

EVALUATION EX ANTE DE SYSTEMES D'ETIQUETAGE NUTRITIONNEL GRAPHIQUE SIMPLIFIE

RAPPORT FINAL DU COMITE SCIENTIFIQUE

Le comité de pilotage de l'évaluation *ex ante*, en conditions réelles d'achat, de divers systèmes graphiques d'étiquetage nutritionnel simplifié avait, par lettre du 12 février 2016, donné au comité scientifique, dont la composition actualisée figure en annexe 1, mission d'établir le protocole de l'intervention à conduire – ce qui a été l'objet de son rapport du 11 avril 2016 -, d'en suivre l'exécution et enfin de lui faire rapport sur les résultats obtenus.

C'est l'objet du présent document.

La conduite opérationnelle de l'intervention, et donc la mise en œuvre du protocole défini par le comité scientifique, avait été dévolue, conformément au protocole de gouvernance de l'évaluation, au Fonds français pour l'alimentation et la santé (FFAS). Le FFAS a établi un rapport d'exécution qu'il a transmis au comité de pilotage et au comité scientifique.

Le traitement des données d'achat a été réalisé par une équipe dont la composition figure en annexe 2. Les membres de cette équipe indépendante ont été choisis nominativement par le ministère de la santé avec l'accord du comité scientifique, conformément à ce qui était prévu dans le protocole.

Le comité scientifique avait enfin souhaité que, de manière complémentaire à l'intervention portant sur les comportements d'achat, une étude déclarative réalisée sous forme de questionnaire étudie notamment l'aptitude des divers systèmes d'étiquetage à attirer l'attention des consommateurs et à être correctement compris. Le FFAS a chargé le CREDOC de cette étude.

Le comité scientifique, considérant la bonne qualité de ces trois rapports estime qu'ils constituent en eux-mêmes le résultat de l'intervention dont il avait défini le protocole. Dès lors ils figurent tous trois en annexes du présent rapport. Il apparaît au comité qu'il est inutile d'établir un résumé exhaustif de ces trois rapports, ou de les compléter par ses propres analyses. Ceci d'autant que le temps qui lui est imparti est très court.

Il se bornera donc, au cours des brefs développements qui vont suivre, à présenter ces rapports et à appeler, le cas échéant, l'attention du comité de pilotage sur ceux de leurs aspects qui lui paraissent devoir être particulièrement signalés.

A- Le déroulement de l'intervention

Le rapport d'exécution établi par le FFAS figure en annexe 3. Il est lui-même accompagné de quatre annexes qui, en raison notamment du poids de l'une d'entre elles, ne sont pas jointes, mais qui peuvent être consultées et téléchargées sur le site du FFAS.

Ce rapport rend compte de manière détaillée du déroulement de l'ensemble de l'intervention.

- 1- On en retiendra avant tout que ce déroulement s'est effectué de façon satisfaisante, grâce à la mobilisation du porteur opérationnel de projet et des prestataires qu'il a recrutés, au premier rang desquels l'agence LinkUp qui a assuré, sous le contrôle du FFAS, la coordination des prestataires techniques.

Aucun des 60 magasins sélectionnés à l'issue de la randomisation effectuée, conformément au protocole, par le comité scientifique lui-même, n'a fait défection et, après un démarrage un peu difficile pour un petit nombre d'entre eux (3 sur les 40 dans lesquels étaient étiquetés des produits) l'opération s'y est partout déroulée de façon au moins acceptable au regard des critères exigeants de qualité qui avaient été fixés.

Le déroulement satisfaisant tient également à ce que la proportion des produits étiquetés dans les catégories retenues a été importante : 79% pour les produits traiteur frais, 86% pour le pain et les viennoiseries et 67% pour les plats cuisinés épicerie.

Le comité scientifique retient enfin du rapport que le nombre d'incidents d'étiquetage a été plutôt faible. Le tableau des événements magasins (annexe C au rapport du FFAS) recense ces incidents de manière exhaustive, par jour et par magasin. Ils consistent le plus souvent dans le fait que l'étiquetage, ce jour-là, n'avait pas pu être complètement terminé faute de temps.

- 2- L'intervention a été de grande ampleur.

