelles ils se rattachent ^a
Eléments tangibles	D1	.72	4	63 75 62 47
Fiabilité	D2	.83	5	74 56 73 71 47
Serviabilité	D3	.84	5	60 73 59 76 66
Communication Crédibilité Sécurité Compétence Courtoisie	D4	.79	4	35 53 66 56
	D5	.85	7	41 62 47 50 75 52 54
Compréhension/ connaissance du consommateur	D6	.85	4	80 76 62 77
Accessibilité	D7	.78	5	57 50 75 52 71
Fiabilité de la combinaison linéaire (fiabilité de l'échelle totale)		.94		

a. Les chiffres représentent les *loadings* multipliés par 100. Les *loadings* des items sur les dimensions auxquelles ils n'appartiennent pas étaient inférieurs à 0,3. Le pourcentage de variance extrait par les 7 facteurs était de 61,7 %.

en items de plusieurs des cinq dimensions de départ. La corrélation moyenne deux à deux entre les sept facteurs après rotation oblique était de 0,27. Cette valeur relativement faible ainsi que les *loadings* factoriels relativement élevés présentés dans le tableau 1 suggéraient que la qualité de service pouvait avoir sept dimensions relativement uniques.

Les fortes valeurs de alpha indiquaient une bonne cohérence interne entre les items à l'intérieur de chaque dimension. De plus, la fiabilité combinée sur l'échelle composée des 34 items, calculée en utilisant la formule de la fiabilité pour des combinaisons linéaires (Nunnally, 1978) était très forte (0,74). C'est pourquoi on a considéré l'instrument de 34 items comme prêt à être testé avec des données obtenues sur de nouveaux échantillons.

Collecte de données, deuxième étape

Afin de poursuivre l'évaluation des propriétés psychométriques de l'échelle à 34 items, des données portant sur la qualité de service de quatre entreprises de notoriété nationale ont été collectées : une banque, une entreprise de cartes de crédit, une entreprise de réparation et de maintenance de petits appareils électriques et une entreprise de téléphone longue distance. Pour chacune d'entre elles, un échantillon indépendant de 200 consommateurs âgés de 25 ans au moins a été sélectionné par une entreprise d'études marketing dans un espace commerçant d'une importante zone métropolitaine de l'est des Etats-Unis. Le critère filtre pour participer à l'enquête était d'avoir eu recours aux services des entreprises en question dans les trois derniers mois. Chaque échantillon contenait autant d'hommes que de femmes. Enfin, tout comme dans la première phase de collecte de données, les individus ont répondu au questionnaire eux-mêmes.

Purification de l'échelle, deuxième étape

L'un des principaux buts de cette étape était d'évaluer la robustesse de l'échelle composée de 34 items pour mesurer la qualité de service des quatre entreprises. C'est pourquoi les données de chacun des quatre échantillons ont été analysées séparément afin d'obtenir les valeurs de alpha (et les corrélations corrigées item-ensemble). A la suite d'une rotation oblique pré-cisant une solution de sept facteurs, on a obtenu une matrice des *loadings* factoriels. Les résultats pour chacun des échantillons ont facilité la validation croisée des résultats issus des autres échantillons.

Les résultats des quatre ensembles d'analyses étaient relativement cohérents, mais présentaient des différences avec les résultats du tableau 1. Deux différences en particulier sont apparues. Tout d'abord, les corrélations corrigées item-ensemble pour plusieurs items (notamment parmi les items composant les dimensions D₄ et D₇ du tableau 1), ainsi que les alphas des dimensions correspondantes étaient plus faibles que ceux obtenus lors de la première étape. Ensuite, les matrices de *loadings* factoriels obtenues à partir des quatre analyses montraient des chevauchements nettement plus

importants entre les dimensions D₄ et D₅ d'une part, et D₆ et D₇ d'autre part. En raison de l'existence de ces différences de façon cohérente pour les quatre échantillons et ensembles de données indépendants, une étape supplémentaire de purification de l'échelle composée de 34 items a été jugée nécessaire.

Quelques items ayant des corrélations items-ensemble relativement faibles ont été supprimés. De plus, ainsi que le suggéraient les analyses factorielles, les items restants en D₄ et D₅, aussi bien que ceux en D₆ et D₇ ont été combinés pour former deux dimensions distinctes. Pour chaque échantillon, on a procédé au calcul des nouvelles valeurs de alpha et aux analyses factorielles (pour l'extraction de cinq facteurs par rotation oblique) sur l'ensemble réduit à cinq dimensions. L'étude des résultats a été menée selon la même séquence itérative que celle de l'étape 5 de la figure 1. Cette procédure a permis d'obtenir une échelle purifiée (« SERVQUAL ») composée de 22 items répartis en cinq dimensions (D₁, D₂, D₃, une combinaison de D₄ et D₅, et une combinaison de D₆ et D₇). Les énoncés relatifs aux attentes et aux perceptions dans l'instrument final SERVQUAL sont présentés en annexe.

L'interprétation du contenu des items finaux composant les cinq dimensions de SERVQUAL (dont trois dimensions d'origine et deux dimensions combinées) a suggéré les intitulés et les définitions concises des dimensions suivants :

- Eléments tangibles : installations physiques, équipement et apparence du personnel;
- Fiabilité : capacité à réaliser le service promis en toute confiance et de manière précise;
- Serviabilité : bonne volonté pour répondre aux consommateurs et offrir un service prompt;
- Assurance : compétence et courtoisie des employés ainsi que leur capacité à inspirer confiance;
- Empathie : prise en considération, attention individualisée que l'entreprise accorde à ses consommateurs.

Les deux dernières dimensions (assurance et empathie) contiennent des items représentant les sept dimensions originelles — communication, crédibilité, sécurité, compétence, courtoisie, compréhension/connaissance des consommateurs et accessibilité — qui ne sont pas restées distinctes après les deux étapes de purification. C'est pourquoi, si SERVQUAL se compose seulement de cinq dimensions distinctes, celles-ci capturent pourtant les facettes des 10 dimensions conceptualisées à l'origine.

