L'année psychologique

L'analyse des variables indépendantes continues et catégorielles : alternatives à la dichotomisation

Markus Brauer

Abstract

Summary : The analysis of continuous and categorical independent variables : Alternatives to dichotomization.
Transforming continuous independent variables into categorial ones makes the statistical analyses simpler because after dichotomization, the effects of these variables can be examined via an analysis of variance (ANOVA) rather than a multiple regression analysis. However, this simplicity comes at a high price. When a continuous variable is dichotomized, one artificially introduces random error which decreases the statistical power of the inferential analyses. Assuming a normal distribution, the decrease in statistical power is equivalent to the exclusion of approximately 38 % of the participants. In this article, we present the problems associated with the dichotomization of continuous variables and we discuss various strategies that allow researchers to analyze experimental designs with continuous and categorical independent variables.
Key words : continuous variables, quantitative variables, regression analysis, mean deviation form, dichotomization.

Résumé

Résumé
Le fait de transformer une variable indépendante continue en une variable catégorielle à deux modalités rend les analyses statistiques plus faciles car après dichotomisation, les effets de cette variable peuvent être examinés grâce à une analyse de variance (ANOVA) plutôt qu'une analyse de régression multiple. Or, cette facilité a un prix élevé. En dichotomisant une variable continue, on introduit artificiellement de l'erreur aléatoire ce qui diminue la puissance statistique des analyses inférentielles. Supposant une distribution normale, la diminution de la puissance statistique correspond à l'exclusion d'environ 38 % des participants. Dans cet article, nous présentons les problèmes associés à la dichotomisation des variables continues et nous discutons différentes stratégies statistiques permettant d'analyser des plans expérimentaux contenant des variables indépendantes continues et catégorielles.
Mots-clés : variables continues, variables quantitatives, analyse de régression, forme centrée, dichotomisation.

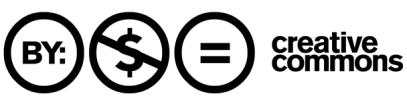
Citer ce document / Cite this document :

Brauer Markus. L'analyse des variables indépendantes continues et catégorielles : alternatives à la dichotomisation. In: L'année psychologique. 2002 vol. 102, n°3. pp. 449-484.

doi : 10.3406/psy.2002.29602

http://www.persee.fr/doc/psy_0003-5033_2002_num_102_3_29602

Document généré le 24/09/2015


BY:
creative
commons

L'Année psychologique, 2002, *102*, 449-484

NOTE MÉTHODOLOGIQUE

Laboratoire de Psychologie Sociale de la Cognition
Université Blaise-Pascal[1]
CNRS UNR 6024

L'ANALYSE DES VARIABLES INDÉPENDANTES CONTINUES ET CATÉGORIELLES : ALTERNATIVES À LA DICHOTOMISATION

Markus BRAUER[2, 3]

Persée creative commons

SUMMARY : The analysis of continuous and categorical independent variables : Alternatives to dichotomization.

Transforming continuous independent variables into categorial ones makes the statistical analyses simpler because after dichotomization, the effects of these variables can be examined via an analysis of variance (ANOVA) rather than a multiple regression analysis. However, this simplicity comes at a high price. When a continuous variable is dichotomized, one artificially introduces random error which decreases the statistical power of the inferential analyses. Assuming a normal distribution, the decrease in statistical power is equivalent to the exclusion of approximately 38 % of the participants. In this article, we present the problems associated with the dichotomization of continuous variables and we discuss various strategies that allow researchers to analyze experimental designs with continuous and categorical independent variables.

Key words : continuous variables, quantitative variables, regression analysis, mean deviation form, dichotomization.

1. 34, avenue Carnot, 63037 Clermont-Ferrand Cedex.
2. E-mail : Markus.Brauer@srvpsy.univ-bpclermont.fr.
3. Je voudrais remercier Gary McClelland, Patrick Lemaire, Sophie Berjot, Frédérick Grouzet, Sophie Brunot, Ghislaine Richard, Christelle Maisonneuve, Charles Judd, ainsi qu'un rapporteur anonyme pour leurs commentaires sur des versions antérieures de cet article.

Dans les sciences sociales, une pratique courante est de dichotomiser les variables indépendantes continues avant d'analyser leur effet sur la variable dépendante. À partir du fameux « median split », les participants sont divisés en deux groupes, et les différences entre ces deux groupes sont examinées avec une analyse de variance (ANOVA). Depuis longtemps, cette façon de procéder a été critiquée par des statisticiens comportementalistes et des psychologues quantitativistes (Peters et Van Voorhis, 1940 ; Humphreys et Fleishman, 1974 ; Humphreys, 1978 ; Cohen, 1983, 1990 ; Maxwell et Delaney, 1993). En effet, Cohen (1983) a démontré qu'en cas de distribution normale, le fait de dichotomiser une variable indépendante continue entraîne une diminution de la puissance statistique du test qui correspond à l'exclusion de 38 % des participants. Cette perte de puissance statistique se produit parce que le chercheur introduit artificiellement de l'erreur aléatoire dans les données quand il transforme une variable indépendante continue en variable catégorielle dichotomique. Le but de cet article est d'identifier le problème et de suggérer des techniques statistiques permettant d'analyser des plans expérimentaux contenant des variables continues et catégorielles.

LE PRIX DE LA DICHOTOMISATION

Pour illustrer le problème, prenons un exemple simple avec une seule variable indépendante. Imaginons un psycholinguiste qui étudie l'effet du niveau d'expertise dans une langue étrangère sur la représentation lexicale des mots dans les langues maternelle et étrangère (Mägiste, 1984 ; Brauer, 1998). Grâce à un test de langue, il mesure le niveau d'anglais de 12 étudiants français qui sont inscrits en première année de psychologie. Ce test fournit une note entre 0 (aucune connaissance en anglais) et 20 (parfaitement bilingue). Ensuite, le psycholinguiste fait faire aux étudiants une version bilingue de la tâche de Stroop (1935) permettant de mesurer à quel point les mots français interfèrent avec la production des mots anglais. Un score d'interférence faible indique que le lexique anglais est stocké indépendamment du lexique français. En revanche, plus le score d'interférence est élevé, plus les deux lexiques sont associative-

ment liés. Le psycholinguiste prédit alors une relation négative entre le niveau d'anglais et l'interférence de façon à ce que mieux un étudiant parle anglais, moins il y aura interférence entre les deux langues. Les données de cette étude hypothétique apparaissent dans le tableau I.

TABLEAU I. — *Les données de l'étude hypothétique sur la relation entre le niveau d'expertise dans une langue étrangère et la représentation lexicale des mots dans les langues maternelle et étrangère*

The data of the hypothetical study on the relationship between proficiency in a foreign language and lexical representation of words in the mother tongue and the foreign language

Numéro de Participant	Note d'Anglais	Score d'Interférence
1	2	95
2	5	45
3	6	80
4	7	40
5	8	45
6	8	20
7	10	70
8	10	35
9	11	40
10	12	20
11	13	30
12	16	5
Moyenne	9.00	43.75
Ecart-type	3.81	26.21

Pour analyser ses données, le psycholinguiste commence par dichotomiser sa variable indépendante. Pour cela, sur la base de la note médiane au test d'anglais, il sépare ses participants en 2 groupes : tous les participants avec une note en anglais en dessous de la médiane (participants 1 à 6) obtiennent un score de 1,

et constituent le groupe des « faibles ». Tous ceux en dessus de la médiane (participants 7 à 12) obtiennent un score de 2, et constituent le groupe des « bons ». Ensuite, le psycholinguiste fait des analyses descriptives et découvrent que les « faibles » ont un score d'interférence moyen de 54,17 (s = 27,82) alors que celui des « bons » est de 33,33 (s = 21,83). Finalement, il effectue une ANOVA uni-factorielle avec un facteur interparticipant (*i.e.*, niveau d'expertise en anglais) et les scores d'interférence comme variable dépendante. À sa déception, la différence entre les « faibles » et les « bons » n'est pas statistiquement significative, $F(1,10) = 2,08$, $p = .18$[1].

Comme Cohen (1983) et d'autres statisticiens nous l'ont fait remarquer, le fait de dichotomiser une variable indépendante continue est coûteux car cela réduit notre puissance statistique de détecter un effet. Le psycholinguiste aurait mieux fait d'analyser sa variable indépendante telle qu'elle est, c'est-à-dire en tant que variable continue. Pour cela, il aurait pu faire une analyse de régression dans laquelle le score d'interférence est régressé sur la note en anglais. En faisant cela, on aurait obtenu l'équation de régression suivante : $\text{INTERFER} = b_0 + b_1\text{ANGLAIS} = 90,43 - 5,19\text{ANGLAIS}$. Dans cette équation, INTERFER correspond au score d'interférence et ANGLAIS à la note dans le test d'anglais. L'analyse inférentielle montre que b_1 est significativement inférieur à zéro, $F(1,10) = 13,24$, $p = .005$. En conclusion, la relation négative entre le score d'interférence et la note en anglais est significative, et l'hypothèse du psycholinguiste semble exacte.

Pourquoi ces conclusions différentes ? Pourquoi le psycholinguiste obtient-il un résultat non significatif quand il traite sa variable indépendante en tant que variable catégorielle alors que ce même résultat est significatif quand il la traite en tant que variable continue ? Comme nous allons le démontrer ci-dessous, le fait de dichotomiser une variable indépendante introduit artificiellement de l'erreur aléatoire (voir aussi Irwin et McClelland, 2002). Les données de l'étude hypothétique sont montrées dans la figure 1. La note en anglais (la variable indépendante) est sur l'abscisse, le score d'interférence (la variable

1. Bien sûr, il aurait pu faire un test t avec deux échantillons indépendants qui aurait donné exactement le même résultat : $t(10) = 1,44$, $p = .18$ (le F étant égal à t^2).

dépendante) sur l'ordonnée. Apparaît également dans la figure 1 la droite de régression dont l'équation de régression est mentionnée plus haut (voir ligne noire).

