



HAL
open science

Reconsidérer la discrétisation des variables quantitatives : vers une nouvelle analyse de modération dans la recherche expérimentale

Romain Cadario, Béatrice Parguel

► To cite this version:

Romain Cadario, Béatrice Parguel. Reconsidérer la discrétisation des variables quantitatives : vers une nouvelle analyse de modération dans la recherche expérimentale. Recherche et Applications en Marketing (French Edition), 2014, 29 (4). halshs-00820942v3

HAL Id: halshs-00820942

<https://shs.hal.science/halshs-00820942v3>

Submitted on 17 Jul 2013 (v3), last revised 20 Jan 2014 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reconsidérer la discrétisation des variables quantitatives : vers une nouvelle analyse de modération dans la recherche expérimentale*

Romain Cadario** & Béatrice Parguel***

Working Paper - Version 2.1 (Juillet 2013)

Résumé

En recherche expérimentale, les chercheurs en marketing discrétisent souvent les variables quantitatives pour tester leur caractère modérateur. Cet article pédagogique explique les limites d'une telle pratique et illustre les étapes des analyses *spotlight* et *floodlight* à conduire suivant que la variable modératrice considérée présente ou non des valeurs focales significatives. **Mots-clés** : Effet modérateur, discrétisation, *median-split*, analyse *spotlight*, analyse *floodlight*, recherche expérimentale.

1 Introduction

Critiquée en 2001 par Irwin et McClelland, la discrétisation des variables quantitatives reste une pratique prégnante au sein de la communauté académique en marketing lorsqu'il s'agit de tester leur effet modérateur, en particulier dans l'analyse de données expérimentales. De fait, les chercheurs sont en priorité formés à la description et à la présentation de designs expérimentaux dans lesquels les variables indépendantes sont qualitatives (Irwin, 2001 ; Fitzsimons, 2008) et il est plus difficile d'expliquer et de représenter un effet d'interaction¹ entre des variables indépendantes de nature quantitative (Irwin et McClelland, 2001). Dès lors, pour tester le caractère modérateur de nombreuses variables sociodémographiques (e.g., âge, revenu), cognitives (e.g., familiarité, implication), affectives (e.g., attachement à la marque), comportementales (e.g., niveau de consommation) ou de personnalité (e.g., besoin de cognition, locus de contrôle)

qu'il convient d'appréhender au moyen de mesures quantitatives, ils recourent le plus souvent à la discrétisation pour ne pas s'écarter du design de recherche le plus courant et qu'ils maîtrisent le mieux.

Le déplorant au quotidien dans les manuscrits soumis à *Journal of Consumer Research*, Fitzsimons appelle en 2008 les chercheurs en comportement du consommateur à en finir avec cette pratique dans un éditorial intitulé « *Death to dichotomizing* ». A la place, il recommande d'utiliser une méthodologie d'étude fondée sur des régressions multiples et conclue par une analyse *spotlight* qui consiste à « placer le projecteur sur le modèle depuis différents angles » (Irwin et McClelland, 2001, p. 106).

La littérature offre des pistes pour comprendre cette méthodologie d'étude des variables modératrices quantitatives. Très technique, l'article séminal d'Irwin et McClelland (2001) pénalise toutefois sa bonne com-

* Les auteurs tiennent à remercier Pierre Valette-Florence pour sa relecture de la première version de cet article au cours de l'été 2012 et la qualité de ses commentaires. Ils adressent également leurs remerciements à Aradhna Krishna et Priya Raghurir pour avoir attiré leur attention sur le problème de la discrétisation des variables quantitatives dans l'analyse de données expérimentales et les avoir encouragés à rédiger cet article.

** Doctorant, Université Paris Dauphine, DRM (UMR CNRS 7088). Contact : romain.cadario@dauphine.fr

*** Chargée de recherche CNRS, Université Paris Dauphine, DRM (UMR CNRS 7088). Contact : beatrice.parguel@dauphine.fr

¹ Si les termes « effet modérateur » et « effet d'interaction » semblent parfois utilisés l'un pour l'autre, ils se distinguent toutefois sur le plan conceptuel. Le terme d'« effet modérateur de la variable M » caractérise l'influence de M sur la relation entre une variable indépendante X et une variable dépendante Y. Il s'emploie dans le cadre d'une discussion d'ordre théorique pour caractériser l'influence d'une variable sur une relation déjà établie. Le terme d'« effet d'interaction entre les variables X et M » caractérise l'influence combinée de deux variables indépendantes sur une variable dépendante Y. Il s'emploie dans le cadre de l'analyse statistique d'un potentiel effet modérateur.

préhension. Plus accessible à un public non spécialiste, l'article de Chumpitaz-Caceres et Vanhamme (2003) ne présente pas l'analyse *spotlight*, bien qu'elle apparaisse comme l'étape fondamentale de cette méthodologie. En 2008, l'éditorial de Fitzsimons (2008) apporte des précisions sur les normes qui régissent l'utilisation de l'analyse *spotlight* au sein de la communauté anglo-saxonne, mais il ne l'aborde que de manière littéraire, sans expliciter clairement les équations à estimer ni recourir à un quelconque exemple numérique. Le codage des données qu'il préconise contient par ailleurs une erreur au regard de celui présenté dans l'article de référence d'Irwin et McClelland (2001)². L'absence de synthèse didactique a finalement contribué au manque de diffusion de cette méthodologie ces dix dernières années. Pour illustration, en France, De Pechpeyrou et Odou sont à notre connaissance les premiers chercheurs en marketing à s'y être intéressés en 2012.

En 2013, Spiller, Fitzsimons, Lynch Jr. et McClelland publient enfin un article dédié à l'étude des variables modératrices quantitatives dans *Journal of Marketing Research*. Ils y détaillent l'analyse *spotlight*, développent l'analyse *floodlight* complémentaire et précisent leurs conditions d'application respectives (i.e., existence ou non de valeurs focales significatives). Leur article reste toutefois complexe sur le plan technique et difficile d'accès pour des chercheurs non avertis en quête d'outils simples à utiliser.

Finalement, le présent article propose aux chercheurs francophones une synthèse didactique pour accompagner l'étude de la modération par des variables quantitatives. Cette contribution de nature pédagogique semble importante à deux égards. D'une part, la mise en évidence d'effets modérateurs constitue l'objet de très nombreux travaux en marketing (Chumpitaz-Caceres et Vanhamme, 2003), notamment dans une recherche expérimentale qui prend régulièrement de l'ampleur au sein de la communauté académique francophone. D'autre part, l'aspiration de nombreux chercheurs à publier dans les meilleures revues du champ suppose de maîtriser les analyses statistiques appropriées, notamment dans les cas de modération par des variables quantitatives (Fitzsimons, 2008).

Pour atteindre son objectif pédagogique, le présent article expose dans une première partie les différents

² Fitzsimons (2008, p.6) indique que la pente de la variable modératrice au niveau "présent" de la variable manipulée peut être examinée lorsque la variable manipulée est codée 0=présence / 1=absence. Or, Irwin et McClelland (2001, p.107) montrent, exemple numérique à l'appui, que dans cette situation la variable manipulée doit être codée de façon inverse, soit 0=absence / 1=présence.

types d'analyses de modération en recherche expérimentale, illustre concrètement les limites de la discrétisation des variables quantitatives et appelle à lui préférer une analyse *spotlight* ou *floodlight*, suivant que la variable modératrice étudiée présente ou non des valeurs focales significatives. Les deuxième et troisième parties exposent successivement et de manière illustrée l'analyse *spotlight* et l'analyse *floodlight*, en détaillant les équations à estimer à chaque étape, ainsi que la manière d'interpréter et de rédiger les résultats. Pour faciliter le propos, toutes les illustrations sont tirées d'un seul et même article de Yang et Raghurir, paru en 2005 dans *Journal of Retailing* et rendu accessible en français en 2006 dans le cadre de la rubrique Sélection internationale de *Recherche et Applications en Marketing*. De plus, une macro originale, simple d'utilisation, est proposée afin d'accompagner l'utilisateur dans la conduite des analyses appropriées sous SPSS.