SERVQUAL : FIABILITÉ ET STRUCTURE FACTORIELLE

Le tableau 2 montre les coefficients de fiabilité des composantes et de la totalité de l'instrument SERVQUAL pour chacun des quatre échantillons. Les alpha sont élevés pour les quatre échantillons, à l'exception possible d'un

TABLEAU 2. — *Cohérences internes des cinq dimensions de qualité de service après la deuxième étape de purification de l'échelle*

Dimensions	Intitulé	Nombre d'items	Coefficients de fiabilité (Alphas) ^a				Items ^b
			B	CC	R & M	LDT	
Eléments tangibles	F ₁	4	.52	.62	.64	.64	Q ₁ Q ₂ Q ₃ Q ₄
Fiabilité	F ₂	5	.80	.78	.84	.74	Q ₅ Q ₆ Q ₇ Q ₈ Q ₉
Serviabilité	F ₃	4	.72	.69	.76	.70	Q ₁₀ Q ₁₁ Q ₁₂ Q ₁₃
Assurance	F ₄	4	.84	.80	.87	.84	Q ₁₄ Q ₁₅ Q ₁₆ Q ₁₇
Empathie	F ₅	5	.71	.80	.72	.76	Q ₁₈ Q ₁₉ Q ₂₀ Q ₂₁ Q ₂₂
Fiabilité de la combinaison linéaire (fiabilité de l'échelle totale)			.87	.89	.90	.88	

^a B = Banque; CC = Entreprise de cartes de crédit; R & M = Entreprise de réparation et de maintenance; LDT = Compagnie de téléphone longue distance.

^b Les numéros des items sont ceux des énoncés relatifs aux attentes et aux perceptions situés en appendice.

couple de valeurs relatives à la dimension « éléments tangibles ». La fiabilité pour l'échelle totale (*i.e.* la fiabilité de la combinaison linéaire) est proche de 0,9 dans chacun des quatre cas.

Les résultats des analyses factorielles des données issues des quatre échantillons sont synthétisés dans le tableau 3. La forme globale des *loadings* factoriels est remarquablement similaire à travers les quatre groupes indépendants de résultats. A quelques exceptions près, les items affectés à chaque dimension ont de façon cohérente des *loadings* élevés uniquement sur l'un des cinq facteurs extraits. Le caractère distinct des cinq dimensions de SERVQUAL suggéré par les résultats du tableau 3 a été confirmé ensuite par les intercorrélations relativement faibles des cinq facteurs entre eux — les corrélations moyennes deux à deux entre facteurs après rotation oblique étaient de 0,21, 0,24, 0,26 et 0,23 pour les échantillons relatifs à la banque, aux cartes de

TABLEAU 3. — *Matrices des loadings factoriels après rotation oblique sur une analyse à cinq facteurs*

Items	« LOADINGS » FACTORIELS																			
	Banque					Entreprise de cartes de crédit					Entreprise de réparation/maintenance					Compagnie de téléphone longue distance				
	F ₁	F ₂	F ₃	F ₄	F ₅	F ₁	F ₂	F ₃	F ₄	F ₅	F ₁	F ₂	F ₃	F ₄	F ₅	F ₁	F ₂	F ₃	F ₄	F ₅
Q ₁	34	28				36		35			34					42				
Q ₂	64					70					70					72				
Q ₃	39			28		52					53					51				
Q ₄	28			28		52					65					59			30	
Q ₅		72					54					73					52			
Q ₆		63					43	27				51					40			
Q ₇		71					87					84					79			
Q ₈		80					83					88					59			
Q ₉		39					49					29		30			54			
Q ₁₀			37					43		26			56					39		
Q ₁₁			55					48					52					43		
Q ₁₂			62					54					74					92		
Q ₁₃			69					33					71					53		
Q ₁₄				68					65					86					69	
Q ₁₅				84					76					89					81	
Q ₁₆				72					73					65					61	
Q ₁₇				64					61					64					66	
Q ₁₈					37					64					42					59
Q ₁₉					48					72					61					79
Q ₂₀					41					63		28	34		46					55
Q ₂₁					33					59					32					36
Q ₂₂					68					64					61					59

Tous les chiffres de ce tableau représentent des *loadings* multipliés par 100. Les *loadings* inférieurs ou égaux à 0,25 ne sont pas indiqués. Le pourcentage de variance extrait par les cinq facteurs dans les échantillons représentatifs des entreprises bancaires, de cartes de crédit, de réparation/maintenance et de téléphone longue distance était de 56 %, 57,5 %, 61,6 % et 56,2 % respectivement.

crédit, à la réparation et la maintenance et au téléphone longue distance, respectivement¹.

En guise de vérification supplémentaire des coefficients de fiabilité et de la structure factorielle de SERVQUAL, le premier ensemble de données ayant abouti à un instrument de 34 items en sept dimensions a été analysé de nouveau après suppression des 12 items abandonnés lors de la seconde phase de purification de l'échelle. Les résultats de cette analyse sont résumés dans le tableau 4 et apportent une confirmation supplémentaire des forts degrés de fiabilité et de distinction entre les dimensions de l'échelle. La corrélation moyenne entre les cinq facteurs après rotation oblique était de 0,35.

