

Mesurer la qualité perçue d'un site web : une application de la procédure C-OAR-SE

Yves Cinotti

PR.AG. économie et gestion à l'Université Toulouse II - Doctorant à l'IAE de Toulouse I

Adresse professionnelle

Département CETIA
Université Toulouse II
5, allées Antonio Machado
31058 TOULOUSE Cedex 9

Adresse personnelle :

38, rue du Bousquet
31650 SAINT-ORENS DE GAMEVILLE
☎ 06 70 63 28 70
Courriel : yvcinotti@free.fr

Mesurer la qualité perçue d'un site web : une application de la procédure C-OAR-SE

Résumé

L'objectif de cette recherche est de réaliser une application de la procédure C-OAR-SE (Rossiter, 2002). Une échelle de mesure de la qualité d'un site Web informatif est développée. La comparaison de l'échelle QSR avec une échelle « churchillienne » s'avère délicate. Quelques adaptations possibles de la procédure C-OAR-SE sont suggérées.

Mots-clés : mesure, échelle, qualité de service, site Web informatif

Measuring the Perceived Quality of a Website: an Application of the C-OAR-SE Procedure

Abstract

The purpose of this study is to carry out a first application of the C-OAR-SE procedure (Rossiter, 2002). A scale was developed to measure the perceived quality of informational website. Comparing QSR with a "churchillian" scale seems delicate. Some adaptations of the C-OAR-SE procedure are suggested.

Keywords : *measure, scale, service quality, informational website*

L'évolution du marketing vers plus de scientificité est certainement liée à l'utilisation d'instruments de mesure performants (Churchill, 1979). Or, la mesure d'un construit marketing n'est pas bonne parce que quelqu'un l'a dit mais parce qu'elle satisfait les critères de validité, de fiabilité et de sensibilité (Jacoby, 1978). La procédure complète de développement d'une échelle de mesure des construits marketing, proposée par Churchill (1979) a constitué une véritable avancée car elle a poussé les chercheurs à étudier les qualités psychométriques des échelles. C'est en réaction à l'utilisation systématique par les chercheurs, encouragée sans doute par les relecteurs des revues académiques, de la procédure de Churchill (1979) que Rossiter (2002) a recommandé la procédure C-OAR-SE.

Diamantopoulos (2005) et Finn & Kayandé (2005) félicitent Rossiter de sa contribution à la littérature concernant le développement d'échelle. Mais ils remettent largement en question la procédure C-OAR-SE. Il est donc nécessaire de chercher à développer une échelle de mesure en ayant recours à la procédure C-OAR-SE afin de disposer d'une application empirique, de relever les difficultés rencontrées et d'étudier les critiques. C'est l'objet de cette recherche, dans laquelle est développée une échelle de mesure de la qualité perçue des sites Web informatifs. Jusqu'à présent, n'ont été proposées que des échelles de mesure de la qualité des sites transactionnels. Il s'agit donc de répondre à la question : la procédure C-OAR-SE permet-elle de développer une échelle de mesure valide et fiable de la qualité perçue d'un site Web informatif ?

Cette recherche permettra, au plan méthodologique, d'étudier les différentes méthodes de développement d'une échelle de mesure en marketing et, au plan managérial, de proposer un outil d'analyse de la qualité d'un site Web informatif. Après une revue de littérature concernant les procédures de développement d'une échelle de mesure et la mesure de la qualité, une échelle de la qualité de service des sites Web informatifs de restaurant sera développée en utilisant la procédure C-OAR-SE puis les résultats discutés afin de répondre à la question posée.

ÉVOLUTIONS DE LA MESURE EN MARKETING

Le développement d'échelle de mesure en marketing a connu trois grandes phases (Finn & Kayandé, 2005). Avant 1979 étaient utilisées des échelles *ad hoc*, créées sans procédure logique et systématique de définition et de validation empirique du construit.

Le numéro de février 1979 de la revue *Journal of Marketing Research* a constitué une étape importante de l'amélioration de la qualité des mesures en marketing (Finn & Kayandé, 1997) puisqu'il comportait les deux articles séminaux de Peter (1979) concernant la théorie classique de la fiabilité et de Churchill (1979) présentant une procédure de mesure des construits marketing. Churchill (1979) a en fait proposé d'adopter des méthodes de construction d'échelle déjà utilisées en psychologie (Finn & Kayandé, 1997) et constate que cette discipline a été confrontée, vingt ans avant le marketing, au même problème. Et il utilise comme exemples la satisfaction du client et la satisfaction au travail.

Les *marketers*, les chercheurs et les enseignants en marketing ont donc adopté et fait adopter le paradigme churchillien pour la construction de toutes les échelles comme s'il s'agissait toujours d'étudier la psychologie d'un individu. Pourtant on ne cherche pas systématiquement, en marketing, à mesurer tel ou tel trait d'un individu. L'objet peut être également une entreprise, une publicité ou une marque (Finn & Kayandé, 1997). La procédure proposée par Churchill (1979) a donc été appliquée de manière machinale et sans réflexion, sans doute parce qu'il s'agissait d'une démarche méthodique de définition et de validation empirique du construit (Finn & Kayandé, 2005). Mais avec le temps les chercheurs semblent avoir eu le souci de maximiser les scores de fiabilité et de validité plutôt que de soigner la conceptualisation du construit. En effet, les chercheurs en sciences sociales sont souvent épris de nombres et de calcul et prennent rarement le temps de se demander : « Qu'y a-t-il derrière ces données ? » (*ibidem*). Ne consacrent-ils pas beaucoup d'efforts et de temps à travailler à partir de données obtenues à l'aide d'instruments de mesure invalides et ne finissent-ils pas par pratiquer *nolens volens* le *GIGO* (*garbage in garbage out*) (Churchill, 1979) ? La facilité de calcul du coefficient alpha à l'aide des outils informatiques aujourd'hui disponibles pourrait expliquer le succès auprès des chercheurs de cet indice de la fiabilité. Cependant l'objectivité en marketing a progressé grâce à la procédure proposée par Churchill, mais à tel point que celle-ci est considérée comme un véritable *imprimatur* (Hunt, 1993 ; Flynn & Percy, 2001).

La procédure de Churchill (1979) a bien sûr donné lieu à des analyses et des recherches qui se sont intéressées en particulier à la question de la fiabilité voire ont suggéré des amendements (Churchill & Peter, 1984 ; Peterson, 1994 ; Finn & Kayandé, 1997 ; Smith, 1999 ; Flynn & Percy, 2001). Mais aucune nouvelle procédure complète de développement d'échelle n'a été proposée.

Rossiter (2002), estimant que le recours sans discernement à la procédure de Churchill (1979) a mené à la construction d'échelles de mesure suspectes, recommande une solution théorique

et une procédure adaptée à la construction d'échelles : la procédure C-OAR-SE. Il en appelle à la raison. Il invite les chercheurs à réfléchir minutieusement à la nature du construit en commençant par donner une définition solide de celui-ci et à soigner les items et l'ancrage verbal des questions ainsi que le compte rendu des résultats. Pour autant, il ne rejette pas la procédure de Churchill : il estime qu'elle ne doit être utilisée que pour la construction de certaines échelles.

La problématique est ambitieuse puisque Rossiter (2002) affirme que la procédure C-OAR-SE est applicable dans tous les domaines des sciences sociales. L'auteur est parfaitement conscient que sa proposition, au moins au début, va susciter des réticences dans la communauté académique d'autant plus qu'il affirme que la définition et la mesure de construits majeurs en marketing sont suspectes et qu'elles ont mené à des conclusions floues et douteuses.

La procédure C-OAR-SE n'a pas encore fait l'objet de nombreuses critiques. Le numéro 22 de *International Journal of Research in Marketing* de mars 2005 comporte néanmoins deux articles à son propos (Diamantopoulos, 2005 ; Finn & Kayandé, 2005) et une courte réponse de Rossiter (2005). Finn & Kayandé (2005) proposent en fait une alternative à la procédure C-OAR-SE : la procédure s'appuyant sur la théorie de la généralisabilité multivariée (*Multivariate Generalizability Theory*) dérivée de la théorie de la généralisabilité univariée (Finn & Kayandé, 1997).

Rossiter (2002) évoque à plusieurs reprises la qualité de service. Il estime que SERVQUAL (Parasuraman, Zeithaml & Berry, 1988 ; Parasuraman, Berry & Zeithaml, 1991) n'aurait pas dû être développée en suivant la procédure de Churchill (1979). Il est donc intéressant de créer une échelle de mesure de la qualité d'un service en appliquant intégralement la procédure C-OAR-SE. Le développement exponentiel d'Internet a soulevé la question de la qualité du service perçue *on-line*. La partie suivante s'y intéresse donc.

