5. Les modèles de base de méta-analyses (Raudenbush et Bryk, 1985)

Deux modèles de base ont été proposés. Le premier est le modèle dit fixe qui considère que la taille de l'effet est commune à la population de toutes les recherches retenues. Le second modèle, généralement plus réaliste, considère que la taille de l'effet observée se distribue selon une distribution de probabilité et que sa variance doit être prise en compte dans les calculs. On se retrouve alors dans le contexte d'une analyse à niveaux multiples où les observations au niveau 1 ne sont malheureusement pas disponibles : seulement celles aux niveaux 2 ou plus sont disponibles.

5.1 Modèle fixe

Dans le modèle fixe, comme dans la théorie classique des tests, on considère que la valeur de la taille de l'effet observée (d_j) dans l'étude j est égale à la valeur de la taille de l'effet réelle latente (δ) commune à toutes les études plus de l'erreur (e_i) spécifique à chaque étude. La taille de l'effet observée d_j se distribue selon une loi de probabilité $N(\delta, \sigma_e^2)$ où σ_e^2 est la variance de l'erreur.

$$d_j = \delta + e_j \qquad (10)$$

Pour plusieurs ce modèle est irréaliste, car les recherches retenues ne sont qu'un échantillon de toutes les recherches possibles. Il faudrait plutôt considérer que la taille de l'effet réelle est une valeur aléatoire tirée d'un échantillon de j recherches. Un modèle à coefficients aléatoires devient donc plus approprié (Hox, 2002, chapitre 8; Hox et de Leeuw, 2003; Raudenbush et Bryk, 2002, chapitre 7).

Tiré de :

Raîche, G. (2017). Méta-analyse, méta-régression et méga méta-analyse. Dans G. Raîche (dir.), *Analyse de donnes quantitative en éducation* (notes de cours) (chapitre 14). Montréal, Québec : Université du Québec à Montréal.

5.2 Modèle à coefficients aléatoires (multi niveaux)

Nous allons maintenant présenter la formulation à niveaux multiples proposée par Hox ainsi que par Raudenbush et Bryk. Toutefois, la représentation que ceux-ci utilisent porte à confusion et nous allons aborder la structuration à niveaux multiples en distinguant mieux les indices et variables utilisés.

Dans ce modèle on considère que la valeur prédite de la taille de l'effet pour chaque individu *i* ayant participé aux *j* études est égale à $\hat{d}_{ij}$. Cette prédiction est tributaire de la taille de l'effet réelle latente $\bar{\delta}$ (soit la moyenne des tailles de l'effet réelles prédites $\hat{\delta}_j$) qui est maintenant considérée comme une autre variable expliquée par une distribution de probabilité qui lui est propre.

La représentation selon le modèle multi niveaux est alors la suivante :

$$\hat{d}_{ij} = \hat{\delta}_j + e_{ij} \tag{11}$$

$$\hat{\delta}_j = b_0 + u_j \tag{12}$$

Ce qui donne la représentation en modèle mixte constituée d'un coefficient fixe b_0 et de deux coefficients aléatoires σ_e^2 et σ_u^2 , ces derniers coefficients étant associés aux deux erreurs u_j et e_{ij}.

$$\hat{d}_{ij} = b_0 + u_j + e_{ij} \tag{13}$$

L'erreur de mesure spécifique à la taille de l'effet observée d_{ij} se distribue selon une loi de probabilité $N(0, \sigma_e^2)$, tandis que l'erreur de mesure des coefficients associés à la taille de l'effet réelle pour chacune des études $\hat{\delta}_j$ se distribue selon une distribution de probabilité $N(\,0, \sigma_u^2)$. La variance totale devient ainsi la somme de ces deux sources de variance.

Tiré de :

Raîche, G. (2017). Méta-analyse, méta-régression et méga méta-analyse. Dans G. Raîche (dir.), *Analyse de donnes quantitative en éducation* (notes de cours) (chapitre 14). Montréal, Québec : Université du Québec à Montréal.

5.4 Méta-régression

Il est possible d'aller un peu plus loin qu'une méta-analyse simple en tentant d'expliquer la taille de l'effet observée par une ou des variables indépendantes λ identifiables dans chacune des j études retenues. On a ainsi proposé une modélisation qui n'est qu'une adaptation de la régression linéaire à niveaux multiples.

La représentation selon le modèle multi niveaux est la suivante :

$$\hat{d}_{ij} = \hat{\delta}_j + e_{ij} \tag{14}$$

$$\hat{\delta}_j = b_0 + b_1\lambda_j + u_j \tag{15}$$

Ce qui donne pour la représentation en modèle mixte :

$$\hat{d}_{ij} = b_0 + b_1\lambda_j + u_j + e_{ij} \tag{16}$$

Comme pour le cas précédent, deux sources de variance sont encore à considérer σ_e^2 et σ_u^2.

On peut, bien sûr, penser aussi à une modélisation où on aurait plus que deux niveaux. Ce serait le cas, par exemple, on distinguait les études selon l'université à l'intérieur desquelles elles avaient été réalisées. On pourrait ainsi utiliser des variables d'établissement universitaire pour tenter d'expliquer la variation de la taille de l'effet d'une étude à une autre.

Tiré de :

Raîche, G. (2017). Méta-analyse, méta-régression et méga méta-analyse. Dans G. Raîche (dir.), *Analyse de donnes quantitative en éducation* (notes de cours) (chapitre 14). Montréal, Québec : Université du Québec à Montréal.