TABLEAU 4. — Réanalyse des données de la première collecte pour l'échelle à cinq dimensions

Dimensions	Inti- tulé	Nom- bre d'items	Coeffi- cients de fia- bilité (Alphas)	Items	Loadings factoriels des items sur les dimensions auxquelles ils se rattachent ^a
Eléments tangibles	F ₁	4	.72	Q ₁	69
				Q ₂	68
				Q ₃	64
				Q ₄	51
Fiabilité	F ₂	5	.83	Q ₅	75
				Q ₆	63
				Q ₇	71
				Q ₈	75
				Q ₉	50
Serviabilité	F ₃	4	.82	Q ₁₀	51
				Q ₁₁	77
				Q ₁₂	66
				Q ₁₃	86
Assurance	F ₄	4	.81	Q ₁₄	38
				Q ₁₅	72
				Q ₁₆	80
				Q ₁₇	45
Empathie	F ₅	5	.86	Q ₁₈	78
				Q ₁₉	81
				Q ₂₀	59
				Q ₂₁	71
				Q ₂₂	68
Fiabilité de la combinaison linéaire (fiabilité de l'échelle totale)			.92		

a. Les chiffres représentent les *loadings* multipliés par 100. Les *loadings* des items sur les dimensions auxquelles ils n'appartiennent pas étaient inférieurs à 0,3. Le pourcentage de variance extrait par les cinq facteurs était de 63,2 %.

1. Les matrices complètes des corrélations interfacteurs peuvent s'obtenir sur demande au premier auteur.

Il convient de remarquer que la procédure itérative utilisée pour raffiner l'instrument initial a été guidée par des critères empiriques et par l'objectif de départ qui était d'obtenir une échelle concise dont les items soient pertinents pour une variété de prestataires de services. Les coefficients de fiabilité et les structures factorielles indiquent que l'échelle finale composée de 22 items articulés en cinq dimensions présente des propriétés psychométriques saines et stables. De plus, par construction, la procédure itérative n'a retenu que les items qui sont communs à, et pertinents pour, toutes les entreprises de service comprises dans cette étude. De ce fait même, cependant, il est possible que cette procédure ait supprimé certains items « corrects » pertinents pour certaines des entreprises mais pas pour l'ensemble. C'est pourquoi, si SERVQUAL peut être utilisé sous sa forme actuelle pour évaluer et comparer la qualité de service sur une grande variété de firmes ou d'unités à l'intérieur d'une firme, il convient de procéder à une adaptation appropriée de l'instrument quand il s'agit d'étudier une catégorie unique de service. En particulier, il est possible de reformuler les items des cinq dimensions, et/ou d'en accroître le nombre, pour les rendre plus appropriés au contexte d'utilisation de l'instrument.

ÉVALUATION DE LA VALIDITÉ DE SERVQUAL

Les forts coefficients de fiabilité et les structures factorielles cohérentes de SERVQUAL à travers des échantillons indépendants constituent des éléments en faveur de la validité de trait (Campbell, 1960; Peter, 1981). Néanmoins, si de forts niveaux de fiabilité et de cohérences internes constituent des conditions nécessaires à la validité de construit de l'échelle — *i.e.* dans quelle mesure une échelle capture complètement et sans ambiguïté le construit sous-jacent, inobservable, qu'elle est censée mesurer? — ces conditions ne sont pas suffisantes (Churchill, 1979). L'échelle doit satisfaire d'autres critères conceptuels et empiriques avant de pouvoir être considérée comme présentant une bonne validité de construit.

Le critère conceptuel fondamental concernant la validité de construit est la validité de contenu (Est-ce que l'échelle mesure effectivement ce qu'elle est supposée mesurer? Est-ce que les items de l'échelle capturent les facettes essentielles du construit inobservable que l'on mesure?). L'évaluation de la validité de contenu d'une échelle est nécessairement qualitative plutôt que quantitative. Elle implique deux choses : 1) la minutiosité et la profondeur d'explicitation du construit et de son domaine, et 2) le degré de représentation du domaine du construit par les items de l'échelle. Ainsi que le suggèrent les sections précédentes, les procédures utilisées pour développer SERVQUAL satisfont à ces deux exigences. On peut donc considérer l'échelle comme possédant la validité de construit.

La validité de l'échelle doit aussi être évaluée empiriquement en examinant sa validité convergente — *i.e.* l'association entre les scores SERVQUAL et les réponses à une question qui demandait aux consommateurs d'apporter

une note globale de la qualité de service de l'entreprise qu'ils évaluaient. Les répondants de la seconde étape de collecte de données ont évalué la qualité globale de l'entreprise de services (appelée « Q globale » dans les lignes qui suivent) en choisissant parmi les quatre catégories — excellente, bonne, correcte, mauvaise. La correspondance entre les évaluations de Q globale et les scores de SERVQUAL a été analysée au moyen d'une ANOVA univariée. La variable correspondant au facteur dans l'ANOVA était la Q globale composée de trois niveaux au lieu de quatre en raison de la faible fréquence de choix, la catégorie « mauvaise » qualité globale; cela a conduit à créer une catégorie combinant qualité globale « correcte-mauvaise ». La variable dépendante était le score de différence moyen (*i.e.* score égal aux perceptions moins attentes) pour chaque dimension de SERVQUAL, ainsi que pour l'échelle SERVQUAL totale (on a procédé à des ANOVA séparées pour chaque dimension et pour l'échelle totale). Les résultats significatifs des ANOVA ont été complétés par le test des intervalles multiples de Duncan afin d'identifier les différences significatives entre les catégories de Q globale. Les résultats de ces analyses pour chacun des quatre échantillons sont résumés dans le tableau 5 intitulé « Q globale ».

Les chiffres reportés dans ce tableau représentent les scores moyens sur SERVQUAL dans chaque catégorie de Q globale, mesurés à l'aide d'une échelle graduée de — 6 à +6, et sur laquelle plus le score est élevé (moins il est négatif), plus le niveau de qualité de service perçue est élevé. Dans chacun des quatre échantillons, le score combiné SERVQUAL pour les répondants ayant choisi la catégorie qualité de service « excellente » est significativement plus élevé (moins négatif) que pour ceux ayant trouvé une qualité de service « bonne ». Parallèlement, les répondants situés dans la catégorie « bonne » qualité globale ont un score combiné SERVQUAL significativement plus élevé que ceux de la catégorie qualité globale « correcte-mauvaise ». De façon évidente, on trouve aussi une structure de résultats similaire pour les scores sur les dimensions individuelles de SERVQUAL. La force et la persistance du lien entre les catégories de Q globale et les scores sur SERVQUAL à travers quatre échantillons indépendants traduisent clairement la bonne validité convergente de SERVQUAL.