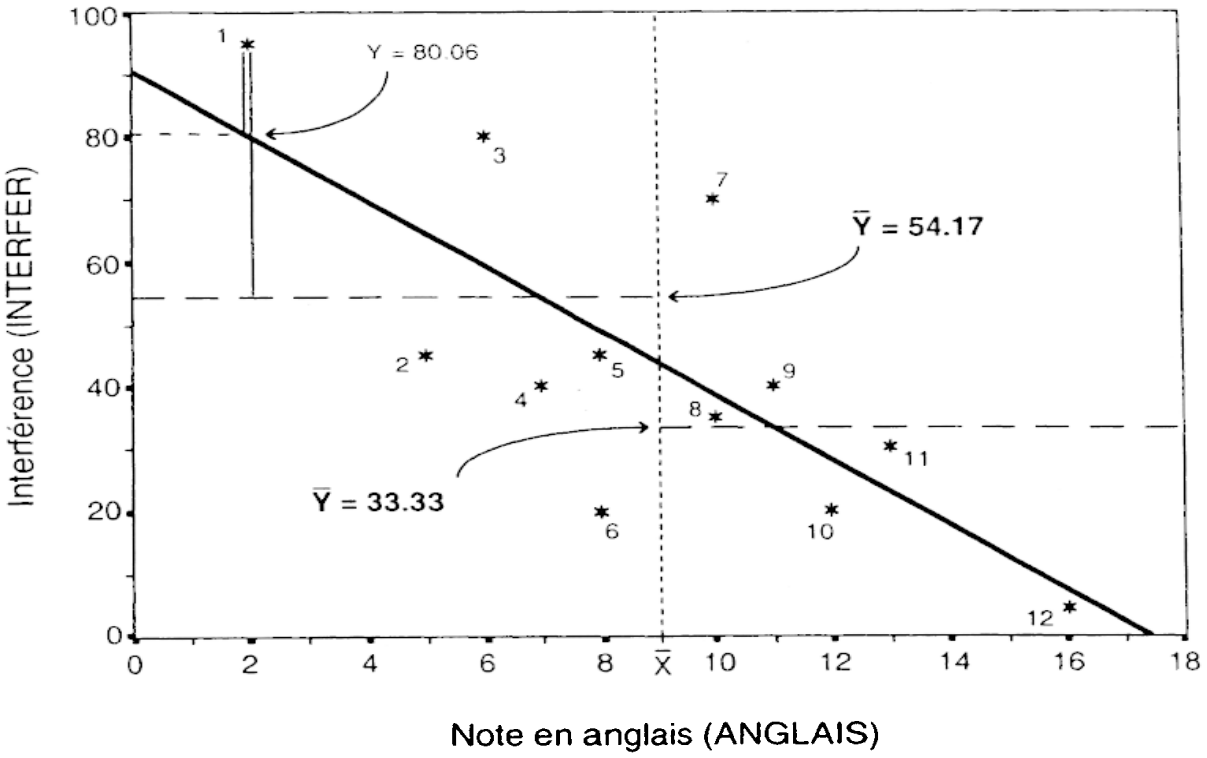


Fig. 1. — Les résultats de l'étude hypothétique sur la relation entre le niveau d'expertise dans une langue étrangère et la représentation lexicale des mots dans les langues maternelle et étrangère

The results of the hypothetical study on the relationship between proficiency in a foreign language and lexical representation of words in the mother tongue and the foreign language

Rappelons que la logique d'une régression est la suivante : connaissant le score d'un participant sur la variable indépendante, quelle prédiction peut-on faire quant à son score sur la variable dépendante ? La droite de régression correspond aux prédictions du modèle de régression. Par définition, cette droite de régression est celle qui fait les « meilleures » prédictions, c'est-à-dire qu'il est impossible de trouver une autre droite qui minimise la variance résiduelle autant que la droite de régression (Cohen et Cohen, 1983 ; Judd et McClelland, 1989). La variance résiduelle correspond à la somme des carrés de toutes les diffé-

rences entre les prédictions du modèle de régression et les observations réelles. Par exemple, le résidu pour le participant n° 1 est (95 – 80,06) = 14,94. Les résidus de tous les participants sont élevés au carré et additionnés. Ils permettent d'estimer la partie de la variance de la variable dépendante qui n'est pas expliquée par la variable indépendante (*i.e.*, la somme des carrés résidus ou la variance résiduelle). Bien sûr, plus la variable indépendante explique de la variance dans la variable dépendante, plus l'effet est grand. Et plus on introduit de l'erreur aléatoire artificiellement, plus la variance résiduelle augmente, et plus l'effet est petit.

Que se passe-t-il alors quand on dichotomise une variable indépendante pour se retrouver avec deux groupes, les « bons » et les « faibles » ? L'ANOVA s'intéresse au score moyen à l'intérieur de chaque groupe, et toute déviation de ce score moyen est considérée comme résidu. Si l'on utilise le langage de la régression, on pourrait dire que la dichotomisation nous amène à générer deux droites de régression, une pour les participants en dessous de la médiane et une pour les participants en dessus de la médiane. Ces droites sont par définition horizontales car tous les participants à l'intérieur de chaque groupe sont considérés être pareils. Et la hauteur de chacune des droites de régression correspond à la moyenne du groupe expérimental. En effet, la meilleure prédiction que l'ANOVA peut faire pour un participant est le score moyen du groupe expérimental auquel il appartient. Les deux droites de régression correspondant aux prédictions de l'ANOVA après dichotomisation sont également montrées dans la figure 1 (voir lignes pointillées).

Représentée de cette manière, il est facile de voir que la dichotomisation augmente la variance résiduelle de façon artificielle. Prenons le participant n° 1. Alors que son résidu était de 14,94 dans l'analyse où la note en anglais est traitée en tant que variable continue, il passe à 40,83 (= 95 – 54,17) dans l'analyse où la note en anglais a été dichotomisée. Certes, les résidus de certains participants sont plus grands dans la première analyse que dans la deuxième (par exemple celui du participant n° 2), mais cela n'empêche pas que la somme de tous les carrés des résidus (la variance résiduelle) est plus grande si la note en anglais est dichotomisée que si elle ne l'est pas. De plus, nous savons que ce n'est pas une spécificité de notre exemple

hypothétique car la droite de régression produit par définition la variance résiduelle la plus petite. En dichotomisant, on fait la même prédiction pour quelqu'un qui parle relativement bien l'anglais (*e.g.*, le participant n° 8) que pour quelqu'un qui est parfaitement bilingue (*e.g.*, le participant n° 12). Supposant une distribution normale, il est évident que l'analyse de régression qui se fonde sur la note exacte dans le test d'anglais et non pas sur des catégories aussi grossières que les « bons » et les « faibles » fera de meilleures prédictions.

Le but de cet exemple était d'illustrer le prix à payer pour la dichotomisation. Certes, l'analyse de données est plus facile si nous transformons nos variables continues en variables catégorielles. Ceci est d'autant plus vrai si l'on a affaire à des plans multifactoriels avec plusieurs variables catégorielles et continues. En revanche, nous payons cette facilité par une augmentation artificielle de la variance résiduelle et, donc, par une diminution de nos chances de détecter un effet qui serait présent. Ceci revient à jeter un dé pour chaque participant et à ajouter à son score le chiffre indiqué sur le dé. Cohen (1983) a pu démontrer qu'en cas d'une distribution normale, la variance expliquée par une variable indépendante dichotomisée correspond à 64,7 % de la variance expliquée par cette même variable si elle est traitée en tant que variable continue. Cette perte de puissance statistique, nous l'avons déjà dit, correspond à l'exclusion d'environ 38 % des participants.

Avant de suggérer des solutions à ce problème, un commentaire s'impose. La plupart des lecteurs pensent peut-être que jamais ils n'auraient utilisé une ANOVA pour analyser les données du psycholinguiste. Après tout, il est évident qu'une analyse de régression, ou même une analyse de corrélation, est la procédure statistique la plus appropriée. Le but de l'exemple hypothétique était de démontrer le coût de la dichotomisation. Nous voulions montrer comment de l'erreur aléatoire est introduite artificiellement en traitant une variable indépendante continue en tant que variable catégorielle. Pour cela, nous avons délibérément choisi un exemple simple avec une seule variable indépendante. En revanche, le même raisonnement s'applique à des plans expérimentaux avec plusieurs variables indépendantes continues et catégorielles. Or, c'est justement dans ces plans expérimentaux multifactoriels que les psychologues ont le plus

souvent recours à la dichotomisation. Dans le reste de cet article, nous allons aborder plusieurs plans multifactoriels et proposer des stratégies d'analyse.

LA FORME CENTRÉE

La première étape de l'analyse de données est commune à l'ensemble des solutions proposées ci-dessous : la variable indépendante continue est transformée en « forme centrée » (anglais : mean deviation form). Cette transformation consiste à déduire de la valeur de chaque observation la moyenne générale de cette variable. Dans notre exemple psycholinguistique, la moyenne des notes en anglais est de 9 (voir le tableau I). Donc, pour transformer la variable indépendante en forme centrée on soustraira 9 de la note de chacun des participants. Par exemple, le nouveau score d'anglais du participant n° 1 sera de $ANGLFC_1 = ANGLAIS_1 - 9 = 2 - 9 = -7$[1].

En régressant les scores d'interférence sur les nouvelles notes d'anglais en forme centrée, on obtient l'équation de régression suivante : $INTERFER = b_0 + b_1 ANGLFC = 43{,}75 - 5{,}19 ANGLFC$. En comparant les résultats de cette analyse à l'analyse de régression avec les notes en anglais en forme brute, on se rend compte que peu de choses ont changé. Le coefficient de régression b_1 n'a pas changé, et son interprétation non plus. b_1 a toujours la valeur de $-5{,}19$ ce qui indique que, pour chaque point supplémentaire en anglais, le modèle de régression diminue sa prédiction de 5,19 unités. Le F associé à b_1 est toujours significatif, $F(1,10) = 13{,}24$, $p = .005$, ce qui indique que la transformation en forme centrée n'a pas affecté la significativité de la relation entre la variable indépendante et la variable dépendante. En revanche, le coefficient de régression b_0 (l'ordonnée à l'origine) a changé. Alors qu'il était de 90,44 dans

1. Cette transformation correspond à la première étape du calcul des scores standardisés (*i.e.*, la « forme centrée réduite »). En effet, on pourrait tout simplement standardiser les variables indépendantes continues, car la logique des analyses décrites plus loin sera la même. En revanche, nous ne conseillons pas cette pratique. Les coefficients de régression sont plus faciles à interpréter si la variable indépendante continue est en « forme centrée » plutôt qu'en « forme centrée réduite ».

la première analyse de régression, il prend la valeur de 43,75 dans l'analyse de régression où les notes d'anglais sont en forme centrée. Dans toute analyse de régression, l'ordonnée à l'origine correspond à la prédiction du modèle de régression pour un participant ayant un score de zéro sur la variable indépendante. Par conséquent, la première analyse montre que la meilleure prédiction que nous pouvons faire pour un participant ne parlant pas un mot d'anglais est de 90,44. L'ordonnée à l'origine de la deuxième analyse nous indique la prédiction de notre modèle de régression pour un participant qui a un score de zéro sur la variable indépendante transformée en forme centrée. En d'autres termes, c'est la prédiction pour le « participant moyen ». Elle prend la valeur de 43,75, ce qui correspond à la moyenne de la variable dépendante (voir le tableau I).

Ce que l'on observe ici s'applique à toutes les stratégies analytiques présentées plus loin. La transformation en forme centrée n'affecte pas l'interprétation des coefficients de régression associés à la variable que l'on vient de transformer mais elle affecte l'interprétation des autres coefficients dans l'équation de régression. Ici, b_0 n'est pas pertinent car on ne s'intéresse pas à la question de savoir si le score moyen sur la variable interférence est significativement supérieur à zéro. Mais parfois, les autres coefficients de régression correspondent à une comparaison théoriquement pertinente (par exemple, à l'effet principal d'une deuxième variable indépendante qui, elle, est catégorielle). Dans ce cas, les coefficients de régression ne sont directement interprétables que si la variable indépendante continue est en forme centrée.