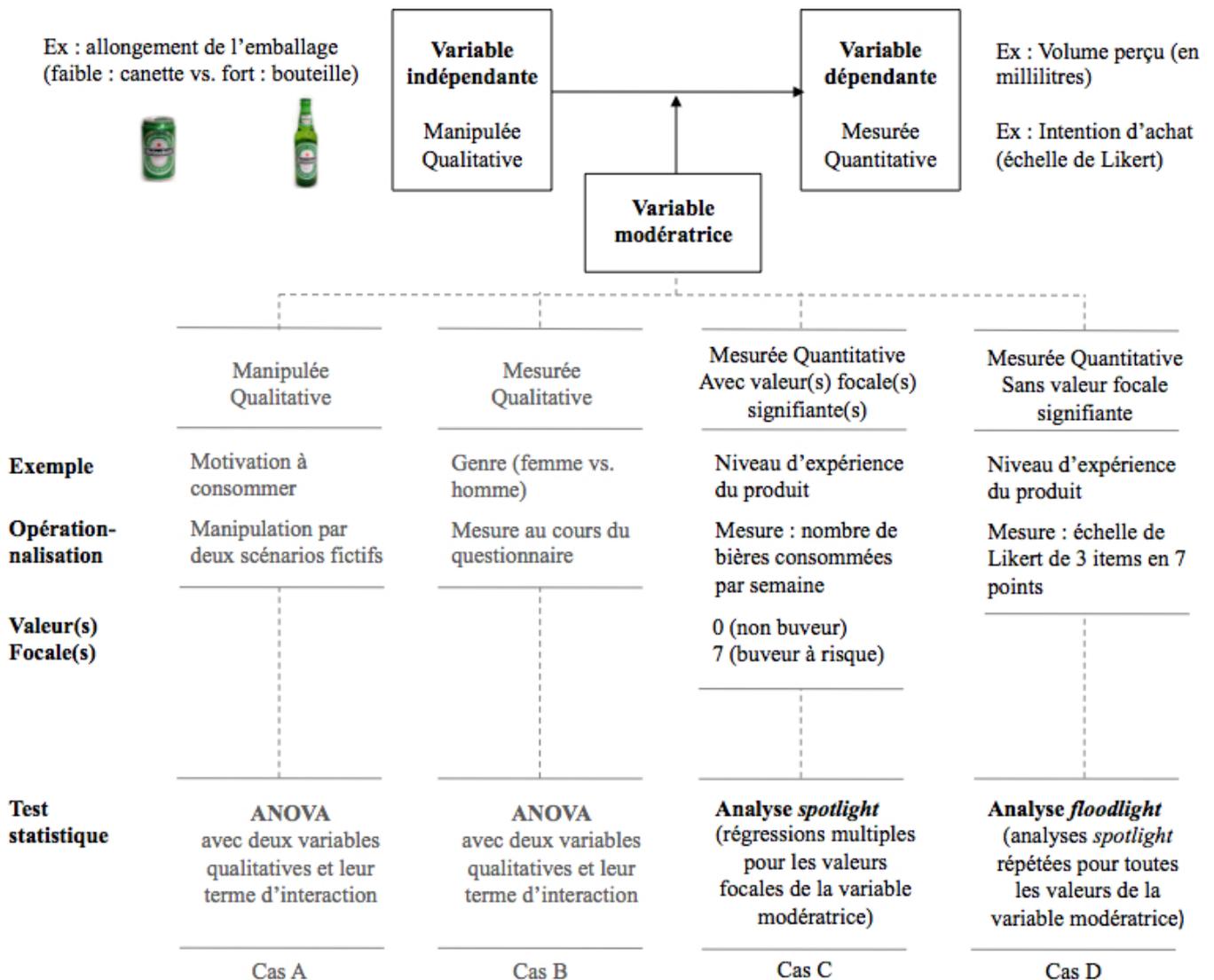
2 Analyses de modération dans la recherche expérimentale

En recherche expérimentale, différentes configurations de modération peuvent se présenter. Cette partie les décrit, les illustre et rappelle les analyses statistiques conduites pour tester l'existence de la modération dans chacune de ces configurations. Elle souligne ensuite les limites de la discrétisation des variables quantitatives et la nécessité du recourt à une méthodologie alternative pour tester leur caractère potentiellement modérateur.

Différentes configurations de modération en recherche expérimentale

Dans la recherche expérimentale, qui vise à contrôler l'influence de chacun des paramètres en n'en faisant varier qu'un seul à la fois, la variable indépendante est par définition manipulée et généralement qualitative. La manipulation d'une variable indépendante de nature quantitative impliquerait un protocole de collecte peu efficient en inter-sujets ou trop biaisé en intra-sujets. La variable dépendante, mesurée, peut être de nature qualitative (e.g., décision d'acheter ou pas, choix entre différentes alternatives) ou quantitative (e.g., perceptions diverses mesurées sur des échelles, montant du panier). Cet article ne traite toutefois que du cas impliquant une variable dépendante quantitative, cas le plus répandu dans l'analyse de données expérimentales en marketing. Le cas impliquant une variable dépendante qualitative fait appel à d'autres techniques d'analyses très spécialisées, relevant notamment des modèles de choix discret

Figure 1 – Différentes configurations de modulation en recherche expérimentale



(Louviere, Hensher et Swait, 2000).

La variable modératrice, qui par définition affecte la direction et/ou l'intensité de la relation entre une variable indépendante et une variable dépendante (Baron et Kenny, 1986), peut, pour sa part, être manipulée – elle est dans ce cas également qualitative – ou mesurée au cours de l'expérimentation, de manière qualitative ou quantitative. La mesure d'une variable modératrice en cours d'expérimentation est une solution efficace pour éviter une manipulation complexe et coûteuse ou susceptible de générer des biais de demande.

Comme présenté en Figure 1, différentes configurations de modulation peuvent être distinguées dans la recherche expérimentale suivant la nature de la variable modératrice (i.e., manipulée qualitative, mesu-

rée qualitative ou mesurée quantitative). L'article de Yang et Raghurir (2005) montre que l'influence de l'allongement de l'emballage d'un produit (en l'occurrence, la bière) sur le volume perçu et la quantité achetée dépend de la motivation contextuelle à consommer le produit, du genre et du niveau d'expérience du produit. Il permet ainsi d'illustrer ces configurations, ainsi que les tests statistiques adaptés pour les analyser.

Le **cas A** considère l'effet modérateur d'une variable manipulée de nature qualitative. Pour l'illustrer, Yang et Raghurir (2005) étudient l'effet modérateur de la motivation contextuelle à consommer le produit (M) sur la relation entre l'allongement de l'emballage (X) et la quantité achetée (Y). Dans ce design de recherche, la variable d'allongement de l'em-

ballage (X), qualitative, est manipulée en inter-sujets au moyen de deux emballages présentant le même volume mais des hauteurs différentes (faible allongement : canette de bière / fort allongement : bouteille de bière). La variable de motivation contextuelle à consommer le produit (M), également qualitative, est manipulée selon deux scénarios (faible motivation : soirée normale la veille des derniers partiels et qui doit finir à 23h / forte motivation : soirée extravagante le lendemain des derniers partiels et qui doit finir à 2h du matin). La variable de quantité achetée (Y), quantitative, est mesurée en nombre d'unités achetées. Dans ce cas, pour tester l'effet modérateur, il faut conduire une ANOVA factorielle, avec la quantité achetée en variable dépendante et les deux variables qualitatives (i.e., allongement de l'emballage et motivation à consommer le produit) ainsi que leur terme d'interaction en variables indépendantes.

Le **cas B** considère l'effet modérateur d'une variable mesurée de nature qualitative. Pour l'illustrer, Yang et Raghurir (2005) étudient l'effet modérateur du genre (M) sur la relation entre l'allongement de l'emballage (X) et le volume perçu (Y). Dans ce design de recherche, la variable de genre (M), qualitative, est mesurée au cours de la collecte de données associée à l'expérimentation (i.e., renseignement explicite par le répondant ou implicite par le chercheur à l'insu du répondant). La variable de volume perçu (Y), quantitative, est mesurée en millilitres (mL). A l'instar du cas précédent, pour tester l'effet modérateur, il faut conduire une ANOVA factorielle, avec le volume perçu en variable dépendante et les deux variables qualitatives (i.e., allongement de l'emballage et genre) ainsi que leur terme d'interaction en variables indépendantes.

Dans le cas d'une variable modératrice quantitative (i.e., **cas C et D**), les auteurs en recherche expérimentale, familiers du cadre de l'analyse de variance, peuvent être incités à discrétiser la variable modératrice pour l'introduire ensuite dans une ANOVA. Pour l'illustrer, Yang et Raghurir (2005) étudient l'effet modérateur du niveau d'expérience du produit (M) sur la relation entre l'allongement de l'emballage (X) et le volume perçu (Y). Dans ce design de recherche, la variable de niveau d'expérience du produit (M), quantitative, est mesurée au cours du questionnaire par une invitation des répondants à préciser le nombre de bières qu'ils consomment habituellement en une semaine. Cette variable mesurée est ensuite discrétisée, c'est-à-dire transformée en une variable catégorielle, et recodée en trois modalités : « non consommateur » si la consommation hebdomadaire est nulle, « petit consommateur » ou « gros consommateur » suivant que l'individu présente une consommation hebdomadaire de bière inférieure ou supérieure à la médiane.

Les auteurs réalisent ensuite une ANOVA factorielle avec le volume perçu en variable dépendante et l'allongement de l'emballage, le niveau d'expérience du produit discrétisé et leur terme d'interaction en variables indépendantes.

Dans le cas du test du caractère modérateur d'une variable quantitative, la discrétisation, telle que pratiquée ci-dessus, présente de nombreuses limites et doit être remplacée par une analyse *spotlight* ou une analyse *floodlight* suivant que la variable modératrice étudiée présente ou non des valeurs focales significatives. Les deux prochaines sections abordent successivement ces deux points.

Limites de la discrétisation des variables modératrices quantitatives

Si la pratique de la discrétisation des variables quantitatives dont on souhaite tester le caractère modérateur séduit par sa simplicité, elle pose une première question. Si le chercheur choisit de retenir un outil quantitatif pour mesurer une variable, c'est notamment parce qu'au terme de précédents débats, la littérature s'est accordée à le reconnaître comme l'une des manières les plus fidèles de mesurer cette variable au regard de sa conceptualisation. Discrétiser cette variable revient ainsi à ignorer les discussions théoriques et travaux méthodologiques qui ont contribué à établir sa légitimité au sein de la communauté. Au-delà de cette question presque « philosophique » (Irwin, 2001), la pratique de la discrétisation des variables quantitatives rencontre également trois limites sur le plan statistique.

Une première limite réside dans l'obtention de résultats significatifs fallacieux (*spurious* en anglais) : la discrétisation des variables quantitatives peut dans certains cas générer à tort des effets significatifs, soit des erreurs de Type I (Maxwell et Delaney, 1993 ; Vargha et alii, 1996). Ainsi, l'effet d'une variable modératrice discrétisée (M) peut apparaître à tort significatif si M est corrélée avec une variable indépendante (X). Cet effet ne reflète pas l'influence de M , mais l'influence réelle de X sur la variable dépendante (Y). Concrètement, parce que M est corrélée à X , M est aussi partiellement corrélée aux résidus de la régression de Y par X et contribue donc à prédire Y lorsque X est contrôlée. Lorsque X et M sont corrélées, un effet d'interaction peut également apparaître à tort entre X et M dans le cas d'une influence non linéaire de l'une des deux variables sur Y . Le lecteur intéressé pourra trouver une démonstration mathématique de ce résultat chez Maxwell et Delaney (1993, p. 187). L'obtention de résultats fallacieux est d'autant plus probable et grave que les variables indépendantes sont corrélées.