On a poursuivi l'examen de la validité de SERVQUAL par l'étude de l'existence d'un lien empirique entre le construit que l'échelle mesure et d'autres variables qui lui sont liées conceptuellement. Les individus de chaque échantillon ont répondu à deux questions générales destinées à servir de mesure à des variables (intitulées « Recommander » et « Problème » dans le tableau 5) dont on pouvait penser qu'elles étaient reliées conceptuellement à la qualité de service perçue; il s'agit de : 1) les répondants recommanderaient-ils ou non le prestataire de service à un ami; et 2) les répondants ont-ils ou non déjà signalé un problème au sujet des services qu'ils ont reçus de l'entreprise. On pouvait donc faire l'hypothèse que les individus répondant oui à la première question (Recommander) et non à la seconde (Problème) percevaient une qualité de service supérieure à celle perçue par les autres répondants. Ainsi que l'indique le tableau 5, les résultats sont cohérents avec cette hypothèse. Ces résultats apportent une confirmation supplémentaire de la validité de SERVQUAL.

TABLEAU 5. — *Différences significatives des moyennes sur l'échelle pour les répondants. Segmentation selon les variables Q globale, Recommander, Problème*

Dimensions de l'échelle individuelle	Q globale			Recommander		Problème	
	Excel-lente	Bonne	Correcte/mauvaise	Oui	Non	Oui	Non
BANQUE							
Eléments tangibles	— 0,04 ^b	— 0,52 ^c	— 1,08 ^d	— 0,41 ^b	— 0,98 ^b	— 0,75 ^b	— 0,45 ^b
Fiabilité	— 0,25 ^b	— 0,96 ^c	— 2,30 ^d	— 0,82 ^b	— 2,21 ^c	— 1,55 ^b	— 0,92 ^c
Servabilité	— 0,32 ^b	— 0,97 ^c	— 1,54 ^c	— 0,74 ^b	— 1,81 ^c	— 1,22 ^b	— 0,84 ^b
Assurance	— 0,49 ^b	— 1,03 ^c	— 1,98 ^d	— 0,88 ^b	— 2,12 ^c	— 1,52 ^b	— 0,96 ^c
Empathie	— 0,30 ^b	— 1,02 ^c	— 1,52 ^c	— 0,76 ^b	— 1,88 ^c	— 1,07 ^b	— 0,91 ^b
Echelle combinée	— 0,22 ^b	— 0,92 ^c	— 1,61 ^d	— 0,72 ^b	— 1,77 ^c	— 1,22 ^b	— 0,80 ^c
Taille de l'échantillon	46	112	40	164	33	47	151
CARTES DE CRÉDIT							
Eléments tangibles	0,06 ^b	— 0,61 ^c	— 0,79 ^c	— 0,39 ^b	— 0,80 ^b	— 0,76 ^b	— 0,29 ^c
Fiabilité	— 0,42 ^b	— 0,94 ^c	— 2,32 ^d	— 0,82 ^b	— 2,50 ^c	— 1,42 ^b	— 0,82 ^c
Servabilité	— 0,08 ^b	— 1,13 ^c	— 1,71 ^c	— 0,75 ^b	— 2,59 ^c	— 1,31 ^b	— 0,77 ^c
Assurance	— 0,59 ^b	— 1,31 ^c	— 2,29 ^d	— 1,08 ^b	— 2,83 ^c	— 1,49 ^b	— 1,15 ^b
Empathie	— 0,50 ^b	— 1,38 ^c	— 1,94 ^c	— 1,03 ^b	— 2,77 ^c	— 1,62 ^b	— 1,01 ^b
Echelle combinée	— 0,32 ^b	— 1,10 ^c	— 1,79 ^d	— 0,83 ^b	— 2,27 ^c	— 1,29 ^b	— 0,83 ^c
Taille de l'échantillon	60	112	28	183	17	50	149
RÉPARATION/ MAINTENANCE							
Eléments tangibles	— 0,15 ^b	— 0,40 ^{b,c}	— 0,86 ^c	— 0,36 ^b	— 0,85 ^b	— 0,58 ^b	— 0,34 ^b
Fiabilité	— 0,48 ^b	— 1,30 ^c	— 3,20 ^d	— 1,14 ^b	— 3,48 ^c	— 2,14 ^b	— 1,18 ^c
Servabilité	— 0,08 ^b	— 1,08 ^c	— 2,41 ^d	— 0,83 ^b	— 2,54 ^c	— 1,71 ^b	— 0,80 ^c
Assurance	— 0,33 ^b	— 1,35 ^c	— 2,84 ^d	— 1,16 ^b	— 2,91 ^c	— 2,04 ^b	— 1,13 ^c
Empathie	0,15 ^b	— 1,11 ^c	— 2,17 ^d	— 0,85 ^b	— 2,19 ^c	— 1,67 ^b	— 0,74 ^c
Echelle combinée	— 0,16 ^b	— 1,07 ^c	— 2,30 ^d	— 0,88 ^b	— 2,40 ^c	— 1,65 ^b	— 0,85 ^c
Taille de l'échantillon	45	114	40	168	30	65	132
TÉLÉPHONE LONGUE DISTANCE							
Eléments tangibles	— 0,08 ^b	— 0,44 ^c	— 0,50 ^c	— 0,26 ^b	— 0,95 ^c	— 0,42 ^b	— 0,26 ^b
Fiabilité	— 0,45 ^b	— 1,42 ^c	— 2,53 ^d	— 1,05 ^b	— 2,71 ^c	— 1,54 ^b	— 1,03 ^c
Servabilité	— 0,30 ^b	— 1,43 ^c	— 1,90 ^c	— 1,00 ^b	— 2,03 ^c	— 1,46 ^b	— 0,86 ^c
Assurance	— 0,39 ^b	— 1,45 ^c	— 2,10 ^d	— 1,00 ^b	— 2,64 ^c	— 1,62 ^b	— 0,87 ^c
Empathie	— 0,33 ^b	— 1,19 ^c	— 2,10 ^d	— 0,86 ^b	— 2,34 ^c	— 1,16 ^b	— 0,90 ^b
Echelle combinée	— 0,30 ^b	— 1,15 ^c	— 1,83 ^d	— 0,83 ^b	— 2,13 ^c	— 1,24 ^b	— 0,76 ^c
Taille de l'échantillon	69	104	25	178	19	78	120