Ajoutons que les variables indépendantes catégorielles se doivent aussi d'être transformées. En effet, elle doivent être codées en contrastes orthogonaux centrés (anglais : « contrast codes »). Par exemple, une variable indépendante catégorielle à deux modalités doit être codée en – 1 et + 1. Un codage en 0 et 1 ou en 1 et 2 n'est pas approprié car une interprétation directe des coefficients de régression ne sera pas possible. C'est seulement dans le cas où les variables indépendantes catégorielles sont codées en contrastes orthogonaux qu'il est possible d'interpréter les « autres » coefficients de régression, c'est-à-dire les coefficients de régression qui ne sont pas associés à ces variables.

UNE VARIABLE INDÉPENDANTE CONTINUE EN PRÉSENCE D'UNE VARIABLE INDÉPENDANTE CATÉGORIELLE INTERPARTICIPANT

Imaginons que le psycholinguiste dans notre exemple hypothétique ait oublié de nous dire que certains de ses participants ont eu l'occasion de se familiariser avec la version bilingue de la tâche de Stroop avant l'expérience proprement dite alors que les autres participants n'ont pas eu cette occasion. Bien sûr, nous voudrions savoir si cette familiarisation préalable module la relation négative entre le niveau en anglais et l'interférence des mots français sur les mots anglais. Pour l'illustration, imaginons que les participants n° 2, 4, 6, 8, 9 et 11 ont eu l'occasion de se familiariser avec la tâche.

Face à une telle situation, certains chercheurs dichotomiseraient la variable indépendante continue et analyseraient les données via une ANOVA 2 × 2 avec deux facteurs interparticipant : Niveau en anglais (« bons », « faibles ») × Familiarité avec la tâche (« familier », « non familier »). Les moyennes correspondant à ce plan 2 × 2 sont les suivants : M = 35,00 (bons-familier), M = 31,67 (bons - non familier), M = 35,00 (faibles-familier), et M = 73,33 (faibles - non familier). L'effet principal du facteur « niveau en anglais » teste la prédiction du psycholinguiste, $F(1,8) = 2{,}58$, $p = .15$. L'interaction entre les deux facteurs teste si cet effet est modulé par l'occasion de se familiariser avec la tâche, $F(1,8) = 2{,}58$, $p = .15$[1]. À partir de ces analyses, on serait tenté de conclure que les données n'étayent pas la prédiction du psycholinguiste. De plus, les analyses inférentielles ne permettent pas de conclure que la relation entre le niveau d'anglais et les scores d'interférence est modulée par la familiarité avec la tâche.

En revanche, n'oublions pas que ces analyses sont conduites sur des notes d'anglais auxquelles on a ajouté artificiellement de l'erreur aléatoire. Les différents effets mentionnés ci-dessus sont-ils significatifs si l'on analyse les données telles qu'elles sont, c'est-à-dire si l'on traite le niveau d'anglais en tant que variable continue ? Pour cela, il faut effectuer une analyse de

1. C'est un hasard que les deux valeurs pour F et *p* sont identiques.

régression multiple où les scores d'interférence sont régressés sur trois variables indépendantes : 1 / les notes d'anglais en forme centrée (ANGLFC) ; 2 / la familiarité avec la tâche codée en – 1 et + 1 (FAMIL), et 3 / une troisième variable nouvellement créée qui correspond au produit des deux variables précédentes (INTERFC = ANGLFC * FAMIL). Appliqué à notre exemple hypothétique, on se retrouve avec l'équation de régression suivante : INTERFER = b_0 + b_1ANGLFC + b_2FAMIL + b_3INTERFC. La décomposition de la variance de cette analyse de régression multiple est présentée dans le tableau II.

TABLEAU II. — ***La décomposition de la variance de l'étude hypothétique sur la relation entre le niveau d'expertise dans une langue étrangère et la représentation lexicale des mots dans les langues maternelle et étrangère***

The source table of the hypothetical study on the relationship between proficiency in a foreign language and lexical representation of words in the mother tongue and the foreign language

Source de variation	b	Somme des carrés	ddl	*F*	*P*
ANGLAIS EN FORME CENTRÉE					
Modèle		6189.21	3	12.07	.002
[Ordonnée à l'orig.]	43.75	[22969.00]	--	134.42	.0001
ANGLFC	-3.86	1847.96	1	10.81	.011
FAMIL	-8.75	918.75	1	5.38	.049
INTERFC	2.79	964.84	1	5.65	.045
Erreur		1367.04	8		
Total		7556.25	11		

Note : ANGLFC = Note en anglais en forme centrée ; FAMIL = Familiarité avec la tâche ; INTERFC = Interaction entre ANGLFC et FAMIL.

La figure 2 fait apparaître les résultats de cette analyse de régression multiple. Les participants dans la condition « familier » sont représentés par des carrés blancs, ceux dans la condition « non familier » sont représentés par des carrés noirs. Le coefficient de régression b_0 (l'ordonnée à l'origine) est égal à la moyenne générale des scores d'interférence (voir le tableau I). Sa significativité est sans intérêt car la question de savoir si la moyenne générale est significativement supérieure à zéro n'est pas pertinente théoriquement.

Le coefficient de régression b_1 correspond à ce que l'on appellerait dans une ANOVA l'effet principal du niveau d'anglais. En effet, ce coefficient exprime la relation entre les scores d'interférence et le niveau d'anglais pour les participants ayant un score de zéro sur la variable FAMIL. Sachant que la familiarité

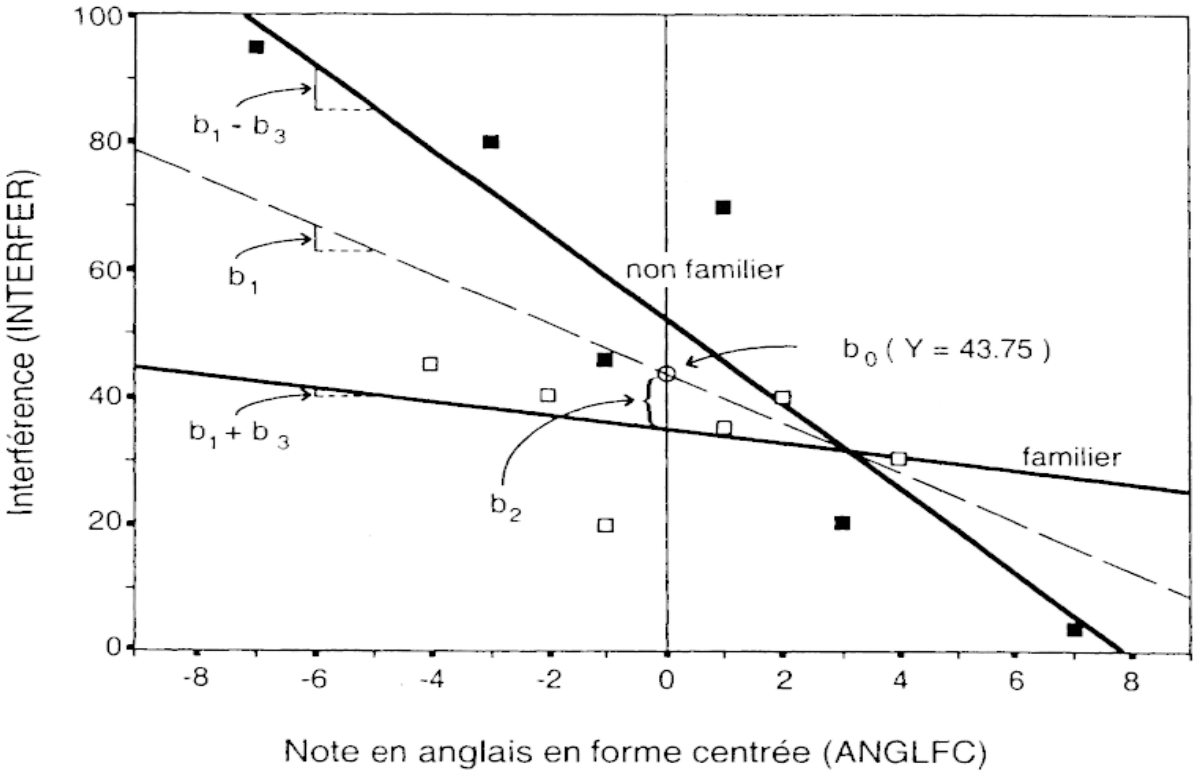


Fig. 2. — Les résultats de l'étude hypothétique sur l'effet interactif du niveau d'expertise dans une langue étrangère et de la familiarité avec la tâche sur la représentation lexicale des mots. La variable indépendante continue (la note en anglais) a été transformée en forme centrée et est traitée en tant que variable continue

The results of the hypothetical study on the interactive effect of proficiency in a foreign language and familiarity with the task on the lexical representation of words. The continuous independent variable (the grade in English) has been transformed into mean deviation form and is treated as a continuous variable

avec la tâche est codée en – 1 et + 1, le score de zéro sur FAMIL correspond au « participant moyen », c'est-à-dire celui qui est conceptuellement à mi-chemin entre la condition « non familier » (codée – 1) et la condition « familier » (codée + 1). Graphiquement, le coefficient b_1 correspond à la pente de la ligne discontinue dans la figure 2. Le coefficient de régression b_2 représente ce que l'on appellerait dans l'ANOVA l'effet principal du facteur familiarité. En d'autres termes, b_2 exprime la relation entre les scores d'interférence et la familiarité avec la tâche pour les participants ayant un score de zéro sur la variable ANGLFC. Et nous savons déjà qu'un score de zéro sur ANGLFC correspond au « participant moyen », c'est-à-dire celui avec un niveau d'expertise moyen en anglais. Graphiquement, b_2 correspond à la distance entre la moyenne générale et le point de croisement de chacune des deux droites de régression et l'ordonnée (voir fig. 2). Le coefficient b_3 correspond à l'effet d'interaction entre les deux variables indépendantes. Sans vouloir rentrer dans les détails, contentons-nous de constater que b_3 exprime à quel point la relation entre les scores d'interférence et le niveau d'anglais dépend de la familiarité avec la tâche. Graphiquement, b_3 correspond à la différence de pente entre la ligne discontinue et chacune des lignes continues.