La deuxième limite concerne le mode de discrétisation de la variable et le codage retenus. Comme Tybout (2001) le suggère, la discrétisation d'une variable quantitative est toujours arbitraire. Ce choix constitue une forme de « degré de liberté du chercheur » (Simmons, Neslon et Simonsohn, 2011). Ainsi, les chercheurs peuvent discrétiser leurs variables à la médiane ou à la moyenne, voire considérer des terciles ou des quartiles, sans que cela ne repose sur la moindre justification théorique. Or, ce choix n'est pas anodin, car il peut influencer le résultat des tests de modération conduits ultérieurement. Pour illustration, la discrétisation d'une variable par affectation des observations égales à la médiane au groupe supérieur ou, alternativement, au groupe inférieur peut influencer la significativité du test de son caractère modérateur (voir Annexe A1 pour une illustration numérique inspirée d'Irwin et McClelland, 2001, p. 106). Nul doute dans ce cas que la probabilité de mettre en évidence des résultats significatifs doit influencer le mode de discrétisation retenu par les chercheurs.

La troisième limite tient au fait que la discrétisation d'une variable modératrice quantitative réduit fortement le pouvoir statistique associé au test (Irwin et McClelland, 2001, 2003). En effet, discrétiser une variable quantitative conduit à une baisse du coefficient de détermination R^2 ainsi qu'à une baisse des statistiques de test pour la variable discrétisée. Dans certains cas, la réduction du pouvoir statistique liée à la discrétisation peut être telle que l'effet d'interaction perd sa significativité au seuil de risque retenu et qu'il apparaisse ainsi une erreur de Type II (voir Annexe A1 pour une illustration numérique). Se traduisant par une perte d'informations équivalente à la moitié des observations utiles (McClelland, 1997), la discrétisation suppose de collecter deux fois plus d'observations pour garder le même pouvoir statistique, voire beaucoup plus d'observations si elle s'effectue suivant les terciles ou quartiles extrêmes (Irwin et McClelland, 2003). Toutes choses égales par ailleurs, elle peut donc aboutir à des résultats « conservateurs » (Irwin et McClelland, 2003) et réduire la capacité des chercheurs à détecter de potentiels effets d'interaction.

En synthèse, la discrétisation des variables quantitatives affecte la capacité des chercheurs à analyser la véritable nature des relations qui existent entre leurs variables. Cette pratique étant fortement déconseillée, les chercheurs sont invités à tester le caractère modérateur de leurs variables quantitatives en utilisant une analyse *spotlight* ou une analyse *floodlight*, suivant que ces variables présentent ou non des valeurs focales significatives.

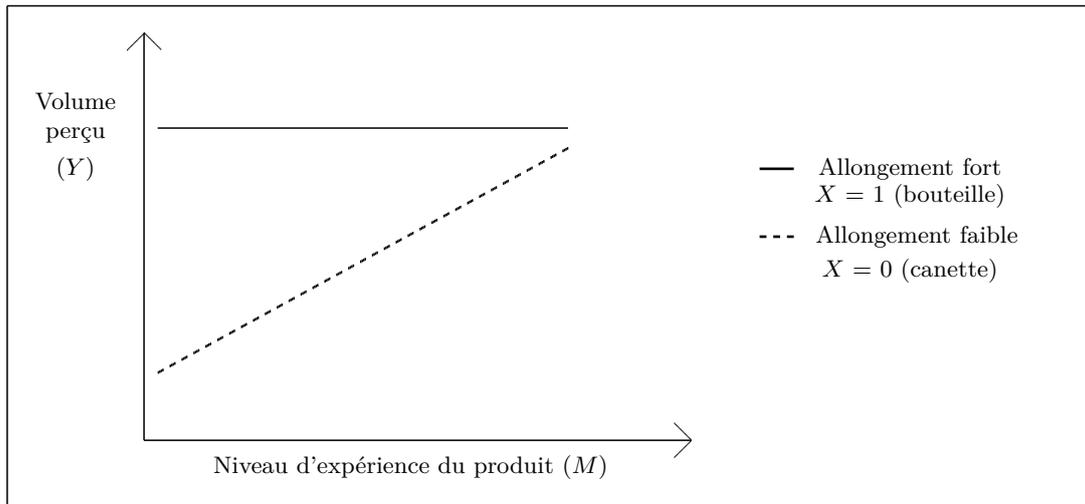
Choix de l'analyse *spotlight* ou de l'analyse *floodlight*

Les analyses *spotlight* et *floodlight* permettent d'éclairer le sens d'un effet d'interaction une fois que celui-ci a été mis en évidence. Elles indiquent concrètement le sens de la relation entre la variable indépendante (X) et la variable dépendante (Y) pour différentes valeurs de la variable modératrice quantitative (M). Comme indiqué en Figure 1, le choix entre ces deux analyses dépend de M : si M présente une ou plusieurs valeur(s) focale(s) significative(s), le chercheur choisira l'analyse *spotlight* et conduira l'analyse pour ces valeurs judicieusement choisies (et non de manière arbitraire à plus ou moins un écart type de la moyenne de M comme y invitait Fitzsimons en 2008). Dans le cas contraire, le chercheur choisira l'analyse *floodlight* (Spiller et alii, 2013).

Une valeur est dite significative si elle est mesurée sur une échelle non arbitraire, c'est-à-dire une échelle qui reflète statistiquement le construit étudié. Elle est dite focale si elle « fait sens » sur le plan théorique, notamment parce qu'elle correspond à un niveau seuil ou plancher reconnu par la littérature. L'article de Yang et Raghurir (2005) permet d'illustrer concrètement ces deux caractéristiques. Il étudie l'effet modérateur du niveau d'expérience du produit (M) sur la relation entre l'allongement de l'emballage (X) et le volume perçu (Y). La variable d'allongement de l'emballage, manipulée en inter-sujets, distingue deux formes d'emballages de même volume : une canette présentant un faible allongement (codée $X = 0$) et une bouteille présentant un fort allongement (codée $X = 1$). La variable de volume perçu est mesurée en mL. Pour précision, l'hypothèse d'un effet modérateur repose sur l'idée que seuls les individus présentant un faible niveau d'expérience du produit utilisent l'heuristique de l'allongement de l'emballage pour évaluer le volume du produit. Cet effet de modération supposé est représenté en Figure 2.

Si le niveau d'expérience du produit est mesuré par le nombre de bières consommées habituellement en une semaine, il est mesuré sur une échelle de type ratio significative et non arbitraire. En effet, cette échelle affiche un « zéro » absolu qui reflète concrètement une absence de consommation de bière. De plus, elle se mesure en unités appréhendables, le passage d'une unité à la suivante représentant concrètement une bière supplémentaire. Pour cette mesure du niveau d'expérience du produit, il est par ailleurs possible d'imaginer au moins deux valeurs focales : « zéro », correspondant à une consommation nulle, et « sept » correspondant à un niveau significatif, si l'on considère qu'un seul verre d'alcool par jour suffit à augmenter les risques de cancers (Nelson et alii, 2013). A contrario, si le niveau d'expérience du produit est mesuré sur l'échelle Likert de Soderlund (2002), en 7

Figure 2 – Effet modérateur de l'exemple (Yang et Raghurir, 2005)



points allant de 0 (« pas du tout d'accord ») à 6 (« tout-à-fait d'accord ») et composée de 3 items (e.g., « je connais de nombreuses bières différentes »), il est mesuré sur une échelle non significative. L'échelle de Soderlund (2002) est en effet une échelle arbitraire : la modalité « pas du tout d'accord » ne distingue pas une expérience réduite d'une absence totale d'expérience et l'unité de l'échelle ne reflète pas l'existence de modalités distinctes sur le plan théorique pour le construire. Cette échelle ne présente par ailleurs pas de valeur focale susceptible de déclencher des changements dans le comportement, les perceptions ou le risque encouru par les consommateurs. Suivant que le niveau d'expérience du produit est mesuré en nombre de bières habituellement consommées par semaine ou sur l'échelle de Soderlund (2002), les deux prochaines parties exposent les différentes étapes des tests à conduire pour réaliser l'analyse *spotlight* et l'analyse *floodlight*, ainsi que la manière d'interpréter et de rédiger les résultats obtenus au terme de l'estimation.

3 Mise en œuvre de l'analyse *spotlight*

Les trois étapes de l'analyse de l'effet modérateur du niveau d'expérience du produit par régressions multiples sont présentées dans les trois prochaines sections. Une base de données fictive téléchargeable sur le site Internet du premier auteur <https://sites.google.com/site/romaincadario/> permet d'illustrer les tests statistiques relatifs à ces trois étapes et de faciliter ainsi leur compréhension. Illustrant l'article de Yang et Raghurir (2005), cette

base de données fictive renseigne notamment le volume perçu (variable « Volume » mesurée en mL) et le niveau d'expérience du produit (variable « Expérience_NbBières » mesurée en nombre de bières habituellement consommées par semaine) pour 80 individus (i.e., 40 observations pour la condition « canette » et 40 observations pour la condition « bouteille »). Le lecteur intéressé est invité à s'en emparer pour suivre la mise en œuvre des différentes étapes de l'analyse et vérifier si les résultats qu'il obtient sont conformes à ceux présentés dans cet article.

Etape 1 : Etablir l'existence de l'effet d'interaction

Etablir l'existence de l'effet d'interaction passe par la validation de l'hypothèse théorique HA³ :

HA : Le niveau d'expérience du produit modère la relation entre l'allongement de l'emballage et le volume perçu.