a. Les chiffres représentent les valeurs moyennes des réponses sur une échelle allant de — 6 à + 6, le zéro impliquant que les perceptions du consommateur coïncident avec ses attentes, des valeurs négatives signifiant que les perceptions sont en dessous des attentes et des valeurs positives signifiant que les perceptions dépassent les attentes.

b, c, d. Les moyennes indiquées par une note identique ne sont pas significativement différentes les unes des autres. Les moyennes indiquées par une note différente le sont.

APPLICATIONS POTENTIELLES DE SERVQUAL

Il est difficile d'imaginer un détaillant qui ne propose absolument aucun service. Certains offrent des services pratiques comme l'assistance lors de l'achat ou encore des services de livraison; cela contribue à vendre les produits. D'autres vendent directement des services. La qualité du service représente donc une question importante pour tous ces détaillants. Les secteurs de distribution hautement concurrentiels (les grands magasins, les supermarchés) peuvent vendre de nombreux produits identiques et la qualité de service constitue alors un moyen élémentaire de différenciation concurrentielle. Quant aux entreprises qui vendent uniquement des services (compagnies de télécommunications, compagnies aériennes), elles ont bien peu à proposer si leur service est médiocre (Berry, 1986).

SERVQUAL est une échelle multi-items, synthétique, disposant de bons niveaux de fiabilité et de validité, qui peut être utilisée par les détaillants afin d'améliorer leur compréhension des attentes et des perceptions des consommateurs en matière de service, et, en conséquence, d'améliorer le service offert. L'instrument a été conçu pour s'appliquer à une large gamme de services. En tant que tel, il permet de dresser une image fondamentale de la qualité de service perçue grâce à sa conception sous un format double attentes/perceptions autour d'énoncés structurés en cinq dimensions de qualité de service. Quand cela est nécessaire, cette image peut être adaptée ou complétée pour correspondre aux caractéristiques d'une organisation particulière ou à ses objectifs de recherche spécifiques.

SERVQUAL trouve une valeur toute particulière quand il est utilisé périodiquement pour repérer les tendances suivies par la qualité de service, et quand il est utilisé en conjonction avec d'autres types de mesure de celle-ci. Un détaillant, par exemple, obtiendrait énormément d'informations sur la qualité de son service et sur ce qu'il conviendrait de mettre en œuvre pour l'améliorer, s'il utilisait à la fois SERVQUAL et une étude auprès des employés trois ou quatre fois par an, tout en sollicitant et analysant les suggestions et les plaintes des consommateurs. L'étude auprès des employés devrait inclure des questions concernant ce qui est perçu comme faisant obstacle à un meilleur service, par exemple : quel est le plus gros problème que vous affrontez quand vous essayez de fournir une haute qualité de service à vos consommateurs ? Si vous étiez PDG pour une journée, quel est l'élément que vous modifieriez dans l'entreprise pour améliorer la qualité de service ?

SERVQUAL peut être utilisé pour évaluer la qualité d'une entreprise donnée sur les cinq dimensions de service par calcul des scores de différences moyens sur les items composant la dimension. On peut aussi déduire de l'instrument une mesure globale de la qualité de service sous la forme d'un score moyen sur les cinq dimensions. Puisque des réponses sensées aux énoncés de perceptions requièrent de la part des répondants une certaine connaissance ou expérience de l'entreprise étudiée, SERVQUAL est limité aux consommateurs habituels ou récents de cette entreprise. Sous cette contrainte, il existe une variété d'applications potentielles.

La première consiste à déterminer *l'importance relative* des cinq dimensions

influençant les perceptions globales de la qualité par les consommateurs. Un moyen de faire cela consiste à régresser les scores globaux de perception de la qualité sur les scores SERVQUAL pour chaque dimension individuelle. Les résultats d'une telle analyse de régression pour les quatre entreprises de cette étude sont reportés dans le tableau 6 (la variable dépendante était la Q globale codifiée de la façon suivante : excellente = 4, bonne = 3, correcte = 2 et mauvaise = 1).

TABLEAU 6. — *Importance relative des cinq dimensions en termes de prédiction de la qualité globale*

Dimensions	Coefficient de régression standardisé	Niveau de signification de la pente ^a	R ² ajusté
BANQUE			
Eléments tangibles	.13	.07	.28 (<i>p</i> < .00)
Fiabilité	.39	.00	
Serviabilité	.07	.35	
Assurance	.13	.09	
Empathie	.01	.89	
CARTES DE CRÉDIT			
Eléments tangibles	.07	.26	.27 (<i>p</i> < .00)
Fiabilité	.33	.00	
Serviabilité	.12	.11	
Assurance	.17	.02	
Empathie	.04	.58	
RÉPARATION/MAINTENANCE			
Eléments tangibles	.04	.48	.52 (<i>p</i> < .00)
Fiabilité	.54	.00	
Serviabilité	.11	.09	
Assurance	.16	.02	
Empathie	.01	.81	
TÉLÉPHONE LONGUE DISTANCE			
Eléments tangibles	.08	.17	.37 (<i>p</i> < .00)
Fiabilité	.45	.00	
Serviabilité	.12	.09	
Assurance	.15	.03	
Empathie	.02	.78	

a. Niveaux de signification pour des tests bilatéraux.

a. Niveaux de signification pour des tests bilatéraux.