MESURER LA QUALITÉ D'UN SITE WEB

Plusieurs auteurs ont étudié le concept de qualité perçue électronique et développé des échelles de mesure (Donthu, 2001 ; Loiacono, Watson & Goodhue, 2002 ; Wolfinbarger M. & Gilly M. C., 2003 ; Jayawardhena, 2004 ; Ming-Hui, 2005 ; Parasuraman, Zeithaml & Malhotra, 2005). Selon Zeithaml, Parasuraman & Malhotra (2002) la qualité perçue électronique repré-

sente « le degré selon lequel un site Web facilite un magasinage, un achat et une livraison efficaces et efficients des produits ou services ». Sabadie & Vernet (2003) définissent un site Web comme l'interface technologique qui soutient la rencontre de service (*service encounter*) entre l'entreprise et ses clients. Bien sûr, les dimensions de la qualité de service « traditionnelle » ne sont pas appropriées dans le domaine des services *on-line* (Bressolles & Nantel, 2005) mais il n'existe pas de consensus sur la nature des dimensions de la qualité perçue d'un site Web (Dubois & Vernet, 2001). Un outil de mesure de la qualité perçue électronique est donc nécessaire.

Bressolles (2004) a proposé Netqu@1, une échelle de mesure multidimensionnelle de la qualité perçue d'une expérience d'achat en ligne. Ses analyses exploratoires et confirmatoires ont mis en évidence cinq dimensions de la qualité perçue dans un contexte de commerce électronique : (1) qualité et quantité des informations disponibles sur le site, (2) facilité d'utilisation de celui-ci, (3) design ou caractère esthétique de l'interface, (4) fiabilité et respect des engagements et (5) sécurité-confidentialité des données personnelles et financières. L'échelle finale comporte dix-huit questions.

Les échelles de mesure de la qualité perçue *on-line* concernent le plus souvent des sites Web transactionnels ou permettant de réaliser des opérations comme des virements bancaires en ligne. Mais il paraît nécessaire de chercher à mesurer la qualité perçue d'un site Web de type particulier : le site Web informatif. Nantel, Mekki Berrada et Bressolles (2005) ont relevé ces deux types de sites. En partant de la définition de la qualité de service électronique proposée par Zeithaml, Parasuraman & Malhotra (2002), on peut définir la qualité perçue d'un site Web informatif comme le **degré selon lequel un site Web informatif facilite de façon efficiente et efficace l'information concernant les produits ou services proposés**. Reste à préciser ce qu'est un site Web informatif.

Reix (2004, p. 227) discerne trois types de positionnements stratégiques selon le rôle que les entreprises confient à Internet : outil de contact, outil transactionnel ou outil d'intégration. Pour Teo & Pian (2003), il existe cinq niveaux d'intégration d'Internet dans une organisation :

0. l'utilisation du courrier électronique ;
1. la présence sur Internet : site fournissant des informations générales sur l'entreprise ;
2. la prospection : site d'information sur l'entreprise et ses produits ou services ;

3. l'intégration d'une stratégie Internet dans la stratégie de l'entreprise ;
4. la transformation : Internet constitue un outil de création de relations et même de valeur.

De nombreuses entreprises sont encore situées au niveau 2 défini par Teo & Pian (2003) : la prospection. Pour elles, l'Internet n'est qu'un outil de contact, un simple média. Il n'est pas intégré dans la stratégie. Il peut s'agir de petites et moyennes entreprises qui manquent de ressources pour passer à un niveau supérieur, d'entreprises de secteurs traditionnels pour lesquelles l'intégration de l'Internet n'est pas envisageable sauf à changer de métier ou encore d'entreprises qui préfèrent privilégier les canaux de distribution traditionnels, par exemple dans le domaine du *B to B*. Le seul service que fournit leur site aux internautes est donc l'information sur l'entreprise et ses produits ou services. Sabadie et Vernet (2003) considèrent que toutes les entreprises qui décident de mettre en ligne un site Web créent une relation de service. On peut donc définir un site Internet informatif comme une **interface technologique qui soutient une relation de service d'information sur l'organisation et ses produits/services**. Mais le site Web d'une banque qui fournirait seulement à ses clients un relevé de leurs comptes ne serait pas un site Web informatif car l'information fournie est personnalisée. Le site informatif se différencie du site transactionnel.

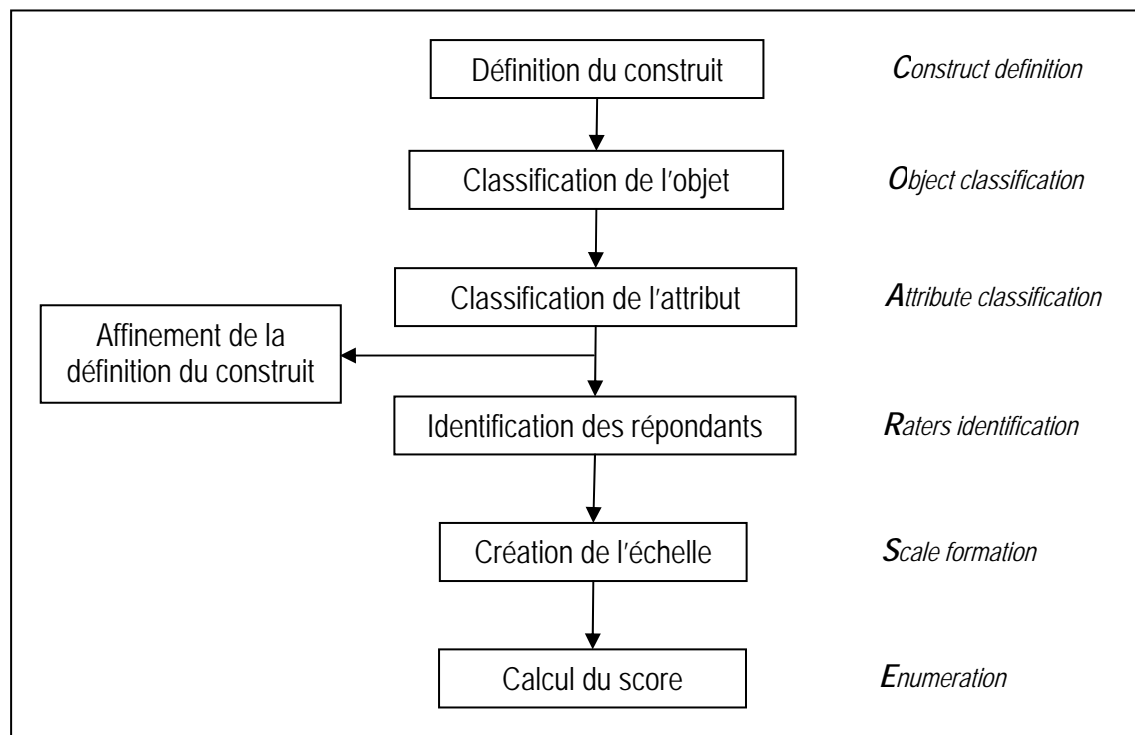
Le site Web informatif offre un service périphérique qui ajoute de la valeur au service de base (Eiglier & Langeard, 1987, p. 83). Il peut même participer à la diminution du risque perçu vis-à-vis du service de base alors que les caractéristiques de la distribution électronique renforcent le sentiment de risque (Sabadie et Vernet, 2003). Trois caractéristiques des services ont été depuis longtemps relevées : intangibilité, hétérogénéité et inséparabilité (Bateson, 1977 ; Lovelock, 1981). Mais il n'y a pas inséparabilité entre la production et la consommation dans le cas d'un site Web informatif. La production correspond à la phase de création ou de mise à jour du site et de mise en ligne. Lorsque l'internaute consomme le service, l'organisation n'a pas la possibilité d'ajuster et de personnaliser l'offre comme elle le fait parfois sur un site Web transactionnel à l'aide de *cookies*. La participation du client au service – autre caractéristique d'un service (Eiglier & Langeard, 1987, p. 31-47) – est bien réelle : il contrôle l'expérience de service en organisant par lui-même sa navigation selon ses attentes et son degré d'expertise : un site Web informatif constitue donc un élément du système de servuction (Sabadie et Vernet, 2003). Sabadie & Vernet (2003) propose comme voie de recherche la définition de normes de qualité de l'offre Internet des entreprises.