En résumé, l'analyse de régression nous donne des résultats qui, dans leur forme, ressemblent à ceux d'une ANOVA 2 × 2. On obtient deux effets simples, qui correspondent aux effets principaux d'une ANOVA, et un effet d'interaction qui teste à quel point la relation entre la variable dépendante et l'une des variables indépendantes est modulée par l'autre variable indépendante. Certes, la procédure est un peu plus compliquée car il a fallu : *a)* transformer la variable indépendante continue en forme centrée ; *b)* recoder la variable indépendante catégorielle en – 1 et + 1 ; et *c)* créer la variable d'interaction en la calculant explicitement dans le logiciel d'analyse de données. Mais le gain de ce surplus de travail est considérable. L'analyse de régression montre que la relation entre les scores d'interférence et le niveau d'anglais est significative, $F(1,8) = 10{,}81$, $p < .02$ (alors qu'elle ne l'était pas dans l'ANOVA portant sur les notes d'anglais dichotomisées, $F(1,8) = 2{,}58$, $p = .15$). De plus, cette analyse de régression montre que cette relation dépend du niveau de familiarité avec la tâche, $F(1,8) = 5{,}65$, $p < .05$ (cette même interaction était loin d'être significative dans l'ANOVA, $F(1,8) = 2{,}58$, $p = .15$).

LA PRÉSENTATION DES RÉSULTATS

Un problème supplémentaire se pose au chercheur : Comment présenter les résultats de l'analyse de régression à la communauté scientifique ? Dans le cas de l'ANOVA, la solution à ce problème est simple car il suffit de calculer les moyennes pour les quatre conditions expérimentales et de rapporter ces moyennes dans un tableau ou dans une figure. En revanche, nous ne pouvons pas procéder de cette manière dans l'analyse de régression car la distinction entre les « bons » et les « faibles » en anglais n'existe plus. Deux possibilités s'offrent au chercheur.

Premièrement, il peut dichotomiser la variable indépendante continue, générer les moyennes des quatre conditions expérimentales, et présenter ce tableau de moyennes dans l'article ou la communication. Il n'y a aucune objection à la façon de procéder qui consiste à traiter la variable indépendante continue en tant que variable catégorielle dans les analyses descriptives mais en tant que variable continue dans les analyses inférentielles[1]. Cette solution est à préférer si l'on s'adresse à un public n'ayant que de faibles connaissances en analyse de régression comme c'est souvent le cas dans la recherche appliquée.

Deuxièmement, le chercheur peut créer un graphique avec plusieurs droites de régression, une pour chacune des modalités de la variable indépendante catégorielle. La plupart des logiciels d'analyse de données permettent de produire un tel graphique en peu de temps. Par exemple, en SPSS, il faut demander un diagramme de dispersion, où la variable dépendante apparaît sur l'axe y, la variable indépendante continue sur l'axe x, et l'on « définit les marques » par la variable indépendante catégorielle. Ensuite, on demande une « courbe d'ajustement » (une droite de régression) pour chacune des modalités de la variable indépendante catégorielle (appelées « sous-groupes » en SPSS). Finalement, il convient d'effacer les données individuelles pour ne garder que les droites de régression. Un autre moyen d'obtenir ce

1. Cela se fait souvent. Ne pensons qu'aux analyses des temps de latence où l'on fait les analyses inférentielles sur des données que l'on a soumises auparavant à une transformation logarithmique (ou une autre), mais où l'on rapporte les données brutes dans les tableaux des moyennes (Fazio, 1990 ; Bargh et Chartrand, 2000).

graphique est de régresser la variable dépendante sur la variable indépendante pour chaque modalité de la variable indépendante catégorielle séparément et de faire apparaître les droites de régression dans un diagramme superposé.

En tant qu'exemple, on peut consulter les figures qui seront présentés ultérieurement dans cet article (voir fig. 3 et 4). Le graphique gagne en lisibilité si : *a)* on utilise la variable indépendante en forme brute et non pas en forme centrée (donc ANGLAIS et non pas ANGLFC dans notre exemple) ; *b)* on ajoute une ligne de référence verticale correspondant au niveau de la moyenne de la variable indépendante continue (car c'est à ce niveau que l'effet principal de la variable indépendante catégorielle est testé) ; *c)* les droites de régression ne sont pas tracées au-delà des valeurs que prend réellement la variable indépendante continue ; *d)* les équations de régression sont indiquées, soit dans la figure soit dans le texte.

UNE VARIABLE INDÉPENDANTE CONTINUE EN PRÉSENCE D'UNE VARIABLE INDÉPENDANTE CATÉGORIELLE INTERPARTICIPANT AVEC PLUS DE DEUX MODALITÉS

Le grand avantage d'une ANOVA est qu'elle gère convenablement des variables indépendantes catégorielles avec plus de deux modalités. Comme nous allons le voir, cette aisance manque dans l'analyse de régression. La plupart des logiciels d'analyse de données contiennent des modules d'analyse de régression multiple qui nous permettent d'obtenir facilement les effets correspondants à des contrastes spécifiques d'une variable catégorielle à plusieurs modalités. En revanche, ces logiciels ne nous donnent pas les effets principaux, c'est-à-dire ceux avec plus d'un degré de liberté. Souvent, il est nécessaire d'effectuer des analyses de régression hiérarchiques pour obtenir la valeur de F d'un effet principal avec deux (ou plus) degrés de liberté.

En tant qu'exemple, imaginons une chercheuse qui voudrait tester l'hypothèse selon laquelle nos jugements sur un objet s'appuient sur la valence des concepts associés à cet objet quand nous sommes tristes ou dans un état émotionnel neutre, mais

que nos jugements sont indépendants des concepts associés quand nous sommes en colère (Bodenhausen, Sheppard et Kramer, 1994). Pour tester sa prédiction, elle présente à ses participants un tableau de peinture et leur demande de porter un jugement sur la valeur artistique de ce tableau en se servant d'une échelle en 9 points où 1 = « très faible » et 9 = « très élevée ». Ce jugement constitue la variable dépendante. Puis, elle demande aux participants de verbaliser à voix haute tout ce qui leur passe par la tête quand ils regardent le tableau. Les verbalisations sont codées ultérieurement par des juges indépendants donnant un score de valence entre 0 et 100 (plus ce score est élevé, plus les concepts évoqués sont positifs). Préalablement, la chercheuse a induit chez ses participants un des trois états émotionnels, « colère », « tristesse », ou « neutre ». La chercheuse prédit que la relation entre les jugements et la valence des concepts évoqués au cours de la verbalisation à voix haute devrait être fortement positive pour les participants dans les conditions « tristesse » et « neutre », mais que cette même relation devrait être plus faible, voire inexistante, pour les participants dans la condition « colère ». Les données de cette étude hypothétique sont répertoriées dans le tableau III.

Une approche courante consiste à dichotomiser la variable « valence » et à créer deux groupes de participants, ceux qui évoquent principalement des concepts positifs et ceux qui évoquent principalement des concepts négatifs au cours de leurs verbalisations. Ensuite, on analyserait les données via une ANOVA 3×2 avec les jugements d'attitude comme variable dépendante : Émotion (« tristesse », « colère », « neutre ») × Valence (« positive », « négative »). Les moyennes correspondant à ce plan expérimental semblent confirmer l'hypothèse de la chercheuse. Alors qu'il semble y avoir une différence entre ceux qui évoquent principalement des concepts positifs et ceux qui évoquent principalement des concepts négatifs dans les conditions « tristesse » (négatif : M = 2,50 ; positif : M = 4,50) et « neutre » (négatif : M = 4,50 ; positif : M = 7,50), les jugements des participants en « colère » semblent être indépendants de la valence des concepts évoqués (négatif : M = 3,50 ; positif : M = 3,00). Dans le cadre de l'ANOVA, il conviendrait de créer deux contrastes (également montré dans le tableau III). Le premier contraste (C1) comparerait le groupe dans lequel on suppose l'occurrence de traitement heuristique

TABLEAU III. — *Les données de l'étude hypothétique sur l'influence des émotions sur la relation entre les jugements sur un objet et la valence des concepts évoqués*

The data of the hypothetical study on the influence of emotions on the relationship between judgments about an object and the valence of concepts associated with this object

Numéro de Participant	Émotion[a]	Valence des concepts	Jugements	C1	C2
1	1	10	3	-2	0
2	1	35	4	-2	0
3	1	60	4	-2	0
4	1	80	2	-2	0
5	2	15	1	+1	-1
6	2	30	4	+1	-1
7	2	55	3	+1	-1
8	2	75	6	+1	-1
9	3	25	3	+1	+1
10	3	45	6	+1	+1
11	3	60	6	+1	+1
12	3	90	9	+1	+1
Moyenne		48.33	4.25	0	0
Écart-type		25.97	2.18	--	--

[a] 1 = colère, 2 = tristesse, 3 = neutre.

aux groupes dans lesquels les participants devraient traiter les informations de façon systématique : « colère » (– 2) *versus* « tristesse » (+ 1) et « neutre » (+ 1). Le deuxième contraste (C2) comparerait les deux derniers groupes entre eux : « tristesse » (– 1) *versus* « neutre » (+ 1) alors que la condition « colère » (0) n'est pas considérée. C1 et C2 sont des contrastes orthogonaux.

Les résultats de l'ANOVA apparaissent dans la partie supérieure du tableau IV. Aucun effet n'est significatif. L'effet principal « valence » est loin du seuil de significativité,

TABLEAU IV. — ***Différentes décompositions de la variance de l'étude hypothétique sur l'influence des émotions sur la relation entre les jugements sur un objet et la valence des concepts évoqués***

Different source tables of the hypothetical study on the influence of emotions on the relationship between judgments about an object and the valence of concepts associated with this object

Source de variation	b	Somme des carrés	ddl	*F*	*P*
VALENCE EN TANT QUE VARIABLE CATÉGORIELLE (ANOVA 2 × 3)					
Modèle		31.75	5	1.86	.236
VALENCE		6.75	1	1.98	.210
EMOTION		18.50	2	2.71	.145
C1 (1 vs 2 + 3)		6.00	1	1.76	.233
C2 (2 vs 3)		12.50	1	3.66	.104
INTERACTION		6.50	2	.95	.438
VAL * C1		6.00	1	1.76	.233
VAL * C2		.50	1	.15	.715
Erreur		20.50	6		
Total		52.25	11		
VALENCE EN TANT QUE VARIABLE CONTINUE (ANALYSE DE RÉGRESSION)					
Modèle		44.68	5	1.86	.236
[Ordonnée à l'orig.]	4.15	[198.28]	--	157.08	.0001
VALFC	.05	15.40	1	12.20	.013
C1 (1 vs 2 + 3)	.46	4.98	1	3.95	.094
C2 (2 vs 3)	.81	4.98	1	3.95	.094
INTERFC1	.03	12.84	1	10.17	.019
INTERFC2	.01	.52	1	.41	.546
Erreur		7.57	6		
Total		52.25	11		

Note : VALENCE = Valence des concepts évoqués (dichotomisée) ; ÉMOTION = Émotion ressentie par les participants ; C1 = Contraste 1 ; C2 = Contraste 2 ; VAL * C1 = Interaction entre VALENCE et C1 ; VAL * C2 = Interaction entre VALENCE et C2 ; VALFC = Valence des concepts évoqués en forme centrée ; INTERFC1 = Interaction entre VALFC et C1 ; INTERFC2 = Interaction entre VALFC et C2.