Pour s'en tenir aux chiffres principaux, 1298 produits ont été étiquetés, et 1 748 503 étiquettes apposées, correspondant à environ autant d'achats.

- 3- Le comité scientifique note que le contrôle qualité de l'intervention, dont il avait défini les principes d'organisation dans le protocole, a été mis en œuvre de manière rigoureuse et efficace.

L'équipe d'auditeurs, dont la grille d'évaluation figure au 8 de l'annexe A du rapport, a pu contrôler le bon fonctionnement des contrôles de premier niveau et garantir ainsi la pertinence de l'étiquetage (les erreurs d'étiquetage ont été rares), la disponibilité généralement très satisfaisante du dispositif d'information des consommateurs ainsi que la traçabilité d'ensemble de l'opération.

Les auditeurs ont toutefois noté, dans leurs conclusions, au titre de l'égalité de traitement, que les magasins dans lesquels étaient testés les systèmes dits analytiques (Nutri-Couleurs et Nutri-Repère) avaient connu, du fait de la très grande variété des étiquettes davantage d'incidents d'étiquetage, soit pour rupture de stock de certaines étiquettes, soit pour de trop longs délais d'étiquetage. Le traitement a mesuré les conséquences éventuelles de cette incomplétude.

- 4- Le comité scientifique a enfin pris connaissance avec un grand intérêt des résultats des études qualitatives préparatoires conduites pour optimiser les systèmes d'étiquetage ainsi que les brochures accompagnatrices de compréhension des systèmes.

Ces résultats, qui lui paraissent d'une grande richesse, sont donnés au 3 de l'annexe A au rapport du FFAS.

B- L'observation des comportements d'achat

Le rapport de l'équipe de traitement des données quantitatives recueillies figure en annexe 4.

L'exploitation des données a été confiée à quatre chercheurs appartenant à trois institutions différentes afin de garantir l'indépendance du traitement des données (cf. annexe 2).

Un sous-groupe du comité scientifique a vérifié le déroulement de l'exploitation des données, tant pendant la phase longue de récupération des données (avec de nombreux aller-retour vers les services statistiques des distributeurs), et de la mise en place de la base (ces données d'une très grande richesse devant à terme être accessible aux chercheurs compétents en faisant la demande), que pendant la phase d'exploitation et la phase de restitution des résultats. Les résultats finaux, la base de données et l'ensemble des programmes ont été livrés au sous-groupe du conseil scientifique, qui a donc pu vérifier la reproductibilité et la validité des conclusions de l'équipe de traitement.

L'équipe de traitement a évalué comparativement l'impact des différents systèmes d'étiquetage nutritionnel sur différents indicateurs basés sur le score FSA. D'un point de vue technique, l'évaluation de l'impact des logos s'appuie sur la méthode des différences premières (comparaison la même année entre les scores des magasins avec étiquetage et les magasins témoins) et celle par différences secondes (évaluation des différences entre deux périodes similaires (entre 2015 et 2016) des magasins étiquetés par rapport aux différences observées dans les magasins témoins). Comme le rappelle le rapport, les deux méthodes sont scientifiquement justifiées et reposent sur des hypothèses d'identification différentes qui ont conduit l'équipe de traitement à leur analyse séquentielle. Deux types de données ont également été traités : les données agrégées des ventes par magasins et les données individuelles des porteurs de carte de fidélité. La comparaison des résultats entre ces deux sources montre que l'on peut raisonnablement penser qu'il n'y a pas de biais dus à la sélection des porteurs de carte de fidélité (dans la mesure où les effets mis en évidence sont du même ordre de grandeur et vont globalement tous dans le même sens selon les données).

N'ayant pu obtenir d'information spécifique sur les populations effectivement défavorisées (hors l'indicateur de zone des magasins), l'équipe de traitement a construit un indicateur de sélection des personnes achetant les produits les moins chers, basé d'abord sur les dépenses moyennes, puis sur les prix moyens d'achat les plus faibles (quantile d'ordre 40% des individus payant les prix moyens pondérés par les quantités les plus faibles). L'ensemble ainsi défini est beaucoup plus large que celui des ménages défavorisés, mais outre que les ménages les plus contraints y figurent à coup sûr, l'analyse de cette population permet de savoir si les logos nutritionnels sont pris en compte par les ménages les plus sensibles aux prix.