Dans la deuxième partie de cette recherche, il a donc été montré que, s'il existe déjà des échelles de mesure de la qualité perçue d'un site Web transactionnel, aucun chercheur – semble-t-il – ne s'est penché sur les sites Web informatifs, sans doute parce qu'il s'agit d'un service simple. La troisième partie présentera les étapes de la création d'une échelle de mesure de la qualité perçue d'un site Web informatif en utilisant la procédure C-OAR-SE (Rossiter, 2002).

CRÉER UNE ÉCHELLE DE MESURE À L'AIDE DE LA PROCÉDURE C-OAR-SE

La procédure C-OAR-SE de Rossiter (2002) comporte six étapes (cf. figure 1) qui vont être successivement étudiées.

Figure 1 : les étapes de la procédure C-OAR-SE (Rossiter, 2002)



Définition du construit

Pour Rossiter (2002) un construit est un phénomène d'intérêt théorique dont la définition conceptuelle doit comporter (1) l'objet et ses constituants, (2) l'attributs et ses composants et (3) les répondants afin d'indiquer comment le construit sera mesuré opérationnellement. Il donne comme exemple de construit LA QUALITÉ DE SERVICE D'IBM PERÇUE PAR

DES EXPERTS INDUSTRIELS. Il considère que la qualité de service n'est pas un construit mais un attribut qui a besoin de s'incarner dans un terrain bien précis, un objet, en l'occurrence IBM. Le construit étudié dans cette étude est la QUALITÉ PERÇUE D'UN SITE WEB INFORMATIF dont la définition conceptuelle est : LA QUALITÉ PERÇUE (attributs) PAR LES CLIENTS (répondants) D'UN SITE WEB INFORMATIF DE RESTAURANT (objet).

Diamantopoulos (2005) rappelle que, par nature, un construit est une abstraction formée à partir d'une généralisation de cas particuliers et estime que LA QUALITÉ DE SERVICE D'IBM PERÇUE PAR DES EXPERTS INDUSTRIELS n'est qu'une instance du concept de qualité de service, une définition opérationnelle et non pas conceptuelle. Ne risque-t-on pas, en précisant autant les facettes du construit, d'en revenir à l'ère pré-churchillienne, celle où l'on ne créait que des échelles *ad hoc* (Finn & Kayandé, 2005) ? N'y a-t-il pas aussi d'autres sources de variation de la mesure que l'objet, l'attribut et les répondants ? Le moment de la mesure pourrait aussi être pris en compte (*ibidem*). Et faut-il considérer que LA QUALITÉ D'UN SITE WEB INFORMATIF DE RESTAURANT PERÇUE PAR DES ÉTUDIANTS est un construit différent de LA QUALITÉ D'UN SITE WEB INFORMATIF DE RESTAURANT PERÇUE PAR DES MÉNAGÈRES DE MOINS DE CINQUANTE ANS ? Comment, avec une telle précision dans la définition des construits, pratiquer les généralisations nécessaires pour développer des concepts marketing (*ibidem*) ?

Finalement, c'est peut-être en commençant par se poser la question de l'objectif de la mesure et de l'objet à mesurer que l'on peut définir si un construit est pertinent. Dans bien des cas, implicitement, l'objet de la mesure de la qualité de service est le répondant (Finn & Kayandé, 2005). À force de considérer que l'étude du comportement du consommateur était la quintessence du marketing, les chercheurs n'ont-ils pas eu tendance à l'anthropomorphisme qui les a amenés à considérer l'orientation-marché ou la qualité de service d'une entreprise comme un trait de caractère (Rossiter, 2002) et à utiliser sans discernement la procédure de Churchill (1979) qui n'est à adopter, selon Rossiter, que lorsque les attributs sont révélateurs ?

Si l'on considère que l'objectif de la création d'échelle est de disposer d'un outil permettant de comparer les sites Web de différents restaurants, LA QUALITÉ PERÇUE PAR LES CLIENTS D'UN SITE WEB INFORMATIF DE RESTAURANT est un phénomène d'intérêt seulement managérial. Mais dans cette recherche, un autre construit est utilisé : L'INTENTION D'ACHAT dont la définition conceptuelle est : L'ENVIE DE FRÉQUENTER UN RESTAURANT (attribut) DÉCLARÉE PAR UN CLIENT (objet et ré-

pondant). Et l'objectif est d'étudier le lien entre ces deux construits. Dans ce cas, le premier construit constitue donc bien un phénomène d'intérêt théorique.

Après avoir défini le construit, les étapes suivantes consistent à classer l'objet, l'attribut et les répondants en utilisant la classification OAR (cf. Annexe 2).

Classification de l'objet

Rossiter (2002) distinguent les objets concrets, abstraits collectifs et abstraits formés. Le site Web du restaurant Michel Bras à Laguiole constitue un objet concret doté d'une adresse URL unique. Mais à quoi bon créer une échelle de mesure pour un seul site Internet ? Cela ne pourrait servir que dans le cadre d'une recherche longitudinale. L'objet étudié dans cette recherche, LES SITES WEB INFORMATIFS DE RESTAURANT, est-il un objet abstrait collectif ? Diamantopoulos (2005) trouve que la distinction entre les objets abstraits collectifs et les objets abstraits concrets est floue. Il semblerait en effet plus logique de différencier :

- Les objets concrets singuliers comme IBM ou un INDIVIDU (soi-même dans le cas d'une autoévaluation) ;
- les objets abstraits collectifs prototypiques, comme les PUBLICITÉS EFFRAYANTES ou les RESTAURANTS MCDONALD'S, constitués d'un ensemble représentatif d'objets concrets ;
- les objets abstraits collectifs inclusifs, constitués d'un ensemble exhaustif d'objets concrets, comme les APÉRITIFS À BASE DE VIN qui regroupent les vermouths, vins doux naturels, vins de liqueurs et quinquina.

Si l'on adopte cette classification, les sites Web informatifs de restaurant sont des objets abstraits collectifs prototypiques. Les cinq restaurants, dont les sites ont été étudiés par les répondants durant la phase d'expérimentation, ont d'ailleurs été choisis parce qu'ils sont représentatifs : ils appartiennent aux catégories des restaurants indépendants et restaurants à thème qui représentent respectivement 59 et 18 % de l'ensemble des restaurants français (Soubès, 2004).

Pourquoi créer une échelle de mesure spécifique pour les sites Web informatifs de restaurant ? Ne serait-il pas plus judicieux, plus académique, de créer une échelle pour les sites Web informatifs sans distinction sectorielle ? L'échelle SERVQUAL (Parasuraman, Zeithaml & Berry, 1988 ; Parasuraman, Berry & Zeithaml, 1991), par exemple, a été créée pour tous les services. Il faut se rappeler que, lors de la phase exploratoire, Parasuraman, Zeithaml & Berry

(1985), qui avait mené des analyses qualitatives auprès de cadres et de clients d'entreprises de services bancaires, de cartes de crédit, de courtage de titres et de maintenance et réparation, ont volontairement rejeté les spécificités de ces secteurs d'activité. Leur intention était donc, dès le début du processus, de créer une échelle universelle. Pourtant, n'y a-t-il pas un fossé immense entre les services fournis, par exemple, par un avocat et un restaurant rapide ? Comment leur appliquer la même échelle de mesure ? Plusieurs études ont pointé cette difficulté (Babakus & Boller, 1992 ; Carman, 1990).

Quant à l'autre construit étudié, l'INTENTION D'ACHAT, son objet est concret puisqu'il s'agit du répondant qui évalue sa propre envie de fréquenter un restaurant dont il vient de visiter le site Web.

Après avoir classé l'objet des deux construits, l'étape suivante est la classification de l'attribut.

Classification de l'attribut

D'emblée, la QUALITÉ PERÇUE semble être un attribut abstrait formé, c'est-à-dire, selon la définition de Rossiter (2002), un attribut dont les composants principaux s'agrègent pour former l'attribut. Les composants principaux sont soit des attributs concrets soit des attributs abstraits formés de second ordre eux-mêmes constitués d'attributs concrets. Rossiter donne l'exemple de la qualité de service qui, selon lui, fait partie de cette classe d'attributs.

Pour Rossiter, un attribut est abstrait si, lorsque l'on demande aux répondants de le définir, ils fournissent des réponses légèrement différentes. Les composants d'un attribut abstrait formé ont quatre caractéristiques :

1. Les composants eux-mêmes sont soit des attributs concrets soit des attributs formés de second ordre composé d'attributs concrets, de premier ordre. Les appellations de Rossiter ne sont d'ailleurs pas très claires : tantôt il parle d'attribut et tantôt de composants. Il sera considéré ici qu'un attribut formé de second ordre, comme la qualité de service, est composé d'attributs de premier ordre (fiabilité, assurance, etc.) eux-mêmes formés à partir d'attributs concrets (livraison dans les délais prévus, comportement du personnel inspirant confiance, etc.)
2. Un attribut formé doit inclure les composants principaux et non pas tous les composants possibles. Ce sont les experts qui aideront à décider des composants principaux car il est,

en pratique, impossible de recenser tous les composants contrairement à ce que pensent certains chercheurs. Un composant cité par un tiers environ des experts (Rossiter propose cette proportion sans plus d'explication) peut être considéré comme important.