Les valeurs du R carré ajusté sont statistiquement significatives pour chacun des quatre cas, en particulier si l'on considère que la variable dépendante était constituée de quatre catégories, et que les trois premières rendaient compte de la plupart des réponses. Un résultat surprenant au niveau de l'importance relative des cinq dimensions en termes de prédiction de la qualité globale est que la fiabilité représente la dimension la plus critique.

L'assurance vient en deuxième position d'importance pour les quatre cas. Les éléments tangibles sont plus importants pour ce qui est de la banque que pour les trois autres firmes, mais l'inverse se produit pour la serviabilité. L'empathie représente la dimension la moins importante dans tous les cas. Cependant, les magnitudes relativement faibles des coefficients de régression pour la dimension empathie, ainsi que leur non-significativité statistique devraient être interprétées avec prudence, puisque l'empathie avait des coefficients de corrélations simples significatifs avec la qualité globale et allant de 0,20 pour la banque à 0,40 pour l'entreprise de réparation et de maintenance. L'empathie disposait aussi de corrélations significatives d'une magnitude similaire avec les dimensions de fiabilité et d'assurance (les deux plus importantes), ce qui implique que son importance dans les analyses de régression ait pu être masquée par un phénomène possible de multicollinéarité. En conséquence, si l'empathie est apparemment la dimension la moins importante de SERVQUAL, elle n'est en aucun cas *sans* importance.

Une autre application de l'instrument consiste à l'utiliser pour créer avec les consommateurs d'une entreprise plusieurs catégories de segments de qualité perçue (par ex., haute, moyenne, basse) sur la base de leurs scores individuels sur SERVQUAL. Ces segments peuvent être analysés en termes : 1) de critères démographiques, psychographiques et/ou d'autres critères, 2) d'importance relative de chaque dimension influençant les perceptions de la qualité de service, et 3) de raisons sous-jacentes aux perceptions enregistrées. Par exemple, supposons qu'un grand magasin s'aperçoive qu'un nombre considérable de répondants à SERVQUAL faisant partie du groupe de qualité perçue « moyenne » correspond à sa cible de marché principale déterminée par des critères démographiques et psychographiques. Supposons ensuite que la fiabilité et l'assurance aient été identifiées comme les dimensions de qualité les plus importantes et enfin, sur la base des scores de différences entre perceptions et attentes sur les items de ces dimensions, que les items ayant trait au suivi des dossiers et au comportement du personnel en contact avec le client révèlent les plus grandes différences. À l'aide de ces données, le personnel d'encadrement du grand magasin gagnerait en compréhension de ce qu'il convient de mettre en œuvre pour améliorer son image aux yeux d'un groupe très important de consommateurs — la principale cible de l'entreprise qui émet un score de qualité de service « moyenne » et qui est en position soit de réagir à une amélioration du service de l'entreprise, soit de se tourner vers la concurrence.

SERVQUAL peut aussi être utilisé par des entreprises de distribution organisées en réseaux de points de vente afin de repérer le niveau de service offert par chacun des points de vente de la chaîne. En demandant aux répondants d'indiquer le magasin particulier de la chaîne avec lequel ils sont le plus familiers, et de donner leurs perceptions pour cette unité, le chercheur peut comparer le score moyen de chaque magasin sur SERVQUAL avec ceux des autres magasins. Entre autres usages, les scores de qualité de service peuvent constituer un facteur d'évaluation de la performance et de motivation du gérant du magasin. Par ailleurs, les scores sur SERVQUAL pour les magasins individuels peuvent être employés pour constituer des groupes de points de vente ayant des images de qualités différentes. Une étude attentive

des caractéristiques des points de vente de chaque groupe peut mettre en évidence les attributs clés qui favorisent — ou freinent — la prestation d'un service de haute qualité.

Un détaillant peut aussi utiliser SERVQUAL pour évaluer sa performance en matière de service par rapport à ses principaux concurrents. Le format en deux sections de l'instrument, qui distingue les attentes des perceptions, en fait un outil pratique pour mesurer la qualité de plusieurs entreprises en incluant un ensemble d'énoncés de perceptions pour chaque firme. Il n'est pas nécessaire de répéter la section concernant les attentes pour chacune d'entre elles. A titre d'exemple, une chaîne de supermarchés pourrait prendre en compte ses deux principaux concurrents dans une étude du marché global, en demandant aux répondants de donner leurs évaluations de perception pour chacune des entreprises dont ils ont déjà été clients. Un distributeur utilisant SERVQUAL pour identifier les dimensions de la qualité de service les plus pertinentes pour ses marchés cibles, et pour se comparer avec la concurrence en termes de forces et faiblesses sur ces dimensions particulières, aura une idée de ce que devront être ses priorités en matière de qualité de service.

En résumé, SERVQUAL présente une variété d'applications potentielles. Il peut contribuer à évaluer les attentes des consommateurs en matière de qualité de service ainsi que leurs perceptions de celle-ci pour une large gamme de services et d'organisations de distribution. Il peut aussi permettre d'éclairer les domaines devant attirer l'attention de l'encadrement et faire l'objet d'actions spécifiques d'amélioration de la qualité de service. Enfin, nous espérons que la publication de l'instrument stimulera la mise en œuvre indispensable de recherches empiriques centrées sur la qualité de service, ses antécédents et ses conséquences.