Les évaluations ont d'abord été effectuées de manière brute (sans variables de contrôle) sur l'indicateur privilégié par le comité scientifique à savoir le score FSA pondéré par les calories en utilisant la méthode des différences premières et secondes, puis en ajoutant des variables de contrôle d'autre part. L'équipe de traitement a en particulier contrôlé systématiquement les effets des vagues d'étiquetage, les effets de périodes par quinzaine (pour contrôler notamment l'impact des premières périodes de mise en œuvre de l'expérience), l'effet des prix moyens et des taux d'erreur d'étiquetage. De manière systématique, l'équipe de traitement a également complété son analyse en utilisant d'autres critères d'évaluation comme l'avait suggéré le comité scientifique. Des analyses similaires ont ainsi porté systématiquement sur le score FSA pondéré par les quantités, le score FSA moyen par transaction (sans repondération, intitulé « transaction » dans le rapport et le tableau ci-dessous), le score FSA total du panier, l'apport calorique total en utilisant les méthodes des simples et doubles différences ainsi que sur l'impact sur le score FSA de chaque rayon. Sans a priori, ces méthodes ont été mises en œuvre respectivement sur la population totale des porteurs de carte et celle des individus achetant aux prix les plus bas.

Des résultats secondaires ont été également détaillés dans le rapport pour l'ensemble des produits des rayons étiquetés ou non, pour les produits étiquetés seulement, et pour les produits non étiquetés. Du fait de l'absence de certaines informations nutritionnelles pour les produits non étiquetés, le comité scientifique considère que les résultats sur les produits étiquetés sont plus pertinents et doivent à ce titre être analysés en priorité même si les effets sur les produits non étiquetés peuvent également être intéressants.

Les principaux résultats obtenus sur les produits étiquetés sont synthétisés dans le tableau suivant. Sont exclus de ce tableau les modèles annexes qui ne présentent aucun effet significatif afin de conserver le caractère systématique de l'analyse et de manière à procéder à une méta-analyse des résultats.

Ce tableau fait apparaître :

- le type de produit analysé, agrégé ou par rayon,
- le type de données (clients porteur de carte, clients achetant les produits les moins chers (Clients MC) selon le critère retenu précédemment),
- le score analysé (score FSA pondéré par les quantités (poids) ou les calories, et pour les dernières lignes le score FSA total ou l'apport calorique total),
- la méthode d'estimation retenue (simple ou double différence),

- l'introduction éventuelle de variables de contrôle (prix et/ou taux d'erreur d'étiquetage),

- les significativités des quatre systèmes selon le codage suivant des effets:

(--) baisse de l'indicateur (= amélioration) statistiquement significative à 99%

(-) baisse de l'indicateur (= amélioration) statistiquement significative à 95%

= résultat non statistiquement significatif

(++) hausse de l'indicateur (= détérioration) statistiquement significative à 99%

(+) hausse de l'indicateur (= détérioration) statistiquement significative à 95%

La dernière colonne réfère au numéro de tableau et aux détails des résultats du rapport de l'équipe de traitement.

Afin de pouvoir comparer les systèmes entre eux nous indiquons sur la dernière ligne le nombre de fois ou le système d'étiquetage a un effet significatif respectivement (--) ou (-), ou un effet contraire au sens attendu (+) ou (++) , ceci afin de donner un indicateur agrégé de l'efficacité des systèmes sur différents critères.

Produits	Fichier	Analyse	Pondération	Modèle	NC	NR	NS	SENS	Numéro de Tableau du rapport
----------	---------	---------	-------------	--------	----	----	----	------	------------------------------

Résultats sur la densité du score FSA

Clients	Agrégé	poids	Simple différence	(--)	(--)	(--)			Table 3.5
Clients	Agrégé	calories	Simple différence	(--)	(--)	(--)			Table 3.5
Clients MC	Agrégé	calories	Simple différence	(--)	(--)				Table 3.7
Clients	Agrégé	calories	double différence						Table 3.8
Clients	Agrégé	calories	double différence - vague1-2		(-)	(-)			Table 3.8
Clients	Agrégé	poids	double différence						Table 3.8
Clients	Agrégé	poids	double différence - vague1-2		(-)				Table 3.8
Clients	Agrégé	calories	double différence				(+)		Table