3. Une fois les composants décidés par les experts, ils doivent tous apparaître dans l'échelle de mesure. Il n'est pas question de pratiquer un échantillonnage du domaine pour accroître la fiabilité de l'échelle.
4. Une échelle de mesure d'un construit comportant un attribut abstrait formé ne peut pas être unidimensionnelle, par nature. Le calcul du coefficient alpha ne sert donc à rien. Les composants sont susceptibles d'être corrélés positivement entre eux mais pas fortement. Le chercheur n'a pas à se soucier d'y parvenir.

Afin de déterminer si la QUALITÉ PERÇUE (PAR LES CLIENTS D'UN SITE WEB INFORMATIF DE RESTAURANT) est effectivement un attribut abstrait, un questionnaire a été envoyé par courriel à un échantillon de convenance de vingt-huit individus sans expertise particulière et susceptibles de visiter un site Web informatif de restaurant.

Parmi les réponses obtenues pour définir la qualité d'un site Web de restaurant, certains ont évoqué le fait qu'un tel site doit donner envie de fréquenter l'établissement :

« La qualité d'un site web de restaurant c'est son pouvoir de séduction de populations hétérogènes. »

« Bon, je dirais qu'un site web de restaurant doit très vite donner l'eau à la bouche. »

« [...] sa capacité à donner envie à celui qui le découvre d'y circuler et de consommer le produit qui y est présenté. »

Pour d'autres, le site constitue une « mise en bouche », une préparation à la fréquentation du restaurant :

« Pour moi, la qualité d'un site Web de restaurant réside dans sa capacité à rendre compte de l'ambiance dans laquelle je serai amenée à prendre place, une fois dans le resto. »

« Pour moi un site web de restaurant sera de qualité si, dès le premier coup d'œil, je m'y sens bien et qu'il me donne faim. C'est comme dans la rue : si le premier coup d'œil est séduisant, on s'arrête pour regarder la carte. »

« [...] mise en condition du futur client. »

« Je dirai qu'il s'agit d'abord de la réputation du resto de par son nom, son chef, ses menus et ses prix, mais surtout qu'il s'agit du genre d'endroit que JE fréquenterais. Cela veut dire que pour moi, la qualité du site web, c'est d'abord la qualité du resto lui-même. »

Enfin, certains semblent plus sensibles à l'esthétique du site :

« Un site web sera de qualité s'il me donne envie de faire le violent effort de cliquer une fois pour en savoir plus. »

« La qualité d'un site de restaurant c'est sa faculté d'entretenir l'attention ou de favoriser la propension d'un internaute navigant peu averti sur le réseau. »

Ainsi les perceptions de la qualité d'un site Web de restaurant sont hétérogènes. La qualité perçue est donc bien un attribut abstrait formé.

Grönroos (1984) distingue les qualités technique et fonctionnelle d'un service en s'appuyant sur Swan et Combs (1976) qui suggèrent que la performance perçue d'un produit peut être divisée en performance instrumentale et performance expressive. Pour Grönroos (1984) la qualité technique répond à la question « *Qu'est-ce que le client obtient ?* » alors que la qualité fonctionnelle répond à la question « *Comment le client l'obtient-il ?* » Ming-Hui (2005) a créé une échelle de mesure de la performance d'un site Web. Il discerne les dimensions hédoniques et utilitaires d'un site Web.

Afin de déterminer les composants de premier ordre (attributs concrets) de chacun des deux attributs de second ordre, deux groupes d'experts ont été interrogés par courriel : des webmestres pour la dimension fonctionnelle et des experts de la restauration pour la dimension technique. L'analyse du contenu des onze réponses des webmestres a révélé les éléments de la qualité perçue d'un site Web de restaurant (cf. Annexe 3).

Quant aux neuf experts en matière de restauration, l'analyse du contenu de leurs réponses a permis de déterminer dix attributs de la qualité technique (cf. Annexe 4). Bien qu'interrogés uniquement sur la qualité technique, les experts en restauration ont évoqué des éléments concernant la forme d'un site Web (photographies, liens, navigation, sobriété, police de caractères).

Identification des répondants

Rossiter distingue trois types de répondants possibles : les individus, les experts et les groupes. Dans cette étude, les répondants constituent un groupe. Ce sont des prospects internautes c'est-à-dire des clients potentiels, qui n'ont jamais fréquenté le restaurant et ayant accès à Internet. Il doit s'agir également de répondants n'ayant aucune expertise en matière d'hôtellerie-restauration pour éviter de biaiser les réponses. C'est pourquoi le questionnaire administré (cf. Annexe 1) comporte une question permettant d'éliminer les personnes expertes dans le domaine de l'hôtellerie-restauration.

Des experts peuvent également être utilisés pour évaluer un site Web. Mais Nantel, Mekki Berrada et Bressolles (2004) démontrent l'importance d'utiliser des consommateurs lorsque vient le temps de procéder à l'évaluation d'un site qui leur est destiné. Ils affirment que de nombreux sites peuvent ne pas apparaître sur la liste des meilleurs sites lorsque celle-ci est établie par une série d'experts alors que ces mêmes sites peuvent ressortir parmi les plus efficaces selon l'évaluation des consommateurs. La question du recours aux experts sera traitée *infra*.

L'objet, l'attribut et les répondants ayant été précisés, il est possible à présent d'affiner la définition conceptuelle du construit de la qualité perçue d'un site Web informatif.

Affinement de la définition du construit

Dans la procédure C-OAR-SE (Rossiter, 2002, p. 306-307), la définition du construit est affinée après l'étape de classification de l'attribut (cf. figure 1). Mais il semble plus logique de placer cette étape après l'identification des répondants puisque Rossiter (2002) considère que les répondants sont une composante intrinsèque du construit, sauf à donner raison à Diamantopoulos (2005) pour qui le répondant affecte la mesure, mais non pas le construit.

La définition conceptuelle de la QUALITÉ PERÇUE D'UN SITE WEB INFORMATIF est donc : **LA QUALITÉ TECHNIQUE ET FONCTIONNELLE (attributs) PERÇUE PAR LES PROSPECTS (répondants) D'UN SITE WEB INFORMATIF DE RESTAURANT (objet).**

Quant à L'INTENTION D'ACHAT, sa définition conceptuelle est : **L'ENVIE DE FRÉQUENTER UN RESTAURANT (attribut) DÉCLARÉE PAR DES PROSPECTS AYANT VISITÉ LE SITE WEB DE CE RESTAURANT (objet et répondant).** En suivant

la recommandation de Rossiter (2002), il a été considéré qu'un seul item est nécessaire pour cette échelle. Le questionnaire complet (cf. Annexe 1 **Erreur ! Source du renvoi introuvable.**) comporte également deux autres échelles mono-item pour savoir si le répondant est un professionnel de l'hôtellerie-restauration et s'il a déjà utilisé Internet pour choisir un restaurant.

La partie suivante explique la procédure suivie pour créer l'échelle de mesure de la qualité perçue d'un site Web informatif de restaurant.

Création de l'échelle de mesure

Selon Rossiter (2002), au cours de cette étape de C-OAR-SE il faut veiller à prendre en compte dans la question l'objet et l'attribut, à mener des interviews cognitives afin de présenter l'échelle, à choisir un format de réponse qui évite les ambiguïtés pour le répondant et à randomiser les items.

L'échelle de mesure créée (dénommée *infra* « QSR » pour **q**ualité d'un **s**ite de **r**estaurant) reprend tous les attributs concrets définis à l'aide des experts. Pour chaque question, la modalité de réponse « Je ne sais pas » a été prévue afin de suivre la recommandation de Rossiter (2002).

Les items de l'échelle Netqu@l de Bressolles (2004) ont également été administrés afin de mener une comparaison. Mais seules les questions liées aux dimensions information, facilité d'utilisation et design ont été retenues. Bressolles et Nantel (2004) ont d'ailleurs utilisé Netqu@l également pour évaluer des sites informatifs. Les dimensions fiabilité et sécurité-confidentialité de Netqu@l ne sont pas pertinentes pour un site Web informatif. Certains items sont communs à QSR et Netqu@l.