Equation à estimer

Le test de l'hypothèse HA passe par la régression linéaire de la variable dépendante de volume perçu (Y) sur les variables indépendantes de niveau d'expérience du produit (M), d'allongement de l'emballage

² Il est important de noter que cette hypothèse, qui ne précise pas la nature de l'effet d'interaction et n'est donc pas testable formellement, est à proscrire d'un article de recherche. Elle n'est précisée ici que pour des raisons pédagogiques, dans une optique de comparaison avec les hypothèses de contraste du modèle.

Tableau 1 – Résultats des régressions pour l'analyse *spotlight*

Variables (paramètre)	Equation 1 $M' = M$ $X' = X$	Equation 2 $M' = M$ $X' = 1 - X$	Equation 3.1 $M' = M - 7$ $X' = X$
Constante (<i>a</i>)	26,02* (41,36)	29,80* (47,39)	29,93* (63,69)
Expérience M' (<i>b</i>)	0,56* (4,62)	,15 (1,22)	0,56* (4,62)
Allongement X' (<i>c</i>)	3,78* (4,26)	-3,78* (-4,26)	0,91 (1,37)
Interaction $M'X'$ (<i>d</i>)	-0,41* (-2,40)	0,41* (2,40)	-0,41* (-2,40)

* $p < 0,05$; t-value entre parenthèses, variable dépendante : volume perçu (Y).

(X), ainsi que sur le terme d'interaction entre ces deux variables indépendantes, soit l'estimation de l'équation 1 :

$$(1) Y = a_1 + b_1M + c_1X + d_1MX$$

Résultats

Les résultats de la régression linéaire appliquée à la base de données fictive sont présentés dans le Tableau 1. Ces résultats montrent l'existence d'un effet d'interaction significatif ($d_1 = -0,41$; $t = -2,40$; $p < 0,05$) et invitent à explorer celui-ci plus avant (voir Tableau 2 pour une synthèse sur l'interprétation des coefficients).

Etape 2 : Explorer la nature de l'effet d'interaction à partir de l'analyse des pentes

Explorer le sens de la modération suivant l'allongement de l'emballage passe par le test des hypothèses théoriques HB1 et HB2 :

HB1 : Dans le cas d'un faible allongement de l'emballage (canette), le volume perçu est plus important chez les individus présentant un fort vs. un faible niveau d'expérience du produit.

HB2 : Dans le cas d'un fort allongement de l'emballage (bouteille), le volume perçu n'est pas différent chez les individus présentant un fort vs. un faible niveau d'expérience du produit.

Equation à estimer

Le test statistique de HB1 passe par l'étude de la pente de la courbe de volume perçu (Y) en fonc-

tion du niveau d'expérience du produit (M) lorsque l'allongement de l'emballage (X) est faible (i.e., la canette). Cette pente correspond au paramètre estimé de la variable M à l'issue de la régression linéaire de l'équation 1 dans laquelle la variable d'allongement de l'emballage est codée $X = 0$ pour un allongement faible (i.e., canette).

Or, lorsque $X = 0$ (i.e., canette), l'équation 1 devient $Y = a_1 + b_1M$. L'analyse de la significativité et du signe du paramètre estimé b_1 dans l'équation 1 permet donc de tester l'hypothèse HB1 (voir Tableau 2). Par exemple, un paramètre significatif et positif indique que, lorsque l'allongement est faible (i.e., la canette), le volume perçu augmente lorsque le niveau d'expérience du produit augmente.

De même, le test statistique de HB2 passe par l'étude de la significativité et du signe de la pente de la courbe de volume perçu (Y) en fonction du niveau d'expérience du produit (M) lorsque l'allongement de l'emballage (X) est fort (i.e., la bouteille). Cette pente correspond au paramètre b_2 estimé à l'issue de la régression linéaire de l'équation 2 dans laquelle la variable d'allongement de l'emballage est codée 0 pour un allongement fort. Pour ce faire, le codage de la variable binaire d'allongement de l'emballage (X) est inversé pour la transformer en une nouvelle variable X' , telle que $X' = 1 - X$, avant l'estimation de l'équation 2 :

$$(2) \begin{cases} Y = a_2 + b_2M + c_2X' + d_2MX' \\ X' = 1 - X \end{cases}$$

L'analyse de la significativité et du signe du paramètre estimé b_2 permet de valider, ou non, l'hypothèse HB2 (voir Tableau 2). Par exemple, une pente non significative indique que, lorsque l'allongement de l'emballage est fort (bouteille), le volume perçu ne dépend pas du niveau d'expérience du produit.

Tableau 2 – Synthèse sur l’interprétation des coefficients

	Interprétation du coefficient	Résultats
d_1	Variation de Y en fonction de X suite à l’augmentation d’une unité de M	- d_1 non significatif : absence d’effet d’interaction - d_1 significatif : l’effet d’interaction est établi et doit être décomposé au travers des coefficients b_1 , b_2 et c_3 .
b_1	Pente de la courbe de M en fonction de Y , lorsque $X = 0$ (canette)	- b_1 non significatif : la pente est plate, il n’existe pas de lien entre M et Y , lorsque $X = 0$ - b_1 significatif et positif : la pente est croissante, la variable Y augmente quand M augmente, lorsque $X = 0$ (canette) - b_1 significatif et négatif : la pente est décroissante, la variable Y baisse quand M augmente, lorsque $X = 0$ (canette)
b_2	Pente de la courbe de M en fonction de Y , lorsque $X = 1$ (bouteille)	- b_2 non significatif : la pente est plate, il n’existe pas de lien entre M et Y , lorsque $X = 0$ - b_2 significatif et positif : la pente est croissante, la variable Y augmente quand M augmente, lorsque $X = 1$ (bouteille) - b_2 significatif et négatif : la pente est décroissante, la variable Y baisse quand M augmente, lorsque $X = 1$ (bouteille)
c_3	Distance entre les deux courbes de régression pour $X = 1$ et $X = 0$ au niveau de la zone focale de la variable modératrice telle que $M = M_{focale}$	- c_3 non significatif : Y ne varie pas significativement entre les conditions $X = 1$ et $X = 0$, lorsque $M = M_{focale}$ - c_3 significatif et positif : Y est plus élevé lorsque $X = 1$ que lorsque $X = 0$, quand $M = M_{focale}$ - c_3 significatif et négatif : Y est plus élevé lorsque $X = 0$ que lorsque $X = 1$, quand $M = M_{focale}$

Résultats

Comme les résultats de l’analyse appliquée à la base de données fictive le montrent dans le Tableau 1, la régression linéaire correspondant à l’équation 1 indique un paramètre b_1 significatif et positif ($b_1 = 0,56$; $t = 4,62$; $p < 0,05$). Elle met en évidence que, lorsque l’allongement de l’emballage est faible (i.e., la canette), la pente de la courbe de volume perçu en fonction du niveau d’expérience du produit est positive (i.e., le volume perçu augmente lorsque le niveau d’expérience du produit augmente). L’hypothèse HB1 est donc corroborée.

Par ailleurs, la régression linéaire correspondant à l’équation 2 indique un paramètre b_2 non significatif ($b_2 = 0,15$; $t = 1,22$; $p = 0,22$). Elle met en évidence que, lorsque l’allongement de l’emballage est fort (i.e., la bouteille), la pente de la courbe de volume perçu ne dépend pas du niveau d’expérience du produit. L’hypothèse HB2 est donc corroborée.

Etape 3 : Explorer l’effet d’interaction avec l’analyse de projecteur

Explorer le sens de la modération suivant le niveau d’expérience du produit passe par le test des hypothèses théoriques HC1 et HC2 :

HC1 : Chez les individus présentant un fort niveau d’expérience du produit (i.e., une bière par jour en moyenne), le volume perçu n’est pas différent dans le cas d’un fort allongement de l’emballage (bouteille) vs. faible (canette).

HC2 : Chez les individus présentant un faible niveau d’expérience du produit (i.e., non consommateur), le volume perçu est relativement plus important dans le cas d’un fort allongement de l’emballage (bouteille) vs. faible (canette).

Equation à estimer

Le test statistique de HC1 passe par l’étude de la distance entre les courbes de volume perçu (Y) en fonction de la variable modératrice quantitative (M) selon chaque niveau d’emballage (X) pour chaque valeur focale significative de M (i.e., la distance entre la courbe pleine et la courbe en pointillés sur la Figure 2). Cette distance correspond au paramètre c_3 estimé à l’issue de la régression linéaire dans laquelle la variable modératrice M est transformée en une nouvelle variable M' , telle que $M' = M - M_{focale}$, avant l’estimation de l’équation 3 :

$$(3) \begin{cases} Y = a_3 + b_3M' + c_3X + d_3M'X \\ M' = M - M_{focale} \end{cases}$$

En effet, lorsque l'allongement est fort ($X = 1$), l'équation (3) devient $Y = (a_3 + c_3) + (b_3 + d_3)M'$ et lorsque l'allongement est faible ($X = 0$), (3) devient $Y = a_3 + b_3M'$. En prenant la différence de ces deux équations transformées, on obtient la variation de Y en fonction de la variation de l'emballage : $\Delta Y = c_3 + d_3M'$. De plus, au niveau de la zone focale $M = M_{focale}$ sur laquelle on place le projecteur (*spotlight*), la variable M' devient nulle ($M' = M - M_{focale} = 0$) et on obtient donc finalement $\Delta Y = c_3$. Le coefficient c_3 représente donc la distance entre les deux courbes de régression pour $X = 1$ (i.e., bouteille) et $X = 0$ (i.e., canette), lorsque la variable modératrice prend sa valeur focale ($M = M_{focale}$).