MC								3.9
Clients				double différence - vague1-2				Table 3.9
MC	Agrégé	calories						Table 3.9
Clients								Table 3.9
MC	Agrégé	poids		double différence			(+)	Table 3.9
Clients				double différence - vague1-2				Table 3.9
								Table 3.10
Clients	Quinzaine	calories		Simple différence	(--)	(--)	(--)	Table 3.10
Clients	Quinzaine	poids		Simple différence	(--)	(--)	(--)	Table 3.10
Clients	Quinzaine	calories		double différence				Table 3.11
Clients	Quinzaine	poids		double différence				Table 3.11
								Table 3.18
Clients	Transaction			Simple différence	(--)	(--)	(--)	(--)
Clients								Table 3.18
MC	Transaction			Simple différence		(-)	(-)	Table 3.18
								Table 6.1
Clients	Transaction			Simple différence	(--)	(--)	(--)	(--)
Clients								Table 6.1
MC	Transaction			Simple différence			(-)	Table 6.1
								Table 3.19
Clients	Transaction			double différence	(-)	(-)	(--)	Table 3.19
Clients								Table 3.19
MC	Transaction			double différence			(--)	Table 3.19
								Table 6.2
Clients	Transaction			double différence	(-)	(-)	(--)	Table 6.2
Clients								Table 6.2
MC	Transaction			double différence		(-)	(--)	Table 6.2
Traiteur								Table 3.22
Frais	Clients	Agrégé	calories	Simple différence		(++)	(--)	Table 3.22
Traiteur								Table 3.22
Frais	Clients	Agrégé	poids	Simple différence		(++)	(--)	Table 3.23
Traiteur								Table 3.23
Frais	Clients	Agrégé	calories	double différence	(--)		(--)	Table 3.23
Traiteur								Table 3.23
Frais	Clients	Agrégé	poids	double différence	(--)	(-)	(--)	Table 3.26
Viennoiserie								Table 3.26
Viennoiserie	Clients	Agrégé	calories	Simple différence	(-)			Table 3.26
Viennoiserie								Table 3.27
Viennoiserie	Clients	Agrégé	calories	double différence			(--)	(-)
Viennoiserie								Table 3.27
Viennoiserie	Clients	Agrégé	poids	double différence			(--)	(-)
Viennoiserie								Table 3.27
Plats	Clients	Agrégé	calories	Simple différence	(-)			Table 3.27

cuisinés									3.30
Plats									Table
cuisinés	Clients	Agrégré	poids	Simple différence	(-)				3.30
Plats									Table
cuisinés	Clients	Agrégré	calories	double différence	(--)	(--)	(-)		3.31
Plats									Table
cuisinés	Clients	Agrégré	poids	double différence	(--)	(--)	(-)		3.31
Pains et Pat									Table
Ind	Clients	Agrégré	calories	Simple différence	(++)	(-)	(--)	(-)	3.34
Pains et Pat									Table
Ind	Clients	Agrégré	poids	Simple différence	(++)	(-)	(--)	(-)	3.34
Pains et Pat									Table
Ind	Clients	Agrégré	calories	double différence	(+)		(+)		3.35
Pains et Pat									Table
Ind	Clients	Agrégré	poids	double différence	(+)		(+)		3.35

Résultats sur la quantité totale du score FSA

Clients	Agrégré		simple différence	(--)	(--)	(--)		Table
								4.2
			double différence		(-)	(--)	(-)	Table
								4.3

Résultats sur la quantité totale de calories

Clients	Agrégré		simple différence	(--)	(--)	(--)		Table
								4.5
			double différence		(--)	(--)		Table
								4.6

Nombre d'analyses avec effets négatifs (favorables)	19	14	26	16
Nombre d'analyses avec effets positifs (défavorables)	4	2	0	4

Les principales conclusions que l'on peut extraire des résultats obtenus sont les suivantes:

- 1- Si l'on s'en tient uniquement à l'utilisation du score FSA pondéré par les calories préconisé par le comité scientifique dans son protocole (sur toute la période et sur tous les produits étiquetés pour les rayons confondus), on note que, en différence première,

trois systèmes sur quatre sont significativement efficaces dans la perspective de la réduction du score FSA (voir tableau 3.5 du rapport du groupe de traitement). Le Nutri-Score engendre une réduction du score FSA de -0.267, SENS de -0.198 et Nutri-Couleurs de -0.233, sur un score moyen de l'ordre de 6. Seul Nutri-Repère a un impact non significatif et positif (+0.027) indiquant une dégradation du score nutritionnel. Trois des systèmes testés ont donc un impact sur les achats en situation réelle dans le sens attendu d'un meilleur score FSA.