Rossiter (2002) insiste sur la nécessité d'un pré-test de l'échelle. Selon lui, la meilleure méthode est l'interview cognitive c'est-à-dire un sondage complet au cours duquel on demande à la personne interrogée de répondre à haute voix pour s'assurer que les questions sont bien comprises. Et il recommande d'utiliser la procédure de pré-test proposée par Bolton (1993) qui permet d'identifier les difficultés cognitives des répondants qui peuvent se situer à quatre niveaux de traitements cognitifs : interprétation de la question, extraction de l'information pertinente de la mémoire, traitement de l'information restituée et formation de la réponse. Il n'a pas été possible de mener à bien un pré-test en utilisant cette procédure qui s'appuie sur une analyse statistique, à l'aide du logiciel SALT (*Systematic Analysis of Language Trans-*

cripts), des éléments verbaux et non verbaux du contenu d'interviews cognitives. Le questionnaire a seulement été auto-administré auprès de trois personnes en leur demandant de le lire et de réfléchir à haute voix. Leurs remarques ont été enregistrées puis analysées a posteriori. Cela a permis de constater que certains termes n'étaient pas très clairs (« site bien adapté à l'image », « visuellement attirant », « restaurant commercial »). Le questionnaire a donc été modifié pour clarifier les termes sauf s'il s'agissait de questions de Netqu@l. Certaines questions de Netqu@l ont semblé répétitives aux individus : « L'information sur ce site est pertinente » et « L'information sur ce site est précise » ainsi que « Il est facile de recherche de l'information sur ce site » et « Il est facile de se déplacer et de trouver ce que l'on cherche sur ce site ». Pour éviter que les répondants reproduisent la même réponse, il a été décidé de disperser ces questions dans le questionnaire. D'autre part, les personnes interrogées ont indiqué qu'elles étaient troublées par les items inversés. Il leur semblait que l'auteur du questionnaire, chaque fois que l'item était inversé, leur imposait, en quelque sorte, son opinion. Il est reconnu que les items inversés ont tendance à accroître les erreurs de réponse (Carman, 1990 ; Smith, 1999).

Selon Rossiter, il ne faut pas utiliser une échelle de Likert pour les questions car ce type de format risque fort de générer des réponses ambiguës. Et en particulier, il faut absolument éviter de placer une notion d'intensité dans l'item lui-même. Par exemple, lorsque, en réponse à l'item « Je vais rarement au restaurant », une personne coche « Tout à fait en désaccord », cela signifie-t-elle qu'elle y va très souvent ou qu'elle n'y va jamais ? Rossiter recommande donc que la référence à l'intensité soit placée dans les modalités de réponse. Il a donc été choisi d'utiliser dans QSR l'échelle sémantique différentielle qui a une bonne efficacité méthodologique (Vernette, 1991) et – quand même – l'échelle de Likert parce qu'elle est plus aisément comprise (Évrard, Pras & Roux, 2003, p. 280), mais en veillant à ne pas placer d'expression de l'intensité dans l'item pour tenir compte de la remarque de Rossiter. Les recommandations de Giannelloni & Vernet (2001, p. 263) pour la construction d'échelle ont été prises en compte : sept items sur dix-sept sont inversés. Cela permet de diminuer l'effet de halo (Évrard, Pras & Roux, 2003, p. 269-271) au risque de rendre la tâche plus difficile pour les répondants et d'augmenter les erreurs de réponse (*ibidem*, p. 284).

Selon Rossiter (2002), les items devraient être randomisés afin de minimiser l'effet d'artefact. Cinq versions successives du questionnaire ont donc été mises en ligne à intervalles réguliers.

Rossiter (2002) invite donc à porter une grande attention à la rédaction des items. Ce n'est plus au logiciel de statistiques de choisir les bonnes questions mais au chercheur. Celui-ci

recouvre la primauté dans la compréhension et la mesure du concept étudié (Vernette, 2004). L'étape suivante dans la procédure C-OAR-SE est le calcul des scores après la collecte des données.

Calcul du score

Dans cette partie seront abordées la collecte des données et l'étude de la validité et de la fiabilité.

Un courriel, renvoyant vers un questionnaire en ligne, a été envoyé auprès d'un échantillon de convenance de 829 internautes dont l'adresse électronique avait été préalablement vérifiée. Il s'agissait des proches d'étudiants de licence et master première année hôtellerie-restauration. Le taux de réponse a été de 38 % soit 318 réponses dont 224 réponses utiles (les professionnels de l'hôtellerie-restauration ont été éliminés) parmi lesquelles 59 % d'hommes et 41 % de femmes et 53 % qui affirment avoir déjà utilisé Internet pour choisir un restaurant.

L'enquête *on-line* est logique pour l'étude de la qualité perçue d'un site Web. Elle présente de nombreux avantages et en particulier la rapidité de mise en œuvre, la possibilité de dispersion géographique, la diminution du biais que constituent l'enquêteur et la faiblesse du coût (Giannelloni & Vernette, 2001, p. 302). Mais elle n'est pas exempte de difficultés, en particulier en matière de représentativité de l'échantillon et donc de fiabilité (*ibid.*). Roy & Berger (2005) ont montré que le taux de réponse est beaucoup plus faible en utilisant ce mode d'administration plutôt qu'une enquête postale même si le questionnaire est inclus dans le courriel lui-même ou une loterie proposée pour inciter à répondre. Le taux de réponse constitue certainement un facteur de qualité important dans les enquêtes (Matar, 2005).

Selon Rossiter (2002), pour obtenir les scores, il faudrait calculer un indice pour chaque attribut a en additionnant les composants c . Mais afin de disposer, pour chaque répondant r , de scores standardisés, c'est une moyenne des scores pour les n composants de chaque attribut (a_t = qualité technique, a_f = qualité fonctionnelle) qui a été calculée :

$$a_t = \frac{\sum_{c=1}^n x_{tc}}{n}$$

Puis les scores totaux par attribut A_i et A_j , ont été calculés en faisant la moyenne des scores pour chaque attribut des m répondants :

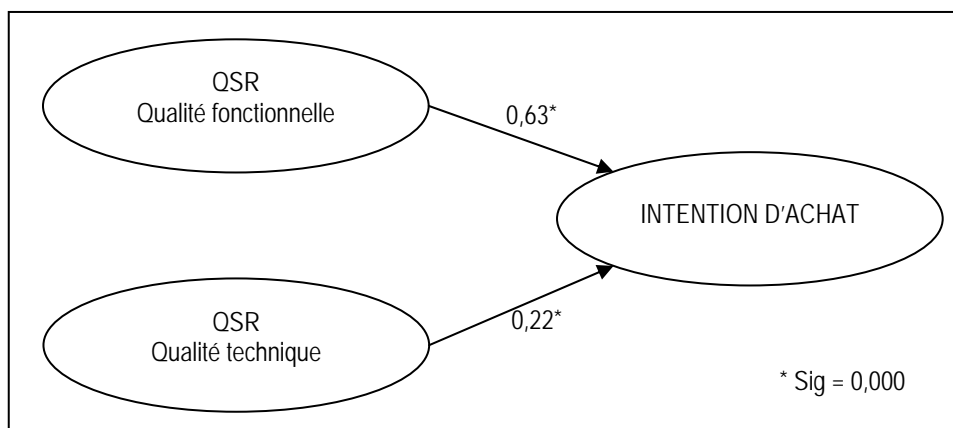
$$A_i = \frac{\sum_{r=1}^m a_{ir}}{m}$$

Le score total à l'échelle QSR est la moyenne arithmétique des attributs qualité technique et qualité fonctionnelle. Les résultats standardisés (de -1 à +1) sont présentés en annexe 5. Le site du restaurant J semble plus apprécié que celui du restaurant E. Mais – comme cela sera vu ci-après – la fiabilité de ces scores est médiocre.

Quelle est la **validité** de l'échelle QSR ? La procédure C-OAR-SE privilégie la recherche de la validité de contenu assurée par le recours aux experts pour définir les attributs. Il est donc hors de question, pour Rossiter (2002), de construire une matrice multitraits-multiméthodes pour juger des validités convergente et discriminante car la validité doit être établie indépendamment des autres échelles. Néanmoins, puisque le questionnaire administré comprend l'échelle Netqu@1 (Bressolles, 2004), dont le score a été calculé de la même manière que pour QSR, il est tentant de calculer un coefficient de corrélation sur les résultats obtenus. Celui-ci atteint 84 %. Que faire d'un tel résultat ? Comme l'affirme Rossiter (2002), il est impossible de déterminer laquelle des deux échelles est la plus valide.