Dans notre exemple, le test statistique de l'hypothèse HC1 passe par l'étude du coefficient c_3 lorsque $M = 7$. Pour ce faire, la variable M est transformée en une variable M' telle que $M' = M - 7$ dans l'équation 3.1.

$$(3.1) \begin{cases} Y = a_3 + b_3M' + c_3X + d_3M'X \\ M' = M - 7 \end{cases}$$

L'analyse de la significativité et du signe du paramètre estimé c_3 permet donc de tester l'hypothèse HC1. Ainsi, une valeur non significative du coefficient estimé signifie que, chez les individus présentant un niveau d'expérience élevé du produit, il n'y a pas de différence de volume perçu dans le cas d'un allongement faible (i.e., la canette) vs. un allongement fort (i.e., la bouteille) de l'emballage.

De même, le test statistique de l'hypothèse HC2 passe par l'étude du coefficient c_3 lorsque $M = 0$. Pour ce faire, la variable M est transformée en une variable M' telle que $M' = M - 0$ dans l'équation 3.2 (nous remarquons immédiatement qu'aucune transformation n'a été effectuée par rapport à l'équation 1 car $M' = M$ au niveau la valeur focale $M = 0$).

$$(3.2) \Leftrightarrow (1) \begin{cases} Y = a_3 + b_3M' + c_3X + d_3M'X \\ M' = M - 0 \end{cases}$$

A nouveau, l'analyse de la significativité et du signe du paramètre estimé c_3 permet donc de tester l'hypothèse HC2 (voir Tableau 2). Ainsi, une valeur significative et positive du coefficient estimé c_3 signifie que, chez les individus présentant un faible niveau d'expérience du produit, le volume perçu est moins important dans le cas d'un allongement faible (i.e., la canette) vs. un allongement fort (i.e., la bouteille) de l'emballage.

Résultats

Comme le montre le Tableau 1, la régression linéaire correspondant à l'équation 3.1 indique un paramètre c_3 non significatif ($c_3 = 0,91; t = 1,37; p = 0,17$). Elle met en évidence que chez les individus présentant un niveau d'expérience élevé du produit (i.e., consommation moyenne d'une bière par jour), il n'y a pas de différence de volume perçu dans le cas d'un allongement faible (i.e., canette) vs. un allongement fort (i.e., bouteille) de l'emballage. L'hypothèse HC1 est donc corroborée.

De plus, la régression linéaire correspondant à l'équation 1 (équivalente à l'équation 3.2) indique un paramètre c_1 positif et significatif ($c_1 = 3,78; t = 4,26; p < 0,05$). Elle met en évidence que chez les individus présentant un faible niveau d'expérience du produit, le volume perçu est relativement plus élevé (3,78 mL de plus) dans le cas d'un allongement fort (i.e., bouteille) vs. un allongement faible (i.e., canette) de l'emballage. L'hypothèse HC2 est donc corroborée.

Utilisation de la macro SPSS « Analyse de modération avec *spotlight* »

Les lecteurs intéressés peuvent télécharger librement la macro « Analyse de modération avec projecteur » développée pour SPSS sur le site Internet du premier auteur : <https://sites.google.com/site/romaincadario/>. Cette macro permet de conduire l'analyse *spotlight* pour chaque valeur focale significative sans avoir à transformer la variable modératrice.

Pour installer cette macro dans SPSS, il suffit d'aller dans 'Utilitaires' > 'Boîte de dialogue personnalisée' > 'Installer une boîte de dialogue personnalisée', avant de sélectionner le fichier « *Spotlight.spd* » et de cliquer sur 'OK'. La boîte de dialogue « Analyse de modération avec *spotlight* » (voir Annexe A2) sera alors accessible en allant dans 'Analyse' > 'Régression' > 'Analyse de modération avec *spotlight*'.

Une fois que la macro est installée, il ne reste plus qu'à faire glisser les variables dans la boîte de dialogue. Dans le cas de notre exemple, il faut donc faire glisser la variable « Allongement » dans la case Variable indépendante X et préciser le codage initial des données, c'est-à-dire qu'un faible allongement de l'emballage (i.e., la canette) est codé $X = 0$ et qu'un fort allongement de l'emballage (i.e., la bouteille) est codé $X = 1$. Il faut ensuite faire glisser la variable « Expérience_NbBières » dans la case Variable modératrice M et préciser la valeur focale significative étudiée. Il faut enfin faire glisser la variable « Volume » dans la case Variable dépendante Y et cliquer sur 'OK'. Il est à noter que la macro permet d'inclure des variables de contrôle pour un meilleur ajustement du modèle.

L'Annexe A3 présente les résultats de l'analyse *spotlight* tels que renvoyés par la macro SPSS pour la base de données fictive. Ces résultats s'articulent autour de l'estimation successive des trois équations présentées plus haut. Un texte d'accompagnement est placé sous les trois tableaux de résultats afin d'accompagner l'utilisateur dans leur interprétation.

Présentation des résultats : rédaction et transparence

Si l'analyse de modération par régressions multiples est à la portée de tout chercheur sur le plan des tests à effectuer, le défi réside véritablement dans la capacité à décrire les résultats obtenus de manière claire. En l'occurrence, il n'est pas nécessaire de faire apparaître les équations et le tableau de résultats des régressions qui ne sont inclus dans cet article que pour des raisons pédagogiques. Les résultats de la recherche peuvent être rédigés comme suit (Fitzsimons, 2008) :

Nous avons effectué une régression du volume perçu sur trois variables indépendantes : (i) le niveau d'expérience du produit ($m = 4,35; \sigma = 2,85; min = 0; max = 10$), (ii) l'allongement de l'emballage (allongement fort : bouteille, codé 1 vs. allongement faible : canette, codé 0), et (iii) le terme d'interaction. Les résultats montrent un effet d'interaction significatif entre le niveau d'expérience du produit et l'allongement de l'emballage ($\beta = -0,41; t = -2,40; p < 0,05$). Pour décomposer cet effet d'interaction, nous avons examiné la pente de niveau d'expérience du produit à chaque niveau d'allongement de l'emballage. Cette pente est significative et positive lorsque l'allongement de l'emballage est faible ($\beta = 0,56; t = 4,62; p < 0,05$), mais non significative lorsqu'il est fort ($p = 0,22$). Une analyse *spotlight* au niveau de la valeur focale d'expérience $M = 7$ montre par ailleurs une différence non significative ($p = 0,17$), de sorte que les individus présentant un niveau d'expérience élevé du produit ne perçoivent pas de différence de volume entre un allongement faible et un allongement fort de l'emballage. De façon similaire, une analyse *spotlight* au niveau de la valeur focale d'expérience $M = 0$ montre une différence significative ($\beta = 3,78; t = 4,26; p < 0,05$), de sorte que les individus présentant un faible niveau d'expérience du produit

perçoivent un volume plus important dans le cas d'un allongement fort vs. faible de l'emballage.

Dans le paragraphe qui précède, nous avons exposé la manière appropriée de présenter les résultats issus des trois étapes de l'analyse de modération par régressions multiples. Ces trois étapes correspondent à une analyse complète de l'effet d'interaction. D'un point de vue pratique, si les hypothèses théoriques formulées par le chercheur correspondent davantage aux hypothèses de contraste mises en avant à la troisième étape (i.e., hypothèses du type HC1 et HC2), la présentation des résultats pourrait se limiter aux étapes 1 (i.e., établir l'existence de l'effet d'interaction) et 3 (i.e., explorer cet effet d'interaction avec l'analyse *spotlight*). Toutefois, à la suite de Simmons et de ses collègues (2011), nous recommandons aux chercheurs une entière transparence dans la présentation de leurs résultats. Pour eux, les chercheurs ne devraient pas avoir la liberté de « sélectionner les comparaisons entre conditions expérimentales qui donnent des résultats conformes à leurs hypothèses » (p. 1363). En conséquence, nous recommandons aux chercheurs d'effectuer et de présenter l'intégralité des résultats des tests associés à l'analyse de modération par régressions multiples.

4 Mise en œuvre de l'analyse *floodlight*

Lorsque la variable modératrice quantitative M ne présente pas de valeur significative ou pas de valeur focale (i.e., cas D), une analyse *floodlight* doit être réalisée pour étudier l'influence de la variable indépendante X sur la variable dépendante Y pour toutes les valeurs possibles de la variable modératrice quantitative M . Pour illustrer cette analyse, nous reprenons l'exemple de Yang et Raghuram (2005) et l'adaptions en imaginant que le niveau d'expérience du produit soit mesuré sur l'échelle en trois items de Soderlund (2002). Dans la base de données fictive, la variable « Expérience_Likert3Items » est calculée comme la moyenne de ces trois items. Comme dans la partie précédente, l'analyse de l'effet modérateur du niveau d'expérience du produit par régressions multiples passe par trois étapes. Les étapes 1 (i.e., établir l'effet d'interaction) et 2 (i.e., explorer la nature de l'effet d'interaction à partir de l'analyse des pentes) sont identiques; nous ne revenons donc pas dessus et présentons directement la troisième étape, celle de l'analyse *floodlight*.

Tableau 3 – Résultats des régressions pour l’analyse *floodlight*

Expérience	Effet*	Erreur standard	<i>t-value</i>	<i>p-value</i>
0	3,52	0,90	3,94	<0,0001
1	2,97	0,68	4,35	<0,0001
2	2,40	0,53	4,52	<0,0001
3	1,84	0,51	3,94	<0,001
4	1,27	0,62	2,06	0,04
5	0,71	0,82	0,87	0,39
6	0,15	1,05	0,14	0,89

*Coefficient c_3 de la régression 3, représentant la différence de Y (volume perçu) entre les deux conditions de X (bouteille vs. canette) pour différentes valeurs de l’expérience. Les sept estimations ont été obtenues avec la macro « Analyse de modération avec *spotlight* »

Explorer l’effet d’interaction avec l’analyse *floodlight*

L’analyse *floodlight* consiste en une répétition d’analyses *spotlight*. Dans l’analyse *spotlight*, on braque le projecteur sur une ou plusieurs valeur(s) focale(s) signifiante(s). Dans l’analyse *floodlight*, on braque le projecteur sur l’ensemble des valeurs prises par la variable modératrice quantitative (M) et on identifie les zones pour lesquelles la variable indépendante manipulée (X) a un effet significatif sur la variable dépendante (Y). La frontière entre ces zones est donnée par le point Johnson-Neyman (Johnson et Neyman, 1936).