$F(1,6) = 1{,}98$, $p = .21$, ce qui est surprenant. Après tout, on aurait pu s'attendre à trouver un lien entre la valence des concepts évoqués et les jugements. La théorie de la chercheuse prédisait une interaction significative entre le premier contrast (C1) et la valence mais cette interaction n'est pas non plus significative, $F(1,6) = 1{,}76$, $p = .23$. En résumé, la chercheuse va conclure que les résultats ne confirment pas son hypothèse concernant les effets des émotions sur le traitement des informations.

Quelle conclusion allons-nous tirer de ces données quand nous analysons la variable « valence » telle qu'elle est, c'est-à-dire comme une variable continue ? Comme nous le savons, la première étape est de transformer la variable indépendante continue en forme centrée. La moyenne des scores de valence étant de 48,33, la nouvelle variable « valence » (VALFC) est créée en déduisant 48,33 des scores de chacun des participants. L'étape suivante consiste à recoder la variable indépendante catégorielle de façon à ce qu'elle soit représentée par des contrastes orthogonaux et centrés. Avec une variable catégorielle à trois modalités, on a besoin de deux contrastes. Ces contrastes correspondent simplement aux contrastes C1 et C2 mentionnés ci-dessus. En d'autres termes, on crée deux nouvelles variables et l'on attribue à chacun des participants deux valeurs qui correspondent au contrastes C1 et C2. Finalement, on est obligé de créer deux variables d'interaction qui correspondent aux produits de la variable indépendante continue en forme centrée avec chacun des deux contrastes : INTERFC1 = VALFC*C1, et INTERFC2 = VALFC*C2.

En régressant le jugement d'attitude sur les cinq variables que l'on vient de créer, on obtient l'équation de régression suivante : $\text{JUGEMENT} = b_0 + b_1\text{VALFC} + b_2\text{C1} + b_3\text{C2} + b_4\text{INTERFC1} + b_5\text{INTERFC2}$. La décomposition de la variance est présentée dans la partie inférieure du tableau IV. Les conclusions de cette analyse sont différentes de celles que l'on a tirées de l'ANOVA. Ici, la valence des concepts évoqués est fortement liée aux jugements, $F(1,6) = 12{,}20$, $p = .01$. On retrouve la relation entre jugements et concepts associés qui a déjà été démontrée dans la littérature. Comme prédit par la chercheuse, cet effet est modulé par une interaction avec le premier contraste, $F(1,6) = 10{,}17$, $p = .02$. Pour les participants dans les conditions « tristesse » et « neutre », il y a une relation positive entre les

jugements et les concepts évoqués lors de la verbalisation à voix haute (voir fig. 3). En revanche, cette relation n'existe pas pour les participants dans la condition « colère ».

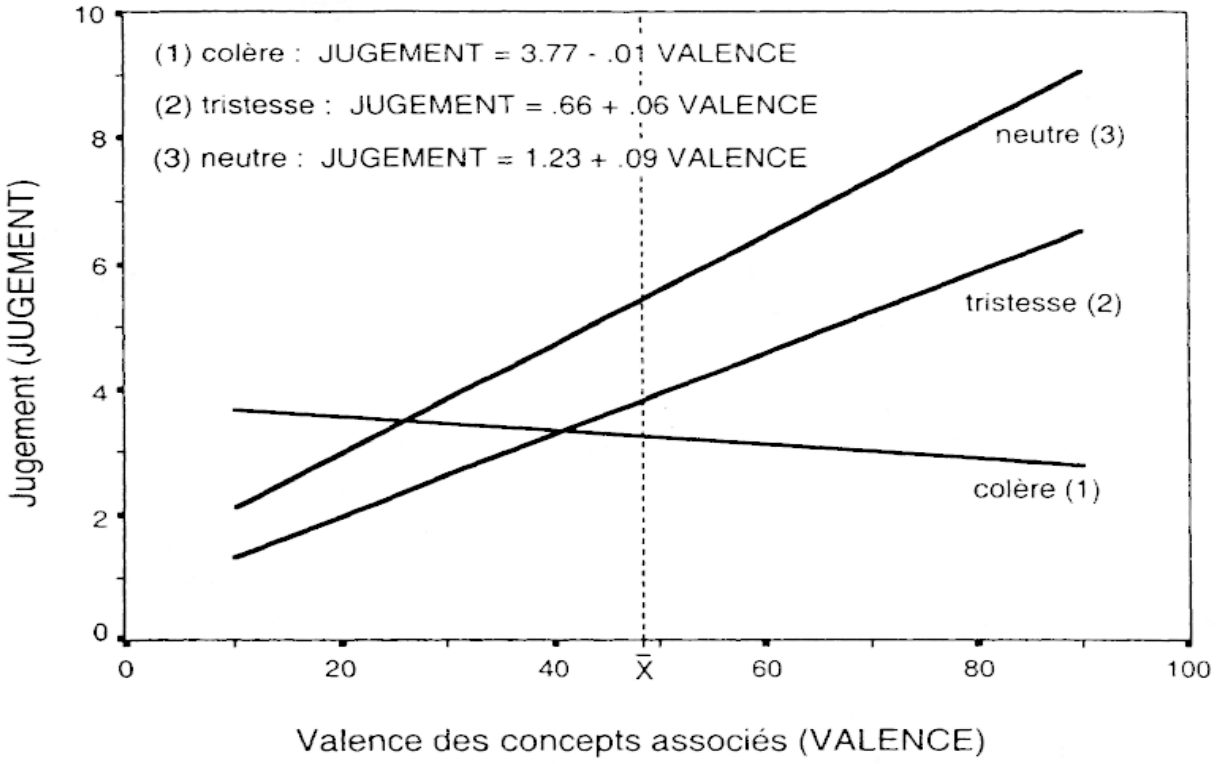


Fig. 3. — Les résultats de l'étude hypothétique sur l'influence des émotions sur la relation entre les jugements sur un objet et la valence des associations liées à cet objet. La variable indépendante continue (valence des concepts évoqués) est traitée en tant que variable continue

The data of the hypothetical study on the influence of emotions on the relationship between judgments about an object and the valence of concepts associated with this object. The continuous independent variable (the valence of the linked associations) is treated as a continuous variable

L'analyse de régression multiple est plus puissante statistiquement, mais elle présente un inconvénient. Elle ne fait pas apparaître les effets principaux des variables catégorielles à plus de deux modalités. Alors que nous avons affirmé ailleurs que les tests avec deux (ou plus) degrés de liberté sont peu informatifs et qu'il vaut mieux examiner tout de suite les contrastes qui testent directement l'hypothèse opérationnelle (Brauer et McClelland, 2002 ; voir aussi, Judd et McClelland, 1989 ; Rosenthal et Rosnow, 1985), certains rapporteurs exigent toujours que l'auteur évoque toutes les comparaisons possibles, y compris les

effets principaux des variables indépendantes à plus de deux modalités. L'analyse de régression mentionnée ci-dessus ne fournit pas ces statistiques. On est alors amené à effectuer deux analyses de régression hiérarchiques.

Pour calculer l'effet principal de la variable « émotion », il convient d'effectuer une analyse de régression hiérarchique avec deux blocs. Dans le deuxième bloc on entrera les deux contrastes C1 et C2, dans le premier bloc on entrera toutes les autres variables (VALFC, INTERFC1 et INTERFC2). La valeur F du deuxième bloc correspond alors à l'effet principal de la variable « émotion » (ici : $F(2,6) = 3{,}91$, $p = .08$). On peut procéder de façon identique pour déterminer si la relation entre les jugements et les concepts évoqués est la même dans les trois conditions expérimentales. On effectue une analyse de régression hiérarchique où l'on entre au premier bloc C1, C2 et VALFC, et au deuxième bloc INTERFC1 et INTERFC2. En examinant la valeur F du deuxième bloc, on constate que l'interaction entre « émotion » et « valence » est significative, $F(2,6) = 5{,}33$, $p = .047$.

UNE VARIABLE INDÉPENDANTE CONTINUE EN PRÉSENCE D'UNE VARIABLE INDÉPENDANTE CATÉGORIELLE INTRAPARTICIPANT

Tesser et Smith (1980) ont montré que les étudiants américains ont tendance à aider un étranger plus qu'un ami quand il s'agit de donner des indices dans un jeu de devinette qui est présenté comme mesurant l'intelligence. Imaginons un psychologue interculturel qui prédit que cette tendance devrait disparaître, ou même être inversée, pour des individus ayant un niveau élevé d'allocentrisme[1]. Le psychologue met en place une expérience pour tester son hypothèse. Après avoir mesuré le niveau d'allocentrisme de six participants (sur une échelle allant de – 4 à + 4), il leur demande de venir au laboratoire accompagné d'un ami. L'ami est conduit dans une autre salle et l'expérimentateur

1. Les individus allocentriques se caractérisent par une focalisation sur autrui, par un concept de soi interdépendant plutôt qu'indépendant, et par une attribution plus importante au succès de l'équipe plutôt qu'au succès individuel (Singelis, 1994 ; Maisonneuve, 2000).

fait croire aux participants que leur ami va jouer à un jeu mesurant l'intelligence contre une troisième personne qu'ils ne connaissent pas et qui se trouve dans une troisième salle (l'étranger). L'ami et l'étranger ne peuvent pas communiquer entre eux, mais lui, le participant, peut envoyer des indices aux deux joueurs en se servant d'une messagerie électronique.

TABLEAU V. — ***Les données des études hypothétiques sur la relation entre l'allocentrisme et les aides apportées à un ami et à un étranger au cours d'un jeu***

The data of the hypothetical studies on the relationship between allocentrism and help given to a friend and to a stranger during a game

Numéro de Participant	Allocentrisme (ALLO)	Ami (AMI)	Étranger (ETR)
1	-4	.20	.35
2	-3	.20	.30
3	-2	.55	.50
4	0	.40	.40
5	2	.70	.55
6	4	.85	.45
Moyenne part. n° 1-6	-.50	.48	.43
Écart-type part. n° 1-6	3.08	.27	.09
7	-3	.20	.00
8	-2	.25	.15
9	-2	.35	.30
10	0	.15	.15
11	1	.65	.40
12	3	.50	.45
Moyenne part. n° 1-12	-.50	.42	.33
Écart-type part. n° 1-12	2.58	.23	.16

Note : Participants 1-6 : implication forte ; participants 7-12 : implication faible ; Ami = utilité des indices envoyés à l'ami ; Étranger = utilité des indices envoyés à l'étranger.

L'utilité des indices (évaluée préalablement sur une échelle de 0.00 à 1.00) est la variable dépendante. Le psychologue se retrouve avec trois variables par participant, un score d'allocentrisme (ALLO), l'utilité des indices envoyés à l'ami (AMI), et l'utilité des indices envoyés à l'étranger (ETR). Les résultats de cette étude hypothétique apparaissent dans la partie supérieure du tableau V (participants 1-6).