- 2- La méthode des différences secondes, critère qui avait été choisi par le comité scientifique pour classer les systèmes, ne permet pas de départager clairement entre Nutri-Score, SENS et Nutri-Couleurs, dans la mesure où aucun des impacts en différence seconde n'est significativement différent de zéro. Même si l'effet de Nutri-Score est statistiquement supérieur à celui de SENS, la différence entre Nutri-Score et Nutri-Couleurs n'est pas statistiquement significative. Il n'est donc pas possible de classer ces trois systèmes avec ce seul critère. En effectuant les analyses par période de deux semaines (afin d'éliminer d'éventuels biais dus aux premières semaines de mise en place de l'expérience), on met en évidence là encore la significativité des mêmes 3 systèmes en différence première (tableau 3.19)
- 3- En ne considérant que la sous-population des acheteurs des produits les moins chers, le Nutri-Score a l'impact le plus important (-0.312) suivi de Nutri-Couleurs (-0.223) en seconde position, en différence première (tableau 3.9). SENS s'avère le plus mauvais système. En différence seconde, SENS entraîne même une dégradation statistiquement significative du score FSA sur cette population (tableau 3.15).
- 4- L'analyse disjointe entre produits étiquetés et non étiquetés montre que la présence des logos a aussi un effet important sur les produits non étiquetés des rayons concernés par l'expérimentation. Plusieurs explications sont proposées par l'équipe de traitement notamment la possibilité de comparer les produits non étiquetés par référence à ceux qui le sont, mais aussi un effet de défiance vis-à-vis des produits non étiquetés qui peut entraîner un effet de substitution. On peut donc sans doute s'attendre à un impact des logos sur l'ensemble d'une famille de produits dès lors qu'un nombre significatif de produits est étiqueté. On observe en outre que les produits non étiquetés ont un score FSA moyen significativement moins bon que les produits étiquetés, ce qui suggère que la qualité nutritionnelle a eu un impact sur la décision de participer à l'expérimentation.
- 5- L'estimation des modèles avec introduction des prix et du pourcentage d'erreur d'étiquetage en variables de contrôle montre d'une part que les quatre systèmes semblent donner des baisses significatives du score FSA. Toutefois, seul Nutri-Score engendre une baisse significative systématique du FSA pour les acheteurs des produits les moins chers (en simple et en double différence) avec ces contrôles.
- 6- L'analyse du critère unique fondé sur le FSA pondéré par les calories ne permettant pas de départager les différents systèmes de manière convaincante, l'équipe de traitement a utilisé les analyses précitées sur de nombreux scores pour voir s'il était possible de départager les systèmes en fonction de leurs effets sur des populations ou des lignes de produits spécifiques. Il est intéressant de constater que les trois systèmes Nutri-Score, SENS et Nutri-Couleurs réduisent significativement l'apport calorique global ou le score FSA total. Toutefois l'analyse de l'ensemble des modèles montre

que seul Nutri-score implique systématiquement soit une baisse non significative soit une baisse significative des indicateurs choisis (voir les deux dernières lignes du tableau précédent) sur l'ensemble de modèle. Les résultats sont plus contrastés sur les autres systèmes.

- 7- Les effets par grandes familles de produits testés montrent une plus grande variabilité. Ce sont toutefois sur les catégories «Traiteurs frais», «Plats préparés» et «Viennoiseries» que les logos ont le plus grand impact. C'est avec Nutri-Score que l'on constate les résultats les plus significatifs sur «Traiteurs frais» et «Viennoiseries» tandis que Nutri-Couleurs et dans une moindre mesure Nutri-Repère sont plus efficaces sur «Plats préparés».