Toute valeur de la variable modératrice quantitative M pour laquelle une analyse *spotlight* révèle une *p-value* exactement égale à 0,05 pour le coefficient c_3 de l’équation 3 (représentant la variation Y entre les deux conditions de X pour cette valeur) est qualifiée de point Johnson-Neyman. Ce point distingue donc deux zones : celles des valeurs de M pour lesquelles il existe un effet significatif de X sur Y et celles des valeurs de M pour lesquelles cet effet n’existe pas. Le signe du coefficient c_3 permet d’interpréter le sens de la relation entre X et Y (voir Tableau 2). Une partie des résultats de l’analyse *floodlight* appliquée à la base de données fictive sont présentés dans le Tableau 3. Ils correspondent à l’estimation de l’équation 3 (en utilisant la macro « Analyse de modération avec *spotlight* ») pour sept valeurs entières de la variable de niveau d’expérience du produit (i.e., de 0 à 6).

Dans le Tableau 3, l’effet correspond au coefficient c_3 de l’équation 3 et reflète la variation de volume perçu entre les deux conditions d’allongement pour chaque niveau d’expérience du produit. Celui-ci est significatif lorsque le niveau d’expérience du produit est inférieur ou égal à 4 et non significatif au-delà, indiquant que le point Johnson-Neyman est compris entre 4 et 5. La macro PROCESS (voir plus loin) permet de réaliser des itérations pour l’ensemble des valeurs possibles de la variable de niveau d’expérience du produit

et de localiser précisément le point Johnson-Neyman à 4,05 (voir Annexe 5). Ces résultats montrent que seuls les individus présentant un faible niveau d’expérience du produit évaluent de manière biaisée le volume perçu. Ils estiment en effet que le volume de l’emballage faible (i.e., la canette) est plus réduit que celui de l’emballage fort (i.e., la bouteille). En revanche, lorsque leur niveau d’expérience du produit passe le seuil de 4,05, les individus n’affichent plus de biais dans leur perception du volume.

Utilisation de la macro SPSS « PROCESS »

Le lecteur intéressé peut télécharger librement la macro « PROCESS » développée pour SPSS par Andrew Hayes et téléchargeable sur son site internet (<http://afhayes.com>). Cette macro permet de conduire facilement les analyses présentées dans le cadre de cette section.

Pour installer cette macro dans SPSS, il suffit d’aller dans ‘Utilitaires’ > ‘Boîte de dialogue personnalisée’ > ‘Installer une boîte de dialogue personnalisée’, avant de sélectionner le fichier « PROCESS.spd » et de cliquer sur ‘OK’. La boîte de dialogue « PROCESS » (voir Annexe A4) sera alors accessible en allant dans ‘Analyse’ > ‘Régression’ > ‘PROCESS’.

Une fois que la macro est installée, il ne reste plus qu’à faire glisser les variables dans la boîte de dialogue. Dans le cas de notre exemple, il faut donc faire glisser la variable « Allongement » dans la case Independent Variable (X). Il faut ensuite faire glisser la variable « Expérience_Likert3Items » dans la case M Variable(s) et « Volume » dans la case Outcome Variable (Y) avant de cliquer sur ‘OK’. Il ne reste enfin qu’à sélectionner le numéro du modèle – préciser « 1 » dans Model Number – et cocher ‘Johnson-Neyman’ dans l’onglet Conditioning, avant de cliquer sur ‘Pour-suivre’ puis ‘OK’.

L’Annexe A5 présente les résultats de l’analyse de

modération tels que renvoyés par la macro SPSS à partir de notre base de données fictive.

Présentation des résultats

Si la macro PROCESS considère toutes les valeurs de la variable modératrice quantitative (M) et offre ainsi des résultats exhaustifs, ceux-ci occupent potentiellement trop de place pour apparaître tels que dans un article (e.g., 22 lignes dans notre exemple). Le chercheur est donc invité à présenter un tableau de résultats tronqué, à l'image du Tableau 3, en n'indiquant les résultats de l'estimation que pour quelques valeurs judicieusement choisies de la variable modératrice quantitative (M). D'après les recommandations adaptées de Fitzsimons (2008) et de Spiller et de ses collègues (2013), les résultats de la recherche doivent être rédigés comme suit :

Nous avons effectué une régression du volume perçu sur trois variables indépendantes : (i) le niveau d'expérience avec la catégorie de produit ($m = 2,68; \sigma = 1,80; \min = 0; \max = 6$), (ii) l'allongement de l'emballage (allongement fort : bouteille, codé 1 vs. allongement faible : canette, codé 0), et (iii) le terme d'interaction. Les résultats montrent un effet d'interaction significatif entre le niveau d'expertise du produit et l'allongement de l'emballage ($\beta = -0,56; t = -2,03; p < 0,05$). Pour décomposer cette interaction, nous avons examiné la pente du niveau d'expérience du produit à chaque niveau d'allongement de l'emballage. La pente est significative et positive lorsque l'allongement de l'emballage est faible ($\beta = 0,80; t = 4,08; p < 0,05$); elle est non significative lorsqu'il est fort ($p = 0,25$). Nous avons par ailleurs mené une analyse *floodlight* en utilisant la technique Johnson-Neyman pour identifier la zone de niveau d'expérience du produit pour laquelle il existe une différence significative de volume perçu entre les conditions expérimentales d'allongement faible (i.e., la canette) vs. fort (i.e., la bouteille) de l'emballage. Chez les individus présentant un niveau d'expérience du produit inférieur à 4,05 ($\beta_{JN} = 1,25; e.s. = 0,63; p = 0,05$), l'effet de l'allongement de l'emballage sur le volume perçu est significatif et positif : ces individus perçoivent un volume plus important dans le cas d'un allongement fort (i.e., la bouteille) vs. faible (i.e., la canette) de l'emballage. Cet effet est non si-

gnificatif chez les individus présentant un niveau d'expérience du produit supérieur à 4,05.

5 Conclusion

A la manière des théories scientifiques, dont Popper (1973) a le premier avancé le caractère fragile et provisoire, les méthodes qu'une communauté académique considère comme légitimes évoluent dans le temps. Si Baron et Kenny (1986) ont régné pendant près d'un quart de siècle sur l'analyse des variables médiatrices, l'heure est désormais à la méthode du bootstrap formulée en 2008 par Preacher et Hayes. Adieu « Baron et Kenny », bonjour « Preacher et Hayes » !

De la même façon, la discrétisation des variables quantitatives lorsqu'il s'agit de tester leur effet modérateur dans le cadre d'une analyse de données expérimentales est aujourd'hui remise en cause au profit des analyses *spotlight* et *floodlight*. Aucun manuscrit en soumission ne pourra bientôt plus faire l'économie de ces deux nouvelles méthodologies d'étude. Désormais, une fois qu'un effet d'interaction impliquant une variable quantitative sera établi, seules ces méthodologies d'étude seront considérées comme légitimes pour l'explorer plus avant. Fondées sur des régressions multiples qui consistent à étudier l'influence de la variable indépendante (X) sur la variable dépendante (Y) pour différents niveaux de la variable modératrice (M), ces méthodologies d'étude permettent de décomposer un effet d'interaction en conservant la nature quantitative de la variable modératrice (M).

Suivant la recommandation de Spiller et de ses collègues (2013), l'analyse *spotlight* est retenue lorsque la variable modératrice (M) présente une ou plusieurs valeur(s) focale(s) significative(s) : elle étudie l'influence de la variable indépendante (X) sur la variable dépendante (Y) au niveau de la ou des valeurs focale(s) significative(s) de la variable modératrice (M). Dans le cas contraire, l'analyse *floodlight* étudie l'influence de la variable indépendante (X) sur la variable dépendante (Y) pour toutes les valeurs possibles de la variable modératrice (M). Dans la pratique, l'analyse *floodlight* n'est donc qu'une répétition d'analyses *spotlight* pour chacune des valeurs de la variable modératrice (M). La plupart des échelles utilisées dans la recherche en marketing sont des échelles arbitraires non significatives ne présentant pas de valeur focale. En effet, si le développement de normes mesurées au sein d'un échantillon raisonnablement représentatif de la population de consommateurs devrait constituer l'étape ultime du développement d'une échelle (Churchill, 1979), cette étape est le plus souvent ignorée dans la recherche en marketing (Spiller et alii, 2013).

Faute d'échelle significative et de valeurs focales, l'analyse *floodlight* devrait donc davantage intéresser les chercheurs en marketing que l'analyse *spotlight*.

Par rapport à la méthode qui consiste à discrétiser la variable quantitative avant de la considérer comme facteur au sein d'une ANOVA, les analyses *spotlight* et *floodlight* présentent différents inconvénients, puisqu'elles imposent une transformation des variables indépendantes, la conduite de plusieurs tests statistiques, voire un certain « coût d'entrée » en termes de compréhension. Par-dessus tout, elles s'éloignent du cadre classique d'analyse des données expérimentales, tel qu'il a été enseigné dans les formations doctorales et diffusé par les travaux de recherche depuis de nombreuses années. Dès lors, et dans la mesure où la conduite d'un travail de recherche nécessite de comprendre le cadre d'analyse retenu et d'être en mesure d'en interpréter et d'en décrire les résultats de la manière la plus claire possible, pourquoi inciter les chercheurs en comportement du consommateur à abandonner un cadre d'analyse dont ils sont particulièrement familiers au profit d'analyses plus complexes et moins instinctives pour eux ?