L'approche traditionnelle serait de dichotomiser la variable allocentrisme et d'effectuer une ANOVA 2 × 2 avec un facteur intraparticipant (cible : « ami » *versus* « étranger ») et un facteur interparticipant (allocentrisme : « fort » *versus* « faible »). Cette analyse ne révèle aucun résultat significatif, ni l'effet principal d'allocentrisme, $F(1,4) = 3{,}09$, $p = .15$, ni l'effet principal de la cible, $F(1,4) = .79$, $p = .42$, ni l'interaction entre les deux facteurs, $F(1,4) = 3{,}63$, $p = .13$.

Pour examiner la même question en traitant l'allocentrisme en tant que variable continue, il convient de transformer la variable ALLO en forme centrée. La moyenne étant de – .50, il faut ajouter .50 à chacun des participants (ALLOFC). L'étape suivante est de créer deux nouvelles variables : la première est simplement la moyenne des deux scores d'utilité (MOYENNE = [AMI + ETR]/2), la deuxième est la différence de ces deux scores (DIFF = AMI – ETR). Ensuite, on effectuera deux analyses de régression. Dans la première on régresse MOYENNE sur ALLOFC, dans la deuxième DIFF sur ALLOFC. La décomposition de la variance de ces deux analyses de régression est montrée dans le tableau VI.

Le coefficient b_1 de la première analyse de régression examine ce que l'on a appelé l'effet principal d'allocentrisme dans l'ANOVA. Il examine s'il y a une relation significative entre le score d'allocentrisme d'un participant et l'utilité des indices qu'il a donnés à ses partenaires, indépendamment du fait qu'il s'agit de l'ami ou de l'étranger. Le résultat montre que plus un individu est allocentrique, plus il a tendance à aider ses partenaires de jeu, $F(1,4) = 12{,}27$, $p = .02$. Le coefficient b_0 de la deuxième analyse de régression (celle avec DIFF en tant que variable dépendante) examine l'effet principal de la cible. Est-ce que les participants aident davantage leur ami que l'étranger ? Les données ne semblent pas étayer cette idée, car l'effet principal de la cible n'est pas significatif, $F(1,4) = 3{,}54$, $p = .13$. Le coefficient b_1 de la même analyse de régression teste l'interaction

TABLEAU VI. — *Différentes décompositions de la variance de la première étude hypothétique sur la relation entre l'allocentrisme et les aides apportées à un ami et à un étranger (participants n° 1-6)*

Different source tables of the first hypothetical study on the relationship between allocentrism and help given to a friend and to a stranger (participants 1-6)

Source de variation	b	Somme des carrés	ddl	*F*	*P*
VARIABLE DÉPENDANTE : MOYENNE					
Modèle		.112	1	12.27	.025
[Ordonnée à l'orig.]	.45	[1.238]	--	134.93	.0003
ALLOFC	.05	.112	1	12.27	.025
Erreur		.037	4		
Total		.149	5		
VARIABLE DÉPENDANTE : DIFF					
Modèle		.174	1	30.17	.005
[Ordonnée à l'orig.]	.06	[.020]	--	3.54	.133
ALLOFC	.06	.174	1	30.17	.005
Erreur		.023	4		
Total		.197	5		

Note : MOYENNE = Utilité moyenne des indices envoyés ; DIFF = Différence d'utilité des indices envoyés à l'ami et des indices envoyés à l'étranger ; ALLOFC = Allocentrisme en forme centrée.

entre les deux facteurs, $F(1,4) = 30{,}17$, $p = .005$. b_1 est positif, ce qui signifie que plus quelqu'un est allocentrique, plus il a tendance à aider davantage son ami que l'étranger.

En régressant séparément chacune des deux variables dépendantes originales (AMI et ETR) sur le score

d'allocentrisme en forme brute (ALLO), on obtient les équations des deux droites de régression correspondant à chacune des cibles : AMI = .523 + .079 ALLO, et ETR = .434 + .018 ALLO. Le graphique montrant les deux droites de régression sera présenté ultérieurement (voir la partie supérieure de la figure 4). Contentons-nous pour l'instant de discuter les différences entre les coefficients de régression. L'allocentrisme semble fortement lié à l'aide apportée à une cible si cette cible est un ami (b_1 = .079, F(1,4) = 20,67, p = .01), tandis que l'allocentrisme semble jouer un rôle beaucoup moins important quand la cible est un étranger (b_1 = .018, F(1,4) = 2.33, p = .20). Tous les participants, indépendamment de leur niveau d'allocentrisme, envoient des indices d'une utilité intermédiaire à l'étranger. En revanche, plus ils sont allocentriques, plus ils aident leurs amis.

UNE VARIABLE INDÉPENDANTE CONTINUE EN PRÉSENCE D'UNE VARIABLE INDÉPENDANTE CATÉGORIELLE INTRAPARTICIPANT ET UNE VARIABLE INDÉPENDANTE CATÉGORIELLE INTERPARTICIPANT

Un rapporteur fait remarquer au psychologue interculturel que Tesser et Smith (1980) ont montré que les étudiants américains aident davantage un étranger quand la tâche implique le soi mais qu'ils aident davantage un ami quand ils attribuent peu d'importance à la tâche. Le psychologue interculturel est alors amené à faire une deuxième expérience dans laquelle il ajoute une manipulation expérimentale. Il dit à la moitié des participants (N = 6) que le jeu est un test d'intelligence (comme dans l'expérience précédente), et à l'autre moitié (N = 6) qu'il s'agit d'un test préalable à l'élaboration d'un jeu éducatif destiné aux enfants. Il s'attend à trouver une interaction entre cible et allocentrisme dans la condition « forte implication du soi » (*i.e.*, il s'attend à répliquer la première expérience), mais il prédit que cette interaction devrait disparaître dans la condition « faible implication du soi », c'est-à-dire dans la condition où le jeu est

présenté comme prétest d'un jeu pour enfants. Ici, il s'attend à ce que tous les participants, indépendamment de leur niveau d'allocentrisme, aident l'ami plus que l'étranger. En somme, il prédit une interaction de niveau deux entre allocentrisme, cible et implication du soi. Les données de cette étude hypothétique sont rapportées dans le tableau V (participants 1-12 ; on suppose que les participants dans la condition « forte implication du soi » donnent les mêmes réponses que ceux de l'étude précédente).

En termes d'ANOVA, il s'agit d'un plan expérimental en $2 \times 2 \times 2$: Allocentrisme (« faible » *versus* « fort », inter) × Cible (« ami » *versus* « étranger », intra) × Implication du Soi (« faible » *versus* « fort », inter). Nous savons maintenant que cette ANOVA réduit notre puissance statistique et n'est pas l'analyse appropriée. Mais, comment tester l'hypothèse du chercheur en traitant l'allocentrisme en tant que variable continue ? La façon de procéder est quasi identique à celle que nous avons suggéré dans le sous-chapitre précédent. Il convient de transformer l'allocentrisme en forme centrée (ALLOFC) et de calculer deux nouvelles variables, une correspondant à la moyenne d'AMI et d'ETR (MOYENNE) et l'autre correspondant à la différence entre ces deux variables (DIFF). On recodera l'implication du soi de façon à ce qu'elle soit codée en contrastes orthogonaux et centrés (IMPLCO). Tous les participants dont le soi est faiblement impliqué (condition « jeu d'enfants ») auront un score de – 1, tous ceux dont le soi est fortement impliqué (condition « test d'intelligence ») auront un score de + 1. On calculera également une variable d'interaction qui correspond au produit entre le score d'allocentrisme en forme centrée et l'implication du soi en contrastes orthogonaux (INTERFC = ALLOFC*IMPLCO). Finalement, on effectue deux analyses de régression, l'une avec MOYENNE, l'autre avec DIFF en tant que variables dépendantes. Les deux analyses de régression ont les mêmes variables indépendantes : ALLOFC, IMPLCO, et INTERFC. Le tableau VII présente la décomposition de la variance des deux analyses de régression.

Dans l'analyse de régression où MOYENNE est la variable dépendante, le coefficient b_1 (celui associé à ALLOFC) examine l'effet principal d'allocentrisme, le coefficient b_2 (celui associé à IMPLCO) teste l'effet principal de l'implication du soi, et le coefficient b_3 (celui associé à INTERFC) représente l'interaction

TABLEAU VII. — ***Différentes décompositions de la variance de la deuxième étude hypothétique sur la relation entre l'allocentrisme et les aides apportées à un ami et à un étranger (participants n° 1-12)***

Different source tables of the second hypothetical study on the relationship between allocentrism and help given to a friend and to a stranger (participants 1-12).

Source de variation	b	Somme des carrés	ddl	*F*	*P*
VARIABLE DÉPENDANTE : MOYENNE					
Modèle		.275	3	7.03	.012
[Ordonnée à l'orig.]	.38	[1.688]	--	129.57	.0001
ALLOFC	.05	.190	1	14.59	.005
IMPLCO	.08	.079	1	5.78	.043
INTERFC	-.00	.005	1	.12	.739
Erreur		.104	8		
Total		.379	11		
VARIABLE DÉPENDANTE : DIFF					
Modèle		.183	3	7.07	.012
[Ordonnée à l'orig.]	.08	[.083]	--	9.67	.015
ALLOFC	.03	.048	1	5.54	.046
IMPLCO	-.03	.008	1	.87	.378
INTERFC	.03	.075	1	8.74	.018
Erreur		.069	8		
Total		.252	11		

Note : MOYENNE = Utilité moyenne des indices envoyés ; DIFF = Différence d'utilité des indices envoyés à l'ami et des indices envoyés à l'étranger ; ALLOFC = Allocentrisme en forme centrée ; IMPLCO = implication du soi (codée en contrastes orthogonaux et centrés) ; INTERFC = Interaction entre ALLOFC et IMPLCO.

entre l'allocentrisme et l'implication du soi. Dans l'analyse de régression où DIFF est la variable dépendante, le coefficient b_0 est l'effet principal de la cible. Le coefficient b_1 (celui associé à ALLOFC) teste l'interaction entre la cible et l'allocentrisme, alors que le coefficient b_2 (celui associé à IMPLCO) teste l'interaction entre la cible et l'implication du soi. Enfin, le coefficient b_3 (celui associé à INTERFC) teste l'interaction de niveau deux, c'est-à-dire l'interaction entre allocentrisme, cible et implication du soi.

À nouveau, il y a une relation positive entre l'allocentrisme des participants et l'utilité des indices qu'ils envoient à leurs partenaires, $F(1,8) = 14{,}59$, $p = .005$. On trouve également un effet principal de l'implication du soi : les participants aident plus quand le jeu est présenté comme un test d'intelligence que quand il est présenté comme un jeu éducatif, $F(1,8) = 5{,}78$, $p = .04$. Dans cette deuxième étude, l'effet principal de la cible est significatif. Les gens aident davantage un ami qu'un étranger, $F(1,8) = 9{,}67$, $p = .01$. Finalement, ces effets principaux sont modulés par deux interactions. La première, entre la cible et l'allocentrisme, indique que la tendance à aider un ami plus qu'un étranger augmente avec l'allocentrisme, $F(1,8) = 5{,}54$, $p < .05$. La deuxième interaction significative implique l'ensemble des trois variables indépendantes, $F(1,8) = 8{,}74$, $p = .02$. C'est cette interaction de niveau deux que le psychologue interculturel avait prédite.