Le tableau récapitulatif permet de comparer les différents systèmes sur ces tests d'efficacité partielle ou spécifique. Il en ressort que le système Nutri-Score apparaît comme le plus efficace. En particulier, il ne produit jamais de résultat dans le sens opposé à celui souhaité (diminution du score FSA) et semble avoir plus d'impact que les autres sur les populations achetant les produits moins chers.

C- L'étude qualitative du CREDOC

Le rapport du CREDOC figure en annexe 5

Le CREDOC, en complément à l'expérimentation comportementale «grandeur nature», a conduit un travail déclaratif sous la forme de questionnaires face-à-face réalisés dans 20 magasins avant et pendant l'expérimentation, puis, sous la forme de questionnaire en ligne, après l'expérimentation.

L'enquête face-à-face avant l'expérimentation a eu lieu en septembre 2016 dans 20 magasins correctement échantillonnés et a porté sur 1844 répondants eux aussi correctement échantillonnés compte tenu des contraintes de l'exercice. L'enquête face-à-face pendant l'expérimentation a eu lieu en novembre et décembre 2016 dans 20 magasins et a porté sur 1737 répondants. L'enquête en ligne après l'expérimentation a porté sur 715 répondants. Les mêmes commentaires s'appliquent à propos des échantillonnages.

Les enquêtes ont été menées, à la connaissance de Comité Scientifique, dans les règles de l'art, de même que l'exploitation statistique des résultats.

Le rapport final réalisé par le CREDOC suite à ces enquêtes est complet. Il rend compte du travail réalisé qui est conforme à ce qui avait été demandé par le Comité Scientifique.

Les objectifs de cette enquête sont relatifs à quatre sujets : (1) la capacité des systèmes et de chacun d'entre eux à attirer l'attention, (2) la capacité des systèmes et de chacun d'entre eux à modifier les croyances en nutrition, (3) la compréhension active des systèmes d'étiquetage en termes d'aptitude à les utiliser pour choisir, (4) le niveau de satisfaction.

- 1- Les consommateurs enquêtés déclarent que les systèmes Nutri-Score et SENS sont les plus visibles (vus, respectivement, par 48% et 46% des interrogés – une différence statistiquement significative), suivis par Nutri-Repère et Nutri-Couleurs

(respectivement 37% et 31%, deux scores statistiquement inférieurs à celui de Nutri-Score).

- 2- L'impact des systèmes d'étiquetage sur les connaissances en nutrition (capacité à identifier quel aliment d'une paire est le plus riche en lipides, glucides, ou protéines) n'est pas significatif. De même, on ne constate pas non plus de différences dans l'importance accordée à la nutrition telle que déclarée par les répondants.
- 3- La compréhension de logos est mesurée objectivement par la capacité à mieux classer la qualité nutritionnelle de trois produits différents à l'aide des logos. Ce test est décisif dans cette enquête. Les performances à ce test sont bonnes pour Nutri-Score et pour SENS : 92,4% des individus ont bien classés les trois produits avec Nutri-Score, 84,1% avec SENS, une différence statistiquement significative. En revanche, Nutri-Couleurs (29% de bons classements) et Nutri-Repère (16,6%) n'ont pas d'impact significatif sur la capacité à bien classer les trois produits, voire même ont un impact contraire à l'objectif poursuivi quand on se focalise sur les catégories ayant un diplôme inférieur au baccalauréat (c'est à dire que ces systèmes réduisent le nombre de consommateurs réussissant à ordonner correctement les trois produits). Dans ce test, les systèmes Nutri-Couleurs et Nutri-Repère déroutent donc les consommateurs dans l'usage des systèmes comme aide à la décision.

Entre les deux systèmes synthétiques, le design et les couleurs de Nutri-Score permettent plus facilement d'ordonner les catégories entre elles que le design et les couleurs de SENS. Même en l'absence de texte (lettres ou recommandations de consommation), 96% des consommateurs ont ordonné correctement les étiquettes Nutri-Score, alors que seuls 79% ont pu le faire pour SENS, une différence statistiquement significative.