Cet article offre une réponse claire : non seulement, les analyses *spotlight* et *floodlight* respectent la nature des variables quantitatives, mais surtout elles sont statistiquement « supérieures » (Iacobucci, 2001, p. 52). La discrétisation affecte clairement la capacité des chercheurs à analyser la véritable nature des relations qui existent entre leurs variables. Afin d'en convaincre concrètement les chercheurs de la communauté, cet article expose les limites de la discrétisation des variables quantitatives en prenant soin aussi souvent que possible d'illustrer empiriquement comment le mode de discrétisation de la variable modératrice (e.g., « median split haut » vs. « median split bas ») peut réduire le pouvoir statistique de l'analyse ou modifier les conclusions de la recherche. Au-delà, le présent article accompagne, de manière pédagogique, les chercheurs francophones vers les analyses *spotlight* et *floodlight*. Considérant que ces nouvelles analyses de modulation ne sont pas très accessibles à travers la littérature existante (i.e., Irwin et McClelland, 2001 ; Fitzsimons, 2008 ; Spiller et alii, 2013), il tente autant que possible de réduire le « coût d'entrée » associé à leur compréhension. Ainsi, il définit et illustre clairement ce qu'est une valeur focale significative. Il détaille ensuite les différentes étapes de l'analyse, les équations à estimer et la manière d'interpréter les résultats obtenus sous SPSS à partir d'exemples concrets tirés ou adaptés d'un seul et même article accessible en français (i.e., Yang et Raghuram, 2005, 2006). Constatant l'utilisation de plus en plus fréquente des macros sous SPSS ainsi que l'absence de macro pour automatiser l'analyse *spotlight*, il en propose une en accès

libre sur le site du premier auteur dont il détaille pas à pas l'installation et le mode d'utilisation.

Finalement, en proposant une démonstration concrète de l'ampleur des limites de la discrétisation des variables quantitatives, cet article souhaite faire du problème de la discrétisation une préoccupation majeure pour les chercheurs (Fitzsimons, 2008). Et en proposant un tutoriel didactique pour conduire les analyses *spotlight* et *floodlight* appropriées, il souhaite lever les freins à l'adoption de ces nouvelles analyses trop souvent ignorées, notamment au niveau hexagonal, et permettre aux chercheurs de mieux satisfaire aux exigences des plus grandes revues académiques. Pour des raisons de concision, cet article présente les analyses *spotlight* et *floodlight* dans le cas standard du test d'une variable modératrice quantitative sur la relation entre une variable présentant deux modalités manipulées en inter-sujets et une variable quantitative. Toutefois, une approche similaire pourrait être suivie en présence d'une variable manipulée présentant plus de deux modalités et/ou en intra-sujets, ou encore avec plusieurs variables indépendantes quantitatives (Spiller et alii, 2013). À l'instar de Chumpitaz-Caceres et Vanhamme (2003), cet article ne considère pas non plus l'étude simultanée de plusieurs relations interdépendantes. Celle-ci appelle la mobilisation d'analyses plus élaborées telles que les modèles d'équations structurelles (pour plus de détails au sujet de ces analyses, le lecteur est renvoyé à la lecture de Little et alii, 2007).

Références

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research : conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1171-1182.
- C Derbaix, C. P. et. (2002). L'attitude de l'enfant envers une nouvelle marque : de la nécessité d'une phase de fixation ? *Recherche et applications en Marketing*, 17(3), 63-79.
- Chumpitaz-Caceres, R., & Vanhamme, J. (2003). Les processus modérateurs et médiateurs : distinction conceptuelle, aspects analytiques et illustrations. *Recherche et applications en Marketing*, 18(2), 67-100.
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64-73.
- De Pechpeyrou, P., & Odou, P. (2012). Scepticisme du consommateur et efficacité promotionnelle. *Recherche et applications en Marketing*, 27(2), 45-69.
- Fitzsimons, G. J. (2008). Death to dichotomizing. *Journal of Consumer Research*, 35(1), 5-8.
- Iacobucci, D. (2001). Continuous and discrete variables. *Journal of Consumer Psychology*, 10(1/2), 37-53.

- Irwin, J. R. (2001). Treating individual difference predictors as continuous or categorical. *Journal of Consumer Psychology*, 10(1/2), 51-52.
- Irwin, J. R., & McClelland, G. H. (2001). Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research*, 38(1), 100-109.
- Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, 40(3), 366-371.
- Johnson, P. O., & Neyman, J. (1936). Tests of certain hypotheses ad their application to some educational problems. *Statistical Research Memoires*, 1, 77-59.
- Little, T. D., Card, N. A., Bovaird, J. A., Preacher, K. J., & Crandall, C. S. (2007). Structural equation modeling of mediation and moderation with contextual factors. In J. A. Bovaird & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Louviere, J. J., Hensher, D. A., & Swait, J. C. (2000). *Stated choice method : analysis and applications*. Cambridge, UK : Cambridge University Press.
- Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median-split and spurious statistical significance. *Psychological Bulletin*, 113, 181-190.
- McClelland, G. H. (1997). Optimal design in psychological research. *Psychological Methods*, 2(1), 3-19.
- Nelson, D. E., Jarman, D. W., Rehm, J., Greenfield, T. K., Rey, G., Kerr, W. C., et al. (2013). Alcohol-attributable cancer deaths and years of potential life lost in the united states. *American Journal of Public Health*, 103(4), 641-648.
- Popper, K. (1973). *La logique de la découverte scientifique*. Paris : Payot.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879-891.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology : undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359-1366.
- Soderlund, M. (2002). Consumer familiarity and its effects on satisfaction and behavioral intentions. *Psychology and Marketing*, 19(10), 861-880.
- Spiller, S. A., Fitzsimons, G. J., Lynch, J. G., & McClelland, G. H. (2013). Spotlights, floodlights, and the magic number zero : Simple effects tests in moderated regression. *Journal of Marketing Research*, 50(2), 277-288.
- Tybout, A. (2010). Treating an individual difference predictor as continuous or categorical. *Journal of Consumer Psychology*, 10(1/2), 48-49.
- Vargha, A., Rudas, T., Delaney, H. D., & Maxwell, S. E. (1996). Dichotomization, partial correlation, and conditional independence. *Journal of Educational and Behavioral Statistics*, 21(3), 264-282.
- Yang, S., & Raghurir, P. (2005). Can bottles speak volumes? The effect of package shape on how much to buy. *Journal of Retailing*, 81(4), 269-281.
- Yang, S., & Raghurir, P. (2006). Les bouteilles peuvent-elles être transcrites en volumes ? L'effet de la forme de l'emballage sur la quantité à acheter. *Recherche et applications en Marketing*, 21(1), 81-100.

Annexes

A1 - Limites de la discrétisation

Pour illustrer les limites de la discrétisation, nous utilisons la base de données fictive téléchargeable sur le site Internet du premier auteur <https://sites.google.com/site/romaincadario/>. Deux manières de discrétiser suivant le critère de la médiane sont possibles :

- Le « médiane split bas » : la variable modératrice quantitative M est codée 0 si la valeur est inférieure ou égale à la médiane, 1 si la valeur est strictement supérieure à la médiane (valeur de la médiane : 4)
- Le « médiane split haut » : la variable modératrice quantitative M est codée 0 si la valeur est strictement inférieure à la médiane, 1 si la valeur est supérieure ou égale à la médiane.

Premièrement, le résultat de la régression de la variable dépendante Y par les variables indépendantes M , X et leur terme d'interaction est sensible au « median split » retenu. Le tableau ci-dessous montre que

l'effet d'interaction est non significatif lorsque la variable M est discrétisée en utilisant un « median split haut », mais qu'il devient significatif ($p < 0,05$) lorsqu'elle est discrétisée en utilisant un « median split bas ».

	$M' = M$	$M' = \text{Med. split bas}$	$M' = \text{Med. split haut}$
Cste	26,02* (41,36)	27,21* (60,80)	27,07* (44,68)
M'	0,56* (4,62)	3,10* (4,39)	2,21* (2,89)
X'	3,78* (4,26)	2,87* (4,54)	3,13* (3,65)
$M'X$	-0,41* (-2,40)	-2,19* (-2,19)	-1,81 (-1,67)
R^2	0,34	0,33	0,23

* $p < 0,05$; t-value entre parenthèses.

Deuxièmement, la discrétisation réduit le pouvoir explicatif du modèle : le coefficient de détermination passe de 0,34 à 0,23 lorsque la variable modératrice quantitative est discrétisée en utilisant un « median split haut ».

A2 - Macro SPSS « Analyse de modération avec *spotlight* »

Analyse de modération avec spotlight

Variables

Expérience_Likert3...

Variable dépendante Y (quantitative)
Volume

Variable indépendante X (qualitative binaire)
Allongement

Modalité correspondant à X=0
Canette

Modalité correspondant à X=1
Bouteille

Variable modératrice M (quantitative)
Expérience_NbBières

Valeur focale
7

Variable(s) de contrôle (binaires ou quantitatives)

Réinitialiser Coller Annuler OK

A3 - Résultats de la macro SPSS « Analyse de modération avec *spotlight* »

 IMPORTANT: Utilisez à vos propres risques, lire l'article d'accompagnement
 IMPORTANT: Cette macro analyse l'effet modérateur d'une variable M (QUANTITATIVE) sur la relation
 entre une variable indépendante X (QUALITATIVE BINAIRE) et une variable dépendante Y (QUANTITATIVE)

 Résumé des variables de votre modèle:
 - variable indépendante X: **Allongement** (Canette=0 et Bouteille=1)
 - variable dépendante Y: **Volume**
 - variable modératrice M: **Expérience_NbBières**
 - valeur focale de la variable modératrice: M = 7
 - variable(s) de contrôle:

Coefficients^a

Modèle		Coefficients non standardisés		Coefficients standardisés	t	Sig.
		A	Erreur standard	Bêta		
1	(Constante)	26,016	,629		41,367	,000
	Expérience_NbBières	,560	,121	,607	4,620	,000
	Allongement	3,788	,889	,725	4,259	,000
	Interaction_Eq1	-,411	,171	-,465	-2,399	,019

a. Variable dépendante : Volume

* Ci-dessus: résultats de l'estimation de l'équation 1
 (1a) Cette estimation permet de tester l'existence de l'effet d'interaction entre les variables **Allongement** (X) et **Expérience_NbBières** (M): l'effet d'interaction est établi si le coefficient de **Interaction_Eq1** est significatif.
 (1b) Cette estimation permet d'explorer l'effet d'interaction lorsque la variable **Allongement** (X) est **canette**, en étudiant le coefficient de **Expérience_NbBières** (M):
 - s'il est significatif et positif, **Volume** (Y) augmente lorsque **Expérience_NbBières** (M) augmente
 - s'il est significatif et négatif, **Volume** (Y) diminue lorsque **Expérience_NbBières** (M) augmente
 - s'il est non significatif, **Volume** (Y) ne dépend pas de **Expérience_NbBières** (M).