Pour examiner l'interaction impliquant les trois variables indépendantes de plus près, on pourrait calculer l'interaction simple entre l'allocentrisme et la cible pour chacune des conditions de l'implication de soi séparément. En d'autres termes, on pourrait effectuer l'analyse de régression $DIFF = b_0 + b_1 ALLOFC$ à deux reprises, une fois avec les participants qui pensent que le jeu est un test d'intelligence et une fois avec les participants auxquels on a dit qu'il s'agissait d'un test préalable à l'élaboration d'un jeu éducatif destiné aux enfants. En faisant cela, on se rend compte que b_1 (l'interaction entre l'allocentrisme et la cible) est significative pour les participants à forte implication de soi, $F(1,4) = 30{,}17$, $p = .005$, mais qu'elle ne l'est pas pour les participants à faible implication de soi, $F(1,4) = .10$, $p = .76$.

Pour tester si les différentes interactions sont bien dans la « bonne » direction et pour représenter les résultats graphique-

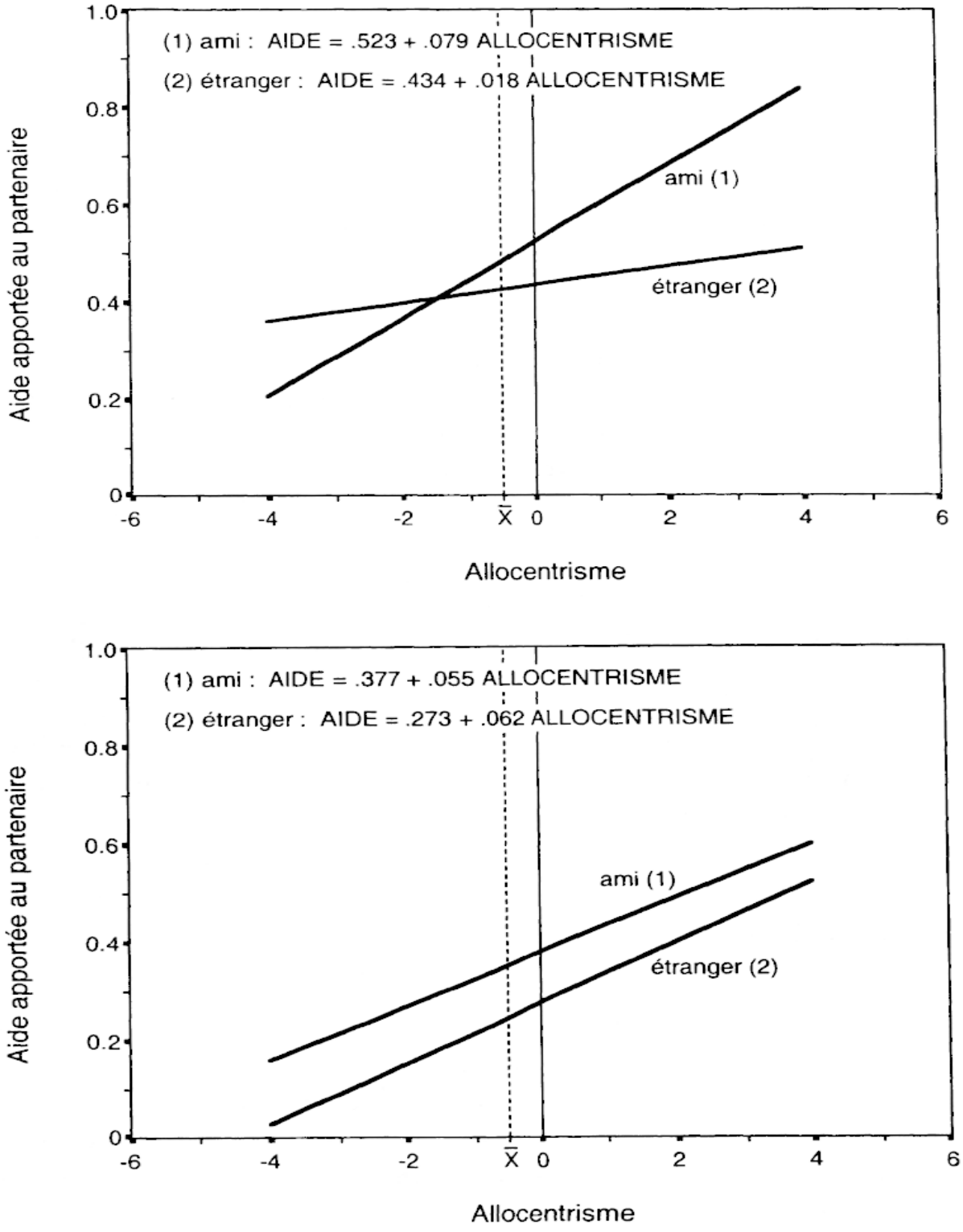


Fig. 4. — Les résultats de l'étude hypothétique sur la relation entre l'allocentrisme et la tendance d'aider un étranger davantage qu'un ami quand les partenaires jouent à un jeu qui est présenté soit comme un test d'intelligence (partie supérieure) soit comme un prétest pour un jeu d'enfants (partie inférieure)

The results of the hypothetical study on the relationship between allocentrism and the tendency to help a stranger more than a friend when the partners play a game that is presented either as an intelligence test (top panel) or as a pretest pour a game for children (bottom panel)

ment, il convient de produire les droites de régression pour chacune des quatre conditions expérimentales créées par le croisement des deux facteurs cible et implication de soi. Les quatre droites de régression apparaissent dans la figure 4. La partie supérieure représente les résultats pour les participants dans la condition forte implication du soi (test d'intelligence). Cette partie supérieure représente également les résultats de la première étude du psychologue interculturel car les données des participants 1 à 6 ont été reprises de façon identique. La partie inférieure de la figure 4 représente les résultats pour les participants dans la condition faible implication du soi (prétest pour un jeu d'enfant). En comparant les deux parties de la figure 4, on voit bien que l'hypothèse du psychologue interculturel est valide. Quand l'implication du soi est faible (voir partie inférieure), tous les participants aident davantage l'ami que l'étranger. En revanche, quand le soi est fortement impliqué (voir partie supérieure) l'aide apportée à l'ami dépend du niveau d'allocentrisme. Les individus fortement allocentriques l'aident beaucoup (plus que l'étranger) alors que les individus faiblement allocentriques envoient à leur ami des indices n'ayant qu'une faible utilité (en tous les cas plus faible que celle des indices envoyés à l'étranger).

CONCLUSION

Le but de cet article était de démontrer les inconvénients de la dichotomisation de variables indépendantes continues et de suggérer des stratégies statistiques permettant d'analyser des plans expérimentaux contenant des variables indépendantes continues et catégorielles. Nous avons vu qu'en dichotomisant une variable continue, on ajoute artificiellement de l'erreur aléatoire. Il est vrai que ce surplus en erreur aléatoire peut être compensé en utilisant un plus grand nombre de participants. Par conséquent, deux possibilités s'offrent à un chercheur qui vient d'interroger 100 participants et qui a une variable indépendante continue dans son plan expérimental : soit il traite la variable continue en tant que telle et effectue les analyses présentées dans cet article, soit il interroge 62 participants supplémentaires (62 étant 38 % de 162), dichotomise sa variable continue et

effectue une ANOVA. Supposant une distribution normale, ces deux stratégies ont une puissance statistique strictement identique.

Nous avons également abordé différents plans expérimentaux comprenant une variable continue et une (ou plusieurs) variable(s) catégorielle(s). Plus précisément, nous avons considéré les cas où la variable continue était en présence : *a)* d'une variable catégorielle interparticipant à deux modalités ; *b)* d'une variable catégorielle interparticipant à trois modalités ; *c)* d'une variable catégorielle intraparticipant ; *d)* d'une variable catégorielle intraparticipant et d'une variable catégorielle interparticipant. Dans chacun des cas, la première étape consistait à transformer la variable indépendante continue en forme centrée, recoder la (ou les) variable(s) indépendante(s) catégorielle(s) en contrastes orthogonaux et centrés, et créer la (ou les) variable(s) d'interaction en multipliant la première avec la (ou les) dernière(s).

Notons qu'il n'y a aucun cas où la dichotomisation est préférable aux stratégies analytiques présentées dans cette article. L'ANOVA étant un cas particulier de l'analyse de régression, toutes contraintes de celle-ci s'appliquent également à celle-là. Dans le modèle linéaire général (qui s'applique aussi bien à la régression qu'à l'ANOVA)

$$Y = b_0 + b_1X1 + b_2X2 + \ldots + e$$

il n'y a qu'une seule supposition distributionnelle : les résidus (e) doivent être distribués de façon normale[1]. Aucune contrainte n'existe quant à la distribution normale des variables indépendantes X1, X2, etc. En d'autres mots, ce serait incorrect de dire que la dichotomisation est préférable si la variable indépendante continue possède une distribution non normale, par exemple une distribution bimodale ou uniforme. La variable indépendante continue peut être traitée en tant que telle indépendamment de la manière dont elle est distribuée. La quantité d'erreur aléatoire artificiellement ajoutée par la dichotomisation varie selon la

1. Pour être exact, deux suppositions distributionnelles existent : 1 / les résidus doivent avoir une distribution normale ; et 2 / les Y doivent avoir une distribution normale à chaque niveau de X1, X2, etc. (*i.e.*, pour chaque modalité de X1, X2, etc., il doit y avoir une distribution normale des valeurs autour de la droite de régression). Malheureusement, nous avons rarement assez d'observations pour pouvoir tester cette deuxième supposition.

forme de la distribution mais il y a une perte de la puissance statistique dans tous les cas de figure. À part la simplicité des analyses inférentielles, aucun argument statistique n'existe en faveur de la dichotomisation.

La dichotomisation réduit également la puissance statistique à détecter un effet expérimental quand on a affaire à des groupes extrêmes. Certes, cette réduction est moins importante que celle d'une dichotomisation « classique » où les participants n'ont pas été présélectionnés. Mais le chercheur diminue néanmoins ses chances de détecter un effet expérimental en dichotomisant sa variable indépendante continue. Prenons l'exemple d'un chercheur qui veut examiner les effets de l'âge sur une variable dépendante quelconque. Il interroge deux groupes extrêmes, l'un avec des participants âgés de 15 à 30 ans, l'autre avec des participants âgés de 60 à 80 ans. Le chercheur a tout intérêt à saisir l'âge exact de ses participants, en transformant sa variable « âge » en forme centrée et à la traiter en tant que variable continue.

