

# Méthodes psychométriques en qualité de vie

Christophe Lalanne

EA 7334 REMES

Unité de Méthodologie des critères d'évaluation

Université Paris-Diderot, Sorbonne Paris-Cité



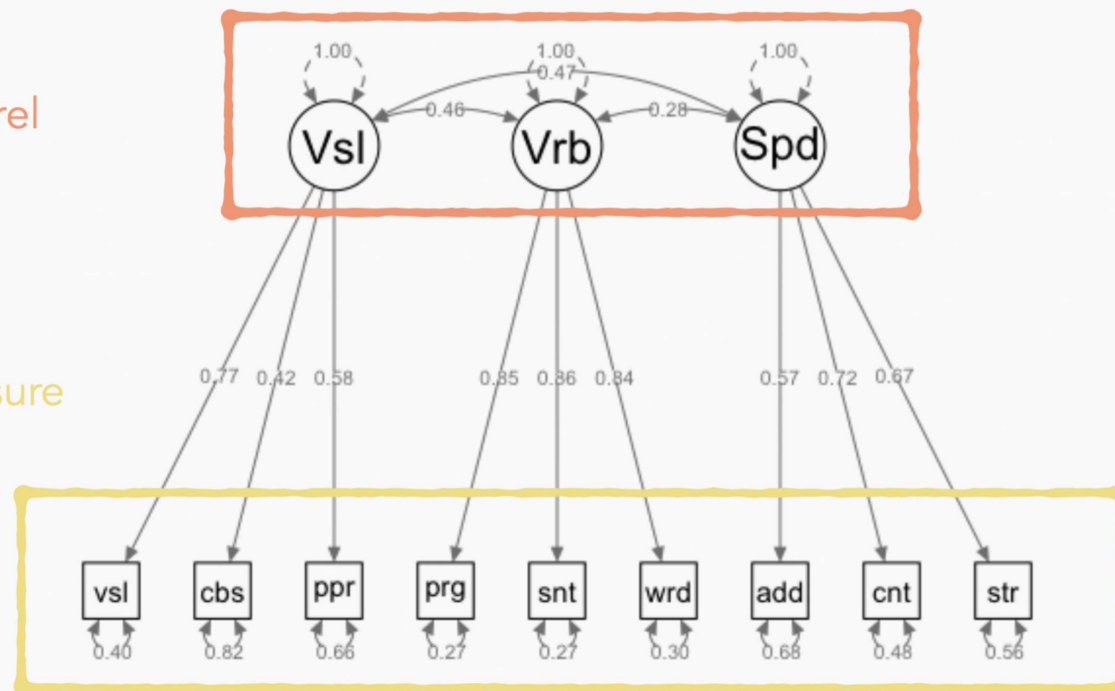
# **Analyses factorielles : extensions**

- Indices de modification
- Approche multi-groupes
- Invariance de mesure
- Cas des données catégorielles