Rossiter (2002) ne rejette pas l'étude de la **validité prédictive** pourvu que l'objectif ne soit pas de la maximiser en procédant à un élagage des items au détriment de la validité de contenu. Le modèle obtenu est présenté à la figure 2. Le coefficient de détermination atteint 0,58. Il faut noter l'importance beaucoup plus grande de la qualité fonctionnelle comme prédicteur de l'intention d'achat.

Figure 2 : validité prédictive QSR - Intention d'achat



Un chercheur « churchillien » qui voudrait juger de la **fiabilité** de l'échelle QSR constaterait que l'alpha de Cronbach est de 0,79 pour la dimension fonctionnelle et de seulement 0,74 pour la dimension technique, sans guère d'amélioration possible en supprimant l'un ou l'autre des items. Pour Rossiter (2002), il n'est pas question de juger de la fiabilité de la sorte. La fiabilité, qui est une estimation de la précision du score obtenu à partir d'une échelle, est liée à la taille de l'échantillon. Lorsque les répondants sont un groupe et qu'il s'agit d'un attribut formé, elle équivaut à 100 % moins l'intervalle de confiance au degré de confiance de 95 % pour une proportion de 50 %. Étant donné que la taille de l'échantillon interrogé dans cette étude concernant le site du restaurant R est de cent douze individus, l'indice de fiabilité est donc de 82 %. Comme seulement dix individus ont répondu à l'échelle QSR concernant le restaurant E, l'indice de fiabilité n'est que de 38 %.

Après avoir mené à bien le développement d'une échelle en suivant la procédure C-OAR-SE, il est possible, dans la partie suivante, de prendre du recul.

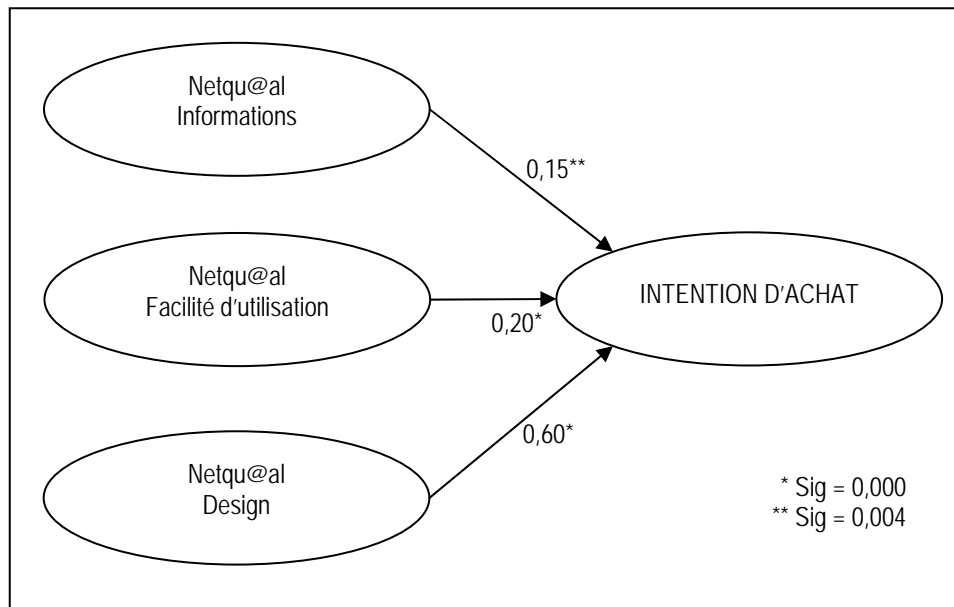
DISCUSSION

Même si Rossiter (2002) rejette l'étude de la validité convergente, il est intéressant de comparer les résultats des échelles QSR et Netqu@1 (Bressolles, 2004) pour pouvoir ensuite porter un regard critique sur la procédure C-OAR-SE. Netqu@1 a été créée en suivant la procédure de Churchill (1979). Sa fiabilité a donc été testée en commençant par une analyse en composante principale. Les dimensions de l'échelle Netqu@1 ont été retrouvées, en pratiquant une rotation Varimax, avec 69,8 % de l'information restituée. Quant aux coefficients alpha, ils atteignent 0,86 pour la dimension design, 0,84 pour la facilité d'utilisation et 0,74 pour l'information. Ces coefficients alpha sont acceptables. Faut-il vraiment s'en étonner ? Smith (1999) affirme qu'un coefficient alpha élevé peut traduire simplement la pauvreté du design de l'instrument de mesure et indiquer que les répondants perçoivent que les items sont redondants. Faut-il s'étonner d'atteindre un coefficient alpha de 0,86 pour la dimension design lorsqu'il est demandé aux répondants si un site est « joli » puis « visuellement attirant » puis s'il « fait preuve de créativité » ?

Il est bien sûr intéressant d'étudier, comme pour QSR, la validité prédictive de l'échelle Netqu@1 (cf. figure 3). Le coefficient de détermination est légèrement supérieur à celui du mo-

dèle QSR : il atteint 0,61. On notera que le design est le meilleur prédicteur de l'intention d'achat.

Figure 3 : validité prédictive Netqu@l - Intention d'achat



L'étude des validités prédictives comparées de QSR et Netqu@l permet tout d'abord de constater que la qualité fonctionnelle (dans QSR) ou le design (dans Netqu@l), c'est-à-dire la forme du site Web, joue un rôle important dans l'intention d'achat. Il est également possible désormais de confronter QSR et Netqu@l. Que compare-t-on en réalité ? D'un côté une échelle construite en suivant la procédure C-OAR-SE et spécialement adaptée aux sites Web de restaurant et d'autre part une échelle « churchillienne » qui a été conçue pour les sites Web de tous les secteurs. Si l'on se fonde uniquement sur le coefficient de détermination, on pourrait affirmer que la validité prédictive de l'échelle Netqu@l est légèrement supérieure à celle de QSR.

On sait que la procédure C-OAR-SE repose entièrement sur les arguments logiques (Rossiter, 2002) et le concours d'experts. Finn & Kayandé (2005), réfutent l'idée que des experts sont capables de simuler toutes les sources de variance et de covariance pour définir les dimensions et établir la validité de contenu d'une échelle. Diamantopoulos (2005) reconnaît que les experts peuvent fournir des informations valides et fiables dans leur champ de compétence mais qu'ils ne sont pas infaillibles. Après tout, parmi les moyens pour générer une batterie d'items, Churchill (1979) suggère une enquête de terrain auprès d'un échantillon non probabiliste de personnes ayant une opinion sur le phénomène. Et il propose, pour mesurer la satisfaction du consommateur, de mener des entretiens avec des responsables de la production, des vendeurs, des distributeurs, des consommateurs, des chercheurs en marketing, des publicitai-

res ainsi que des universitaires et des fonctionnaires disposant d'une expertise particulière dans le domaine. Mais faire reposer la validité d'une échelle uniquement sur des experts, n'est-ce pas dangereux ? Ne faudrait-il pas avoir recours à un groupe Delphi (cf. Vernet, 1994) pour améliorer le consensus parmi les experts ?

De toute façon – si l'on en croit Rossiter (2002) – comparer QSR et Netqu@1 en termes de validité prédictive ne serait pas judicieux voire irrationnel. Il prétend que l'on se trompe si le R^2 est supérieur à 0,50 car la variance inexpliquée est fondamentalement insaisissable : la part d'erreur, c'est la liberté humaine (Kenny, 1979) ! Pour Rossiter, l'étude de la validité prédictive, en plus de la validité de contenu, n'est pertinente que si les deux construits étudiés ont été établis en suivant la procédure C-OAR-SE et pour la même population.

Finn & Kayandé (2005) déplorent l'absence de validation empirique d'une échelle créée en suivant la procédure C-OAR-SE car ils estiment qu'elle révèle si la conceptualisation a abouti. C'est sans doute pour mieux mettre en avant le fait que la théorie de la généralisabilité multivariée, qu'ils présentent comme une alternative, est fondée sur une validation empirique du construit. De toute façon, il est certain que la procédure C-OAR-SE peut sembler un saut dans l'inconnu.