RÉFÉRENCES

- Berry, Leonard L. (1986), Retail Businesses Are Service Businesses, *Journal of Retailing*, 62, (Spring), 3-6.
- Campbell, Donald T. (1960), Recommendations for APA Tests Standards Regarding Construct, Trait or Discriminant Validity, *American Psychologist*, 15 (August), 546-553.
- Churchill, Gilbert A. Jr. (1979), A Paradigm for Developing Better Measures of Marketing Constructs, *Journal of Marketing Research*, 16 (February), 64-73.
- Churchill, Gilbert A. Jr., Ford Neil M. and Walker Orville C. Jr (1974), Measuring the Job Satisfaction of Industrial Salesmen, *Journal of Marketing Research*, 11 (August), 254-260.
- Cronbach, Lee J. (1951), Coefficient Alpha and the Internal Structure of Tests, *Psychometrika*, 16 (October), 297-234.
- Crosby, Philip B. (1979), *Quality is Free : The Art of Making Quality Certain*, New York, New American Library.
- Dodds, William B. and Monroe Kent B. (1984), The Effect of Brand and Price Information on Subjective Product Evaluations, *Advances in Consumer Research* XII.
- Ford, Neil M., Walker Orville C. Jr. and Churchill Gilbert A. Jr. (1975), Expectation-Specific Measures of the Intersender Conflict and Role Ambiguity Experienced by Industrial Salesmen, *Journal of Business Research*, 3 (April), 95-112.
- Garvin, David A. (1983), Quality on the Line, *Harvard Business Review*, 61 (September-October), 65-73.
- Gronroos, Christian (1982), *Strategic Management and Marketing in the Service Sector*, Helsingfors, Swedish School of Economics and Business Administration.
- Harman, Harry H. (1967), *Modern Factor Analysis*, 2nd ed., Chicago, The University of Chicago Press.
- Hjorth-Anderson, Chr. (1984), The Concept of Quality and the Efficiency of Markets for Consumer Products, *Journal of Consumer Research*, 11 (September), 708-718.

- Holbrook, Morris B. and Corfman Kim P. (1985), Quality and Value in the Consumption Experience : Phaldrus Rides Again, in *Perceived Quality*, J. Jacoby and J. Olson (eds.), Lexington, Massachusetts, Lexington Books, 31-57.
- Howard, John and Sheth Jagdish (1969), *The Theory of Buyer Behavior*, New York, John Wiley & Sons.
- Hunt, Keith (1979), *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, Cambridge, Mass., Marketing Science Institute.
- Jacoby, Jacob (1978), Consumer Research : A State of the Art Review, *Journal of Marketing*, 42 (April), 87-96.
- Jacoby, J. and Olson J. (eds.) (1985), *Perceived Quality*, Lexington, Massachusetts, Lexington Books.
- Lehtinen, Uolevi and Lehtinen Jarmo R. (1982), « Service Quality : A Study of Quality Dimensions », unpublished working paper, Helsinki, Service Management Institute, Finland oy.
- Lovelock, Christopher H. (1980), Towards a Classification of Services, in *Emerging Perspectives on Services Marketing*, L. L. Berry, G. L. Shostack and G. Upah (eds.), Chicago, American Marketing Association, 72-76.
- Lovelock, Christopher H. (1983), Classifying Services to Gain Strategic Marketing Insights, *Journal of Marketing*, 47 (Summer), 9-20.
- Nunnally, Jum C. (1978), *Psychometric Theory*, 2nd ed., New York, McGraw-Hill Book Company.
- Oliver, Richard (1981), Measurement and Evaluation of Satisfaction Process in Retail Settings, *Journal of Retailing*, 57 (Fall), 25-48.
- Olshavsky, Richard W. (1985), Perceived Quality in Consumer Decision Making: An Integrated Theoretical Perspective, in *Perceived Quality*, J. Jacoby and J. Olson (eds.), Lexington, Massachusetts, Lexington Books.
- Parasuraman, A., Zeithaml Valarie and Berry Leonard (1985), A Conceptual Model of Service Quality and its Implications for Future Research, *Journal of Marketing* (Fall), 41-50.
- Peter, J. Paul (1981), Construct Validity : A Review of Basic Issues and Marketing Practices, *Journal of Marketing Research*, 18 (May), 133-145.
- Rudie, Mary, Brant Wansley J. and H. (1985), The Merrill Lynch Quality Program, in *Services Marketing in a Changing Environment*, Thomas Bloch, Gregory Upah, and Valarie A. Zeithaml (eds.), Chicago, IL, American Marketing Association.
- Sasser, W. Earl Jr., Olsen R. Paul and Wyckoff D. Dary (1978), *Management of Service Operations : Text and Cases*, Boston, Allyn & Bacon.
- Saxe, Robert and Weitz Barton A. (1982), The soco Scale : A Measure of the Customer Orientation of Salespeople, *Journal of Marketing*, 19 (August), 343-351.
- Thomson, Phillip, DeSouza, Glenn and Gale Bradley T. (1985), *The Strategic Management of Service Quality*, Cambridge, MA, The Strategic Planning Institute, PIMSLETTER, n° 33.
- Zeithaml, Valarie (1987), *Defining and Relating Price, Perceived Quality and Perceived Value*, Report n° 87-101, Cambridge, MA, Marketing Science Institute.

ANNEXE : L'INSTRUMENT SERVQUAL²

Instructions : Cette enquête porte sur vos opinions concernant les services ———. Veuillez indiquer, s'il vous plaît, dans quelle mesure vous pensez que les entreprises proposant ces services devraient posséder les caractéristiques décrites par chaque énoncé. Pour ce faire, choisissez l'un des chiffres proposés avec chaque énoncé. Par exemple, si vous êtes tout à fait d'accord pour que ces entreprises possèdent telle caractéristique, entourez le chiffre 7. Dans le cas contraire, entourez le chiffre 1. Si votre opinion n'est pas clairement départagée, entourez l'un des chiffres intermédiaires. Il n'y a pas de bonnes ou de mauvaises réponses — tout ce qui nous intéresse est le chiffre qui traduit au mieux vos attentes à l'égard des entreprises offrant des services ———.