Coefficients^a

Modèle		Coefficients non standardisés		Coefficients standardisés	t	Sig.
		A	Erreur standard	Bêta		
1	(Constante)	29,804	,629		47,390	,000
	Expérience_NbBières	,149	,121	,161	1,226	,224
	Allongement_inv	-3,788	,889	-,725	-4,259	,000
	Interaction_Eq2	,411	,171	,465	2,399	,019

a. Variable dépendante : Volume

* Ci-dessus: résultats de l'estimation de l'équation 2
 Cette estimation permet d'explorer l'effet d'interaction lorsque la variable **Allongement** (X) est **bouteille**, en étudiant le coefficient de **Expérience_NbBières** (M):
 - s'il est significatif et positif, **Volume** (Y) augmente lorsque **Expérience_NbBières** (M) augmente
 - s'il est significatif et négatif, **Volume** (Y) diminue lorsque **Expérience_NbBières** (M) augmente
 - s'il est non significatif, **Volume** (Y) ne dépend pas de **Expérience_NbBières** (M)

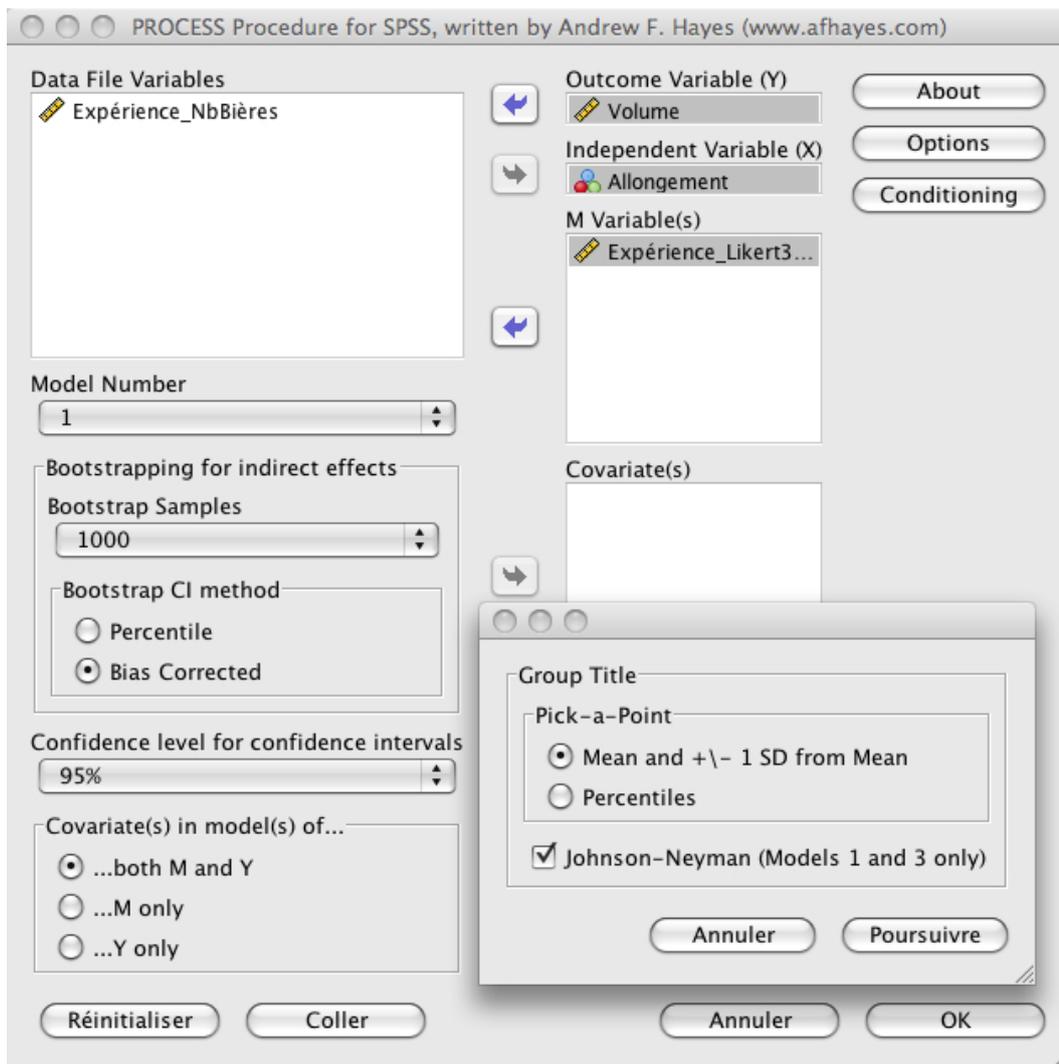
Coefficients^a

Modèle		Coefficients non standardisés		Coefficients standardisés	t	Sig.
		A	Erreur standard	Bêta		
1	(Constante)	29,933	,470		63,691	,000
	Expérience_NbBières_7	,560	,121	,607	4,620	,000
	Allongement	,911	,665	,174	1,370	,175
	Interaction_Eq3	-,411	,171	-,378	-2,399	,019

a. Variable dépendante : Volume

* Ci-dessus: résultats de l'estimation de l'équation 3
 Cette estimation permet d'explorer l'effet d'interaction avec un projecteur sur la zone de **Expérience_NbBières** (M) telle que M = 7 (valeur focale significative), en étudiant le coefficient de **Allongement** (X):
 - s'il est significatif et positif, **Volume** (Y) est plus élevé quand **Allongement** (X) est **bouteille** plutôt que **canette**
 - s'il est significatif et négatif, **Volume** (Y) est plus élevé quand **Allongement** (X) est **canette** plutôt que **bouteille**
 - s'il est non significatif, **Volume** (Y) ne dépend pas de **Allongement** (X).

A4 - Macro SPSS « PROCESS »



A5 - Résultats de la macro SPSS « PROCESS »

 Outcome: Volume

Model Summary

R	R-sq	F	df1	df2	p
,5571	,3104	11,4011	3,0000	76,0000	,0000

Model

	coeff	se	t	p
constant	26,3002	,6331	41,5390	,0000
Expérien	,7965	,1950	4,0843	,0001
Allongem	3,5328	,8960	3,9428	,0002
int_1	-,5643	,2782	-2,0280	,0461

Interactions:

int_1 Allongem X Expérien

R-square increase due to interaction(s):

	R2-chng	F	df1	df2	p
int_1	,0373	4,1130	1,0000	76,0000	,0461

Conditional effect of X on Y at values of the moderator(s)

Expérien	Effect	se	t	p
,8782	3,0373	,7060	4,3021	,0000
2,6784	2,0215	,4977	4,0619	,0001
4,4786	1,0057	,7062	1,4242	,1585

Values for quantitative moderators are the mean and plus/minus one SD from mean

***** JOHNSON-NEYMAN TECHNIQUE *****

Moderator value(s) defining Johnson-Neyman significance region(s)

4,0481

Conditional effect of X on Y at values of the moderator (M)

Expérien	Effect	se	t	p	LLCI	ULCI
,0000	3,5328	,8960	3,9428	,0002	1,7482	5,3174
,3000	3,3635	,8279	4,0627	,0001	1,7146	5,0124
,6000	3,1942	,7629	4,1872	,0001	1,6749	4,7136
,9000	3,0250	,7017	4,3108	,0000	1,6274	4,4225
1,2000	2,8557	,6456	4,4234	,0000	1,5699	4,1415
1,5000	2,6864	,5959	4,5081	,0000	1,4996	3,8732
1,8000	2,5171	,5544	4,5404	,0000	1,4130	3,6213
2,1000	2,3478	,5230	4,4892	,0000	1,3062	3,3895
2,4000	2,1786	,5036	4,3256	,0000	1,1755	3,1816
2,7000	2,0093	,4977	4,0371	,0001	1,0180	3,0005
3,0000	1,8400	,5057	3,6388	,0005	,8329	2,8471
3,3000	1,6707	,5269	3,1709	,0022	,6213	2,7201
3,6000	1,5014	,5599	2,6817	,0090	,3863	2,6165
3,9000	1,3321	,6027	2,2102	,0301	,1317	2,5326
4,0481	1,2486	,6269	1,9917	,0500	,0000	2,4971
4,2000	1,1629	,6534	1,7796	,0791	-,1386	2,4643
4,5000	,9936	,7104	1,3986	,1660	-,4213	2,4085
4,8000	,8243	,7722	1,0675	,2891	-,7136	2,3622
5,1000	,6550	,8377	,7819	,4367	-1,0135	2,3235
5,4000	,4857	,9062	,5360	,5935	-1,3192	2,2907
5,7000	,3165	,9771	,3239	,7469	-1,6295	2,2624
6,0000	,1472	1,0498	,1402	,8889	-1,9436	2,2379

***** ANALYSIS NOTES AND WARNINGS *****