CONCLUSION

Le but de cette recherche était de répondre à la question suivante : la procédure C-OAR-SE permet-elle de développer une échelle de mesure valide et fiable de la qualité perçue d'un site Web informatif ? Au terme de cette étude, l'échelle QSR a été créée puis administrée et ses résultats analysés. Ses validités de contenu et prédictive ainsi que sa fiabilité ont été évaluées. Avec cet outil, il serait possible de comparer différents sites Web informatifs de restaurant. La recherche a montré la primauté de la qualité fonctionnelle d'un site Web de restaurant sur l'intention de fréquenter celui-ci. Mais le but premier était d'ordre méthodologique : il s'agissait d'appliquer complètement la procédure C-OAR-SE pour mesurer un construit marketing. Au terme de cette étude, il apparaît que la validité prédictive de QSR est un peu moins bonne que celle de Netqu@1 (Bressolles, 2004). Mais Rossiter (2002) récuse la possibilité de se livrer à une telle comparaison. De plus, en s'appuyant sur deux critiques de la procédure C-OAR-SE (Diamantopoulos, 2005 ; Finn & Kayandé, 2005), quatre questions ont été posées. Jusqu'à quel degré de précision faut-il aller dans la définition conceptuelle du construit ? La phase d'affinement de la définition conceptuelle du construit ne doit-elle pas être placée après

l'identification des répondants plutôt qu'avant ? La classification des objets est-elle pertinente ? Puisque la validité du contenu repose presque entièrement sur le concours d'experts, comment s'assurer qu'ils ne se tromperont pas ? Des adaptations de la procédure ont été suggérées.

Cette étude présente des limites d'ordre méthodologique. Les répondants ont été interrogés sur les sites de cinq restaurants seulement. Le recours à un échantillon de convenance est source d'erreur de couverture. Un échantillonnage probabiliste serait irréalisable compte tenu de l'impossibilité de disposer d'une base de sondage. On pourrait envisager la sélection d'internautes par la méthode des quotas. Le faible taux de réponse (38 %) accroît le risque d'erreur de non-réponse. Kalafatis, Sarpong & Sharif (2005) voit le développement d'échelles de mesure de concepts marketing comme un processus continu et une activité qui mérite une attention particulière en termes de réplification. La validité externe de cette recherche ne pourra être étudiée qu'après une nouvelle collecte des données sur un échantillon probabiliste plus important.

Pour Duméry (2004), « le meilleur atout de la science nouvelle consiste [...] dans l'homogénéité grandissante de ses méthodes, quels que soient les champs explorés, quelle que soit l'hétérogénéité des objets qu'elle cherche à maîtriser. » La procédure de Churchill (1979) n'a-t-elle pas semblé aux chercheurs un instrument universel assurant l'homogénéité des méthodes ? Le mérite de Rossiter (2002) est de remettre la procédure de Churchill (1979) à sa juste place et de proposer non pas une seule mais des procédures de développement d'échelles adaptées en fonction du type de construit étudié. Il revalorise le rôle du chercheur au détriment du logiciel de statistiques. C'est en cela que la procédure C-OAR-SE peut être considérée comme un souffle d'air frais (Diamantopoulos, 2005).

Il faut maintenant – comme le réclame Rossiter (2002) – créer d'autres échelles, en appliquant la procédure C-OAR-SE, pour des construits comportant d'autres types d'objets, d'attributs et de répondants.

Annexe 1 : questionnaire administré

Échelle QSR

Attributs	Attributs concrets	Item	Support
Qualité fonctionnelle	Navigation aisée	<i>Il est facile de se déplacer et de trouver ce que l'on cherche sur ce site*</i>	Likert**
	Qualités infographiques (charte graphique, couleurs)	<i>Ce site est joli</i>	Likert
	Sobriété	Ce site est... (I)***	trop chargé ⇔ sobre
	Photographies	Les photos de ce site ne sont pas attractives (I)	Likert
	Cohérence du site avec l'image du restaurant	Ce site ne correspond pas à l'image du restaurant (I)	Likert
	Lisibilité	Ce site est... (I)	illisible ⇔ lisible
	Importance page d'accueil	La page d'accueil de ce site ne me donne pas envie de visiter tout le site	Likert
Qualité technique	Prix pratiqués	Ce site ne me donne pas les informations dont j'ai besoin sur les prix pratiqués dans le restaurant (I)	Likert
	Accès	Ce site me donne des informations, pour me rendre dans ce restaurant...	Vraiment très claires ⇔ Très confuses
	Offre	<i>Ce site fournit une information détaillée sur le(s) produit(s) ou service(s) proposé(s)</i>	Likert
	Ambiance et cadre	Ce site me permet de me faire une idée de l'ambiance et du cadre du restaurant	Likert
	Notoriété	Ce site me permet de me faire une idée de la réputation du restaurant	Likert
	Horaires	Ce site ne me permet pas de savoir quels sont les horaires d'ouverture du restaurant (I)	Likert
	Appellations attractives	Les noms des plats de la carte et des menus du restaurant sont...	Très attractifs ⇔ Pas du tout attractifs
	Récence des informations	Je pense que les informations sur ce site sont régulièrement remises à jour	Likert
	Parking	Ce site ne me permet pas de savoir s'il sera facile de garer mon véhicule à proximité du restaurant (I)	Likert
	Conditions de réservation	Ce site me donne les informations dont j'ai besoin pour pouvoir réserver	Likert

* Les items en italique sont communs avec ceux de Netqu@l (Bressolles, 2004)

** Toutes les échelles de Likert sont à sept modalités de réponses de « Tout à fait d'accord » à « Pas du tout d'accord » en passant par « Ni en accord ni en désaccord »

*** (I) = Item inversé

Autres questions

Attribut	Item	Support
INTENTION D'ACHAT	Ce site me donne envie d'aller manger dans le restaurant :	Likert
NON-PROFESSIONNEL	Êtes-vous un professionnel ou un étudiant dans le domaine de l'hôtellerie-restauration ?	Oui / Non

Annexe 2 : classification OAR (Rossiter, 2002)

Objets	Attributs	Répondants
<p>Concret Objet décrit par presque tous les membres d'un échantillon des répondants de la même manière.</p>	<p>Concret Attribut décrit par presque tous les membres d'un échantillon des répondants de la même manière.</p>	<p>Individu Personne évaluant un objet (elle-même) sur un attribut personnel.</p>
<p>Abstrait collectif Ensemble d'objets concrets qui constituent une catégorie.</p>	<p>Abstrait formé Attribut formé de composants principaux qui sont soit des attributs concrets soit des attributs abstraits formés de second ordre eux-mêmes constitués d'attributs concrets.</p>	<p>Groupe Échantillon de consommateurs, de cadres, d'acheteurs, etc. utilisé le plus souvent pour évaluer un objet externe (pas eux-mêmes).</p>
<p>Abstrait formé Objet qui suggère des choses différentes aux répondants et dont des experts vont décider des composants principaux à mesurer.</p>	<p>Abstrait révélateur Trait interne ou état révélé à partir de manifestations extérieures concrètes ou composé d'attributs abstraits révélateurs de second ordre eux-mêmes composés d'un échantillon de manifestations extérieures.</p>	<p>Experts Petit groupe d'évaluateurs disposant d'une expertise concernant le construit et utilisé pour évaluer un objet externe.</p>

Annexe 3 : éléments de la qualité fonctionnelle cités par les webmestres

(Les éléments en gras ont été retenus comme des composants de la qualité fonctionnelle d'un site Web informatif de restaurant)

Éléments	Pourcentage de citation
Navigation aisée	73 %
Qualités infographiques (charte graphique, couleurs)	64 %
Sobriété	45 %
Musique pas vraiment importante	36 %
Photographies	27 %
Cohérence du site avec l'image du restaurant	27 %
Lisibilité	18 %
Importance page d'accueil	18 %
Adaptation à la cible	9 %
Sensualité virtuelle	9 %
Fond d'écran agréable	9 %
Qualité du référencement	9 %
W3C (accessibilité des sites pour les non voyants)	9 %
Site vivant	9 %

Annexe 4 : éléments de la qualité technique cités par les experts en restauration

(Les éléments en gras sont ceux qui ont été retenus comme des composants de la qualité technique d'un site Web informatif de restaurant)

Éléments	Pourcentage de citation
Prix pratiqués	89 %
Accès	67 %
Présentation de l'offre	67 %
Ambiance et cadre	56 %
Notoriété	33 %
Horaires	33 %
Appellations attractives	22 %
Récence des informations	22 %
Parking	22 %
Conditions de réservation	22 %
Photographies	22 %
Adaptation des prestations aux saisons	11 %
Contenants des bouteilles adaptés	11 %
Produits régionaux	11 %
Livre d'or	11 %
Press book	11 %
Promotions	11 %
Réservation en ligne	11 %
Police de caractères	11 %
Fond	11 %
Chercher à utiliser les possibilités d'Internet	11 %
Pertinence des liens	11 %
Navigation aisée	11 %
Sobriété	11 %