- E1. Elles devraient posséder un équipement récent.
- E2. Leurs installations matérielles devraient être visuellement attractives.
- E3. Leurs employés devraient être correctement vêtus et d'apparence soignée.
- E4. L'aspect des installations matérielles de ces entreprises devrait être en rapport avec le type de services offerts.

2. Une échelle à sept points allant de « tout à fait d'accord » (7) à « pas du tout d'accord » (1), sans intitulé pour les points intermédiaires (i.e. de 2 à 6), accompagne chaque énoncé. De plus, les énoncés ont été placés aléatoirement au sein du questionnaire. Le lecteur peut obtenir auprès du premier auteur une liste complète de la version en 34 items de l'instrument utilisée lors de la deuxième étape de collecte de données.

N.d.I.R. — L'échelle a été traduite de l'article d'origine sans prendre en compte — au moyen d'une traduction à rebours notamment — les questions d'équivalence cross-culturelle de l'instrument.

- E5. Quand ces entreprises s'engagent à réaliser quelque chose dans un délai donné, elles devraient s'y tenir.
- E6. Quand les consommateurs rencontrent des problèmes, ces entreprises devraient adopter une attitude sympathique et rassurante.
- E7. Ces entreprises devraient être dignes de confiance.
- E8. Elles devraient fournir leurs services dans les délais sur lesquels elles se sont engagées.
- E9. Elles devraient tenir leurs dossiers précisément.
- E10. On ne devrait pas s'attendre à ce qu'elles aient à annoncer aux consommateurs le moment précis où elles fourniront les services (—)³.
- E11. Il n'est pas réaliste de la part des consommateurs de s'attendre à ce que les employés de ces entreprises rendent des services prompts (—).
- E12. Leurs employés ne sont pas toujours censés être disposés à aider les consommateurs (—).
- E13. Cela passe s'ils sont trop occupés pour répondre aux demandes des consommateurs immédiatement (—).
- E14. Les consommateurs devraient pouvoir faire confiance aux employés de ces entreprises.
- E15. Les consommateurs devraient pouvoir effectuer leurs transactions avec les employés de ces entreprises en toute sécurité.
- E16. Leurs employés devraient être polis.
- E17. Leurs employés devraient obtenir le soutien approprié de ces entreprises pour faire correctement leur travail.
- E18. On ne devrait pas s'attendre à ce que ces entreprises accordent une attention individuelle à chaque consommateur (—).
- E19. On ne peut pas s'attendre à ce que les employés de ces entreprises accordent une attention personnalisée aux consommateurs (—).
- E20. Il n'est pas réaliste de s'attendre à ce que les employés connaissent les besoins de leurs clients (—).
- E21. Il n'est pas réaliste de s'attendre à ce que ces entreprises aient à cœur d'agir dans les meilleurs intérêts des consommateurs (—).
- E22. On ne devrait pas s'attendre à des horaires d'ouverture qui soient pratiques pour les consommateurs.

Instructions : L'ensemble des énoncés suivants concerne vos perceptions à l'égard de XYZ. Pour chaque énoncé, veuillez indiquer, s'il vous plaît, dans quelle mesure vous croyez que XYZ possède les caractéristiques décrites dans chaque énoncé. Une fois de plus, le fait d'entourer le chiffre 7 signifie que vous pensez tout à fait que XYZ possède cette caractéristique. Vous pouvez entourer n'importe quel chiffre intermédiaire afin d'indiquer le degré de vos sentiments. Il n'y a pas de bonne ou de mauvaise réponse — nous sommes intéressés uniquement par le chiffre qui traduit au mieux vos perceptions au sujet de XYZ.

- P1. XYZ possède un équipement récent.
- P2. Les installations matérielles de XYZ sont visuellement attractives.
- P3. Les employés de XYZ sont correctement vêtus et ont une apparence soignée.
- P4. L'aspect des installations matérielles de XYZ est en accord avec le type de services qu'il assure.
- P5. Quand XYZ s'engage à réaliser quelque chose dans un délai donné, il s'y tient.
- P6. Quand vous rencontrez des problèmes, XYZ adopte une attitude sympathique et rassurante.
- P7. XYZ est digne de confiance.
- P8. XYZ fournit ses services dans les délais sur lesquels il s'est engagé.
- P9. XYZ tient ses archives précisément.
- P10. XYZ n'indique pas précisément quand les services seront fournis (—).
- P11. Les employés de XYZ n'assurent pas un service prompt (—).
- P12. Les employés de XYZ ne sont pas toujours disposés à aider les consommateurs (—).
- P13. Les employés de XYZ sont trop occupés pour répondre aux demandes des consommateurs dans l'immédiat (—).
- P14. Vous pouvez faire confiance aux employés de XYZ.
- P15. Vous effectuez vos transactions avec les employés de XYZ en toute sécurité.
- P16. Les employés de XYZ sont polis.
- P17. Les employés obtiennent de la part de XYZ le soutien nécessaire pour faire leur travail correctement.
- P18. Les employés de XYZ ne vous accordent pas une attention individuelle (—).
- P19. Les employés de XYZ ne vous accordent pas une attention personnalisée (—).
- P20. Les employés de XYZ ne connaissent pas vos besoins (—).
- P21. XYZ n'a pas à cœur d'agir selon vos meilleurs intérêts (—).
- P22. Les horaires d'ouverture de XYZ ne sont pas pratiques pour tous ses consommateurs (—).

3. Les scores d'évaluations sur ces énoncés ont été inversés pour l'analyse de données.