Il est également possible d'effectuer les analyses suggérées dans cet article lorsque l'on souhaite examiner les effets d'une variable indépendante résultant d'une combinaison de plusieurs indicateurs différents. Imaginons, par exemple, une étude où un chercheur teste les effets du degré d'extraversion sur une variable dépendante. Après avoir mesuré le degré d'extraversion avec trois indicateurs différents, le chercheur a calculé un score global d'extraversion en faisant la moyenne des trois indicateurs standardisés. Dans un tel cas, le score composite peut être analysé tel quel, sans être dichotomisé et sans subir de transformation en forme centrée (la moyenne de ce score étant, par définition, égale à zéro).

Jusqu'ici, notre argument principal en faveur des analyses suggérées était l'efficacité : il est somme toute plus rapide de créer trois nouvelles variables que d'interroger des participants supplémentaires. Mais il existe un argument plus fondamental en défaveur des analyses dichotomiques. Si dans un plan comprenant seulement une variable indépendante continue, la dichotomisation n'a pas d'autre incidence que de réduire la taille des effets et, par conséquent, les chances de trouver des effets significatifs, la situation est radicalement différente dans un plan mettent en jeu plusieurs variables indépendantes continues. Maxwell et Delaney (1993) ont, en effet, démontré que les

résultats peuvent être biaisés quand on teste l'interaction entre deux variables continues qui ont été dichotomisées. Même une déviation légère de la distribution normale sur une des deux variables continues entraîne des estimations faussées. Avec deux variables continues, l'ANOVA pose donc un problème bien plus grave que celui d'une simple perte de puissance statistique, elle conduit à des résultats erronés !

Enfin, au-delà des considérations purement statistiques, il est probable que la pratique qui consiste à dichotomiser nos variables continues affecte notre réflexion théorique sur les concepts et les phénomènes que nous étudions. La littérature regorge d'exemples de catégories considérées comme qualitativement différentes et mutuellement exclusives, alors qu'elles résultent, en réalité, de la simple dichotomisation d'une variable continue : les individus avec une personnalité de type A *versus* ceux avec une personnalité de type B (Friedman et Rosenman, 1959), les individus à faible *versus* à forte estime de soi (Brown et Smart, 1991), les bons élèves *versus* les élèves faibles (Brunot, 1997), les individus à faibles *versus* à forts préjugés (Devine, 1989), les « incrémentalistes » *versus* les « entitativistes » (Dweck et Leggett, 1988), ceux avec un locus de contrôle interne *versus* ceux avec un locus de contrôle externe (Rotter, 1966 ; Desrumeaux, 1996). Lorsque la caractéristique en question suit une distribution normale (ce qui est le cas le plus fréquent), la plupart des individus se trouvent au centre de la distribution et ne sont donc pas clairement attribuables à une des deux catégories. À tel point, que l'on se retrouve souvent face à l'aberration suivante : un même individu peut se trouver parfois en dessus de la médiane et parfois en dessous, en fonction des caractéristiques des autres participants dans l'échantillon. Par conséquent, ce qui est présenté et considéré comme fondamentalement différent (*i.e.*, une différence qualitative) n'est en fait qu'une question de degré (*i.e.*, une différence quantitative).

Nous avançons l'idée selon laquelle nos modèles théoriques et les questions théoriques posées dans nos recherches sont en partie influencés par la façon dont nous nous représentons les concepts psychologiques sur lesquels nous travaillons. Un chercheur qui considère deux catégories comme qualitativement distinctes se centrera probablement sur les aspects qui différencient les deux groupes d'individus correspondants. Il étudiera, par exemple, les différences de traits de caractère, les différences

dans la façon d'aborder l'environnement social, dans les croyances, les attitudes, ou encore dans la façon de traiter les informations de l'environnement. En revanche, un chercheur qui considère une différence comme quantitative s'intéressera plutôt aux points communs des individus qui se situent à des niveaux différents sur la variable indépendante continue. Il étudiera, dans ce cas, les facteurs qui conduisent un individu à se déplacer sur le continuum dans une direction ou dans une autre. Prenons l'exemple de la réussite scolaire. Un chercheur pensant en termes de « bons élèves » et « élèves faibles » n'a-t-il pas tendance à examiner les différences stables qui existent entre les deux groupes ? Pour lui, une des questions les plus pertinentes est d'étudier ce qui distingue un bon élève d'un élève faible. En revanche, un chercheur qui se représente la réussite scolaire en tant que continuum n'aura-t-il pas tendance à examiner les variables qui permettent à un élève d'augmenter sa réussite scolaire ? Pour lui, la question centrale est alors comment un élève faible peut améliorer ses performances scolaires.

En résumé, nous avons avancé plusieurs raisons susceptibles d'expliquer pourquoi il est préférable de ne pas dichotomiser des variables indépendantes continues. D'abord, c'est une procédure peu efficace car elle introduit artificiellement de l'erreur aléatoire. Ensuite, elle entraîne des estimations biaisées si le plan expérimental contient plusieurs variables continues. Finalement, il y a de fortes chances que conceptualiser une différence quantitative en différence qualitative limite les types de questions théoriques que l'on se pose. Somme toute, il n'existe aucun argument en faveur de la pratique couramment admise qui consiste à dichotomiser nos variables indépendantes continues avant de les soumettre à des analyses inférentielles.

RÉSUMÉ

Le fait de transformer une variable indépendante continue en une variable catégorielle à deux modalités rend les analyses statistiques plus faciles car après dichotomisation, les effets de cette variable peuvent être examinés grâce à une analyse de variance (ANOVA) plutôt qu'une analyse de régression multiple. Or, cette facilité a un prix élevé. En dichotomisant une variable continue, on introduit artificiellement de l'erreur aléatoire ce qui diminue la puissance statistique des analyses inférentielles. Supposant une distribution normale, la diminution de la puissance statistique correspond à l'exclusion d'environ 38 %

des participants. Dans cet article, nous présentons les problèmes associés à la dichotomisation des variables continues et nous discutons différentes stratégies statistiques permettant d'analyser des plans expérimentaux contenant des variables indépendantes continues et catégorielles.

Mots-clés : variables continues, variables quantitatives, analyse de régression, forme centrée, dichotomisation.

BIBLIOGRAPHIE

Bargh J., Chartrand T. — (2000) The measurement of automatic and controlled processes, *in* H. Reis et C. M. Judd (Edit.), *Handbook of research methods in social psychology*, New York (NY), Cambridge University Press, 253-285.

Bodenhausen G. V., Sheppard L. A., Kramer G. P. — (1994) Negative affect and social judgement : The differential impact of anger and sadness, *European Journal of Social Psychology, 24*, 45-62.

Brauer M. — (1998) Stroop interference in bilinguals : The role of similarity between the two languages, *in* A. F. Healy et L. E. Bourne (Edit.), *Foreign language learning : Psycholinguistic studies on training and retention*, Hillsdale (NJ), Erlbaum, 317-337.

Brauer M., McClelland G. H. — (2002) *L'utilisation des contrastes dans l'analyse des données : Comment tester des hypothèses spécifiques dans la recherche en psychologie ?*, manuscrit non publié, Université de Clermont-Ferrand.

Brown J. D., Smart S. A. — (1991) The self and social conduct : Linking self-representations to prosocial behavior, *Journal of Personality and Social Psychology, 60*, 368-375.

Brunot S. — (1997) *Mémoire autobiographique et expériences de réussite et d'échec. Le rôle du contenu de l'histoire personnelle dans l'organisation et la récupération des souvenirs*, thèse de Doctorat nouveau régime, Université de Clermont-Ferrand.

Cohen J. — (1983) The cost of dichotomization, *Applied Psychological Measurement, 7*, 249-253.

Cohen J. — (1990) Things I have learned (so far), *American Psychologist, 45*, 1304-1312.

Cohen J., Cohen P. — (1983) *Applied multiple regression/correlation analysis for the behavioral sciences*, Hillsdale (NJ), Erlbaum.

Desrumaux P. — (1996) Hypothèse de dépendance entre processus d'explications causales internes et externes et l'engagement pro-attitudinal : une application au don du sang, *Revue internationale de Psychologie sociale, 9*, 77-94.

Devine P. G. — (1989) Stereotypes and prejudice : Their automatic and controlled components, *Journal of Personality and Social Psychology, 56*, 5-18.

Dweck C. S., Leggett E. L. — (1988) A social-cognitive approach to motivation and personality, *Psychological Review, 25*, 109-116.

Fazio R. H. — (1990) A practical guide to the use of response latency in social psychological research, *in* C. Hendrick et M. S. Clark (Edit.), *Review of Personality and Social Psychology*, vol. 11, Newbury Park (CA), Sage, 74-97.

Friedman M., Rosenman R. H. — (1959) Association of specific overt behavior pattern with blood and cardiovascular findings : Blood cholesterol level, blood clotting time, incidence of arcus senilis and clinical coronary artery disease, *Journal of the American Medical Association, 169*, 1286-1296.

Humphreys L. G. — (1978) Research on individual differences requires correlational analysis, not Anova, *Intelligence, 2,* 1-5.

Humphreys L. G., Fleishman, A. — (1974) Pseudo-orthogonal and other analysis of variance designs involving individual-differences variables, *Journal of Educational Psychology, 66,* 464-472.

Irwin, J., McClelland G. H. — (2002) *Consequences of Dichotomizing Continuous Variables,* Manuscrit non publié, University of Colorado at Boulder.

Judd C. M., McClelland G. H. — (1989) *Data analysis : A model-comparison approach,* San Diego (CA), Harcourt Brace Jovanovich.

Mägiste E. — (1984) Stroop tasks and Dichotic translation : The development of interference patterns in bilinguals, *Journal of Experimental Psychology : Learning, Memory, and Cognition, 10,* 304-315.

Maisonneuve C. — (2000) *Allocentrisme-Idiocentrisme : d'une perspective différentialiste à une perspective psychosociale,* thèse de doctorat nouveau régime, Université de Clermont-Ferrand 2.

Maxwell S. E., Delaney H. D. — (1993) Bivariate median splits and spurious statistical significance, *Psychological Bulletin, 113,* 181-190.

Peters C. C., Van Voorhis W. R. — (1940) *Statistical procedures and their mathematical bases.* New York (NY), McGraw-Hill.

Rosenthal R., Rosnow R. L. — (1985) *Contrast analysis : Focused comparisons in the analysis of variance,* Cambridge (UK), Cambridge University Press.

Rotter J. B. — (1966) Generalized expectancies for internal versus external control of reinforcement, *Psychological Monographs : General and Applied, 80,* 1, whole n° 609, 1-28.

Singelis T. M. — (1994) The measurement of independent and interdependent self-construals, *Personality and Social Psychology Bulletin, 20,* 580-591.

Stroop J. R. — (1935) Studies of interference in serial verbal reactions, *Journal of Experimental Psychology, 18,* 643-662.

Tesser A., Smith J. — (1980) Some effects of task relevance and friendship on helping : You don't always help the one you like, *Journal of Experimental Social Psychology, 16,* 582-590.

(Accepté le 21 janvier 2002.)