Annexe 5 : scores de l'échelle QSR

Restaurants	B	E	R	S	J	Total
n	20	10	112	50	22	214
QSR fonctionnelle	0,27	0,30	0,26	0,56	0,52	0,36
QSR technique	0,19	0,14	0,09	0,12	0,57	0,16
Qualité perçue d'un site Web de restaurant	0,23	0,22	0,17	0,34	0,55	0,26
Indice de fiabilité	56 %	38 %	82 %	72 %	58 %	

RÉFÉRENCES BIBLIOGRAPHIQUES

- Bateson J. E. G. (1977), Do We Need Service Marketing?, in *Marketing Consumer Services: New Insights*, Eiglier P., Langeard E., Lovelock C.H. and Bateson J. E. G., Cambridge (MA): ed. Marketing Science Institute, 77-115.
- Bolton R. (1993), Pretesting questionnaires: Content Analyses of Respondents'concurrent Verbal Protocols, *Marketing Science*, 12, 3, 280-303.
- Bressolles G. (2004), *La qualité de service électronique, Netqu@l : mesure, conséquences et variables modératrices*, Thèse pour l'obtention du titre de docteur ès sciences de gestion, Université de Toulouse I.
- Bressolles G., Nantel J. (2005), Proposition d'une typologie des sites web commerciaux en fonction des dimensions de la qualité de service électronique, *Actes du XXI^e congrès de l'Association française de marketing*, Nancy, 19 p.
- Carman J. (1990), Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions, *Journal of Retailing*, 66, 1, 33-55.
- Churchill G. A. (1979), A Paradigm for Developing Better Measures of Marketing Constructs, *Journal of Marketing Research*, 64-73.
- Churchill G. A., Peter J. P. (1984), Research Design Effects on the Reliability of Rating Scales: A Meta-Analysis, *Journal of Marketing Research*, 21, 4, 360-375.
- Diamantopoulos A. (2005), The C-OAR-SE procedure for scale development in marketing: A comment. *International Journal of Research in Marketing*, 22, 1, 1-9.
- Donthu N. (2001), Does Your Web Site Measure Up? *Marketing Management*, Nov/Dec, Vol. 10, Issue 4, pp. 29-32.
- Dubois P.-L., Vermette E. (2001), Contribution et pistes pour la recherche en E-marketing, *Recherche et Applications en Marketing*, 16, 3, 1-8.
- Duméry H., Sciences – Introduction, in *Encyclopædia Universalis* [Cédérom], version 10, Paris : Encyclopædia Universalis France SA, 2004.
- Eiglier P., Langeard E. (1987), *Servuction Le marketing des services*, Paris : McGraw-Hill.
- Évrard Y., Pras B., Roux E. (2003), *Market*, 3^e édition, Paris : Dunod.
- Finn A., Kayandé U. (1997), Reliability Assessment and Optimization of Marketing Measurement, *Journal of Marketing Research*, 34, 262-275.
- Finn A., Kayandé U. (2005), How Fine Is C-OAR-SE? A Generalizability Theory Perspective on Rossiter's Procedure? *International Journal of Research in Marketing*, 22, 1, 11-21.
- Flynn L. R., Pearcy D. (2001), Four Subtle Sins in Scale Development: Some Suggestions for Strengthening the Current Paradigm, *International Journal of Market Research*, 43, 4, 409-423
- Giannelloni J.-L., Vermette E. (2001), *Études de marché*, Paris : Librairie Vuibert.
- Grönroos Ch. (1984), A Service Quality Model and Its Marketing Implications, *European Journal of Marketing*, 18, 4, 36-44.
- Hunt S. (1993), Objectivity in Marketing Theory and Research, *Journal of Marketing*, 57, 2, 76-81.

- Jacoby J. (1978), Consumer Research : A State of the Art Review, *Journal of Marketing*, 42, 2, 87-96.
- Jayawardhena Ch. (2004), Measurement of Service Quality in Internet Banking: The Development of an Instrument, *Journal of Marketing Management*, 20, 1/2, 185-207.
- Kalafatis, S. P., Sarpong Jr. S., Sharif K. J. (2005), An Examination of the Stability of Operationalisations of Multi-Item Marketing Scales, *International Journal of Market Research*, 3, 47, 3, 255-266.
- Kenny D. A. (1979), *Correlation and causality*, New York: Wiley.
- Loiacono E. T., Watson R. T., Goodhue D. L. (2002), WebQual™: A Measure of Web Site Quality., *Lycos*, Waltham, Massachusetts, July 2.
- Lovelock Ch. H. (1981), Why Marketing Needs To Be Different for Services, in *Marketing of Services*, J. Donnelly and W. George Eds, Chicago: American Marketing, 5-9.
- Matar A. (2005), Les déterminants du comportement de réponse à un questionnaire : une comparaison des méthodes d'administration postale versus Internet., *Actes du XXI^e congrès de l'Association française de marketing*, Nancy, 27 p.
- Ming-Hui H. (2005), Web Performance Scale, *Information & Management*, 42, 6, 841-852.
- Nantel J., Mekki Berrada A., Bressolles G. (2005), L'efficacité des sites Web : quand les consommateurs s'en mêlent, *Revue internationale de gestion*, 30, 1, 16-23.
- Parasuraman A., Zeithaml V. A., Berry L. L. (1985), A Conceptual Model of Service Quality and Its Implications for Future Research, *Journal of Marketing*, 49, 4, 41-50
- Parasuraman A., Zeithaml V. A., Berry L. L. (1988), SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality, *Journal of Retailing*, 64, 1, 12-40.
- Parasuraman A., Berry L. L., Zeithaml V. A. (1991), Refinement and Reassessment of the SERVQUAL Scale, *Journal of Retailing*, 67, 4, 420-450.
- Parasuraman A., Zeithaml V. A., Malhotra A. (2005), E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality, *Journal of Service Research*, Feb., Vol. 7, Issue 3, pp. 213-233.
- Peter J. P. (1979), Reliability: A Review of Psychometric Basics and Recent Marketing Practices, *Journal of Marketing Research*, 16, 6-17.
- Peterson R.A. (1994), A Meta-Analysis of Cronbach's Coefficient Alpha, *Journal of Consumer Research*, 21, 2, 381-391.
- Reix R. (2004), *Systèmes d'information et management des organisations*, Paris : Librairie Vuibert.
- Rossiter J. R.(2005)., Reminder: a Horse Is a Horse *International Journal of Research in Marketing*, 22, 1, 23-25.
- Rossiter J. R. (2002), The C-OAR-SE Procedure for Scale Development in Marketing, *International Journal of Research in Marketing*, 19, 4, 305-335.
- Roy A., Berger P. D. (2005), E-Mail and Mixed Mode Database Surveys Revisited: Exploratory Analyses of Factors Affecting Response Rates, *Journal of Database Marketing & Customer Strategy Management*, 12, 2, 153-171.

- Sabadie W., Verneette E. (2003), La servuction on line : points communs et spécificités face à la servuction traditionnelle, *Actes de la 2^e journée de recherche sur le e-marketing*, Nantes, 22 p.
- Smith A. M. (1999), Some Problems When Adopting Churchill's Paradigm for the Development of Service Quality Measurement Scales, *Journal of Business Research*, 46, 2, 109-120.
- Soubès S. (2004), État des lieux, *L'Hôtellerie-Restaurant Magazine*, n° 2876, 24-29.
- Swan J. E., Combs L. J. (1976), Product Performance and Consumer Satisfaction: A New Concept, *Journal of Marketing*, 25-33.
- Teo T. SH, Pian Y. (2003), A Contingency Perspective on Internet Adoption and Competitive Advantage, *European Journal of Information Systems*, 12, 78-92.
- Verneette E. (1991), L'efficacité des instruments d'études : évaluation des échelles de mesure, *Recherche et Applications en Marketing*, VI, 2, 43-64.
- Verneette E. (1994), La méthode Delphi, une aide à la prévision marketing, *Décision Marketing*, 1, 97-101.
- Verneette E., (2004), Les études qualitatives à la recherche du sens perdu, *Décisions Marketing*, 36, 7-9.
- Wolfenbarger M., Gilly M. C. (2003), eTailQ: Dimensionalizing, Measuring and Predicting Etail Quality, *Journal of Retailing*, Volume 79, Number 3, pp. 183-198.
- Zeithaml V.A., Parasuraman A., Malhotra A. (2002), Service Quality Delivery Through Web Sites: A Critical Review of Extant Knowledge, *Journal of the Academy of Marketing Science*, 30, 4, 362-375.