- 4- Les acheteurs apprécient Nutri-Score et SENS pour leur aptitude à leur apporter une information à l'échelle de l'aliment, permettant de les guider vers les aliments à favoriser ou à limiter. Nutri-Couleurs et Nutri-Repère sont appréciés pour permettre de suivre les recommandations de santé ou équilibrer son alimentation.

Au total, le Comité scientifique retient que les quatre systèmes en lice peuvent être divisés en deux groupes. Les systèmes synthétiques – Nutri-Score et SENS – d'un côté qui qualifient la qualité nutritionnelle à l'échelle du produit, et les systèmes analytiques – Nutri-Repère et Nutri-Couleurs – de l'autre qui qualifient la qualité nutritionnelle critère par critère.

Les résultats de l'enquête indiquent que les systèmes synthétiques sont sans doute une meilleure aide à la décision dans la mesure où ils permettent sans ambiguïté au consommateur de classer les produits. Les systèmes analytiques sont moins efficaces en cette matière et peuvent induire les consommateurs, notamment ceux ayant le plus faible niveau d'éducation, à des comportements erronés. Entre les deux systèmes synthétiques, Nutri-Score apparaît un peu plus visible que SENS, son usage plus efficace (étant rappelé que la visibilité des étiquettes, optimisée lors des études préparatoires, constitue cependant une caractéristique intrinsèque des différents systèmes).

Les systèmes synthétiques Nutri-Repère et Nutri-Couleurs sont perçus comme plus précis (et de fait ils le sont) et dès lors plus utiles en termes de quantité d'information

délivrée. En contrepartie, ces systèmes analytiques attirent moins l'attention et permettent moins un classement immédiat pertinent des produits.

Un rare point faible de Nutri-Score est que 28% des répondants pensent qu'une étiquette verte indique que l'aliment est issu de l'agriculture biologique (contre 19% pour Nutri-Couleurs et 16% pour SENS, une différence statistiquement significative). A titre de comparaison, seuls 4% des consommateurs pensent qu'un faible score sur Nutri-Repère indique que le produit est bio. Le CS estime qu'il serait utile d'ajouter sur les étiquettes la mention « Un Nutri-Score A ou B (étiquettes vertes) ne garantit pas que le produit est issu de l'agriculture biologique ».

Il est bien étayé aujourd'hui que les consommateurs désirent, de façon complémentaire et non pas alternative, une information détaillée d'une part, un indicateur simple permettant des choix rapides en termes de qualité nutritionnelle d'autre part. L'étude CREDOC montre que les indicateurs synthétiques répondent bien au second volet de ce cahier des charges. Les deux systèmes analytiques testés répondant plus à un objectif d'illustration de la déclaration nutritionnelle obligatoire qui devra bien sûr être maintenue. Cet objectif n'est pas celui dont l'atteinte était l'objet de l'intervention définie dans le protocole scientifique.

Conclusion

Le principal défi que devait relever l'intervention était, dans les limites de temps et sous les contraintes financières assignées à cet exercice, de mettre en évidence des résultats suffisamment fiables pour pouvoir contribuer à éclairer les décisions à prendre par les pouvoirs publics. Le comité scientifique constate qu'à cet égard, l'opération est un succès puisque ces résultats permettent de proposer des réponses objectives aux questions posées.

La première question, implicite mais présente dans tous les esprits, était de savoir si, oui ou non, des systèmes d'étiquetage nutritionnel simplifié étaient susceptibles d'entraîner des modifications dans les comportements d'achat des consommateurs. La réponse est clairement oui. Trois des systèmes testés (Nutri-Couleurs, Nutri-Score et SENS) ont sans ambiguïté un effet positif au regard du critère qu'a retenu le comité scientifique, à savoir le score FSA.

Cet effet, certes statistiquement significatif, n'est toutefois pas de très grande ampleur. Les trois systèmes efficaces (Nutri-Score, Nutri-Couleurs et SENS) améliorent respectivement, en différences premières, le score FSA des achats de produits étiquetés de 0.267, 0.233 et 0.198 points. Ce résultat doit cependant être rapproché du score FSA moyen des achats des consommateurs observés, dont l'ordre de grandeur est de 6 points. En pourcentage, l'amélioration serait donc de l'ordre de 4%. Il faut également rappeler que, pour garantir l'égalité de traitement entre les systèmes, l'intervention n'a été précédée ni accompagnée d'aucune communication sinon l'information, modeste, mise à la disposition des consommateurs sur les lieux de vente.

Par comparaison, le quatrième système testé (Nutri-Repère) n'a pas d'effet. Il semble même avoir, de façon il est vrai statistiquement non significative, un contre effet (+0.025 points de FSA), conduisant certains consommateurs à dégrader la qualité nutritionnelle de leurs achats.

Cet effet positif des trois systèmes efficaces est en outre un effet potentiellement favorable en termes de santé publique. La démonstration est ici indirecte mais le comité scientifique tient à rappeler qu'il avait choisi le score FSA parce qu'il s'agit d'un indicateur validé par des travaux mettant en relation, sur une cohorte importante, son lien aux risques de survenue de cancers, maladies cardiovasculaires, syndrome métabolique et surpoids.

Ce résultat est d'autant plus important que cet effet n'a jusqu'ici jamais été démontré en conditions réelles d'achat en vraie grandeur et en contexte naturel. Le bon aboutissement de l'intervention établit donc, de surcroît, la faisabilité et l'utilité de ce type d'étude.

La deuxième question, celle-ci explicite, était de savoir si ces systèmes pouvaient être classés en fonction de leur efficacité.

Parmi les trois systèmes efficaces, il apparaît que l'analyse suggérée à titre principal par le comité scientifique dans son rapport d'avril 2016 (doubles différences des scores FSA pondérés par les calories, sur les produits étiquetés, et calculées sur l'échantillon total des porteurs de carte) ne permet pas, à elle seule, de mettre en évidence des écarts significatifs entre ces systèmes.

Toutefois, la combinaison d'approches multiples explorées systématiquement (par catégories de produits, d'acheteurs, etc.) fait apparaître une supériorité d'ensemble assez nette pour Nutri-Score qui, à l'occasion de ces diverses mesures, n'entraîne jamais, contrairement aux autres systèmes, de dégradation de la qualité nutritionnelle des achats et produit plus souvent que les autres, de façon statistiquement significative, une amélioration de cette qualité.

Cet avantage de Nutri-Score est plus marqué encore lorsqu'on observe spécifiquement le comportement des consommateurs qui achètent les produits les moins chers.

Ce résultat est renforcé par celui de l'étude qualitative, qui place les systèmes synthétiques, et donc notamment Nutri-Score, clairement devant les systèmes analytiques ;

Au regard de l'ensemble de l'étude, Nutri-Score apparaît donc comme le système le plus efficace.

Annexe 1

Liste des membres du comité scientifique

(au 19 mars 2017)

M. Noël Renaudin : Président du comité scientifique. Président du comité de validation des chartes de progrès nutritionnel (PNNS) et ancien Président du Comité économique des produits de santé.

M. Patrice Bertail, Professeur des universités / Enseignant-Chercheur – Université Paris Ouest Nanterre la défense

M. Pierre Chandon, Professeur de marketing, INSEAD. The L'Oréal Chaired Professor of Marketing – Innovation and creativity

M. Pierre Combris, Directeur de recherche à l'INRA, spécialiste de l'économie de l'alimentation et de l'économie comportementale.

M. Aymery Constant, Maître de conférence des universités, Ecole des hautes études en santé publique, département des sciences humaines et sociales

M. Saadi Lahlou, Professeur de psychologie sociale à la London School of Economics, ancien directeur du département de psychologie sociale de cette université.

Mme Aurélie Lemmens, Associate professor à la Tilburg Scholl of Economics and Management, marketing department

M. Ambroise Martin, Professeur des universités

Mme Natalie Rigal, Psychologue, Enseignant-Chercheur, Université Paris Ouest UFR SPSE 200.

M. Bernard Ruffieux, Professeur à Grenoble INP – Génie industriel, spécialiste de l'économie expérimentale

Annexe 2

Liste des membres de l'équipe de traitement des données

M. Pierre Dubois, directeur scientifique de l'Ecole d'économie de Toulouse

Mme Céline Bonnet, chercheur à l'Ecole d'économie de Toulouse

M. Olivier Allais, directeur du Laboratoire alimentation et sciences sociales de l'INRA

M. Paulo Albuquerque, chercheur à l'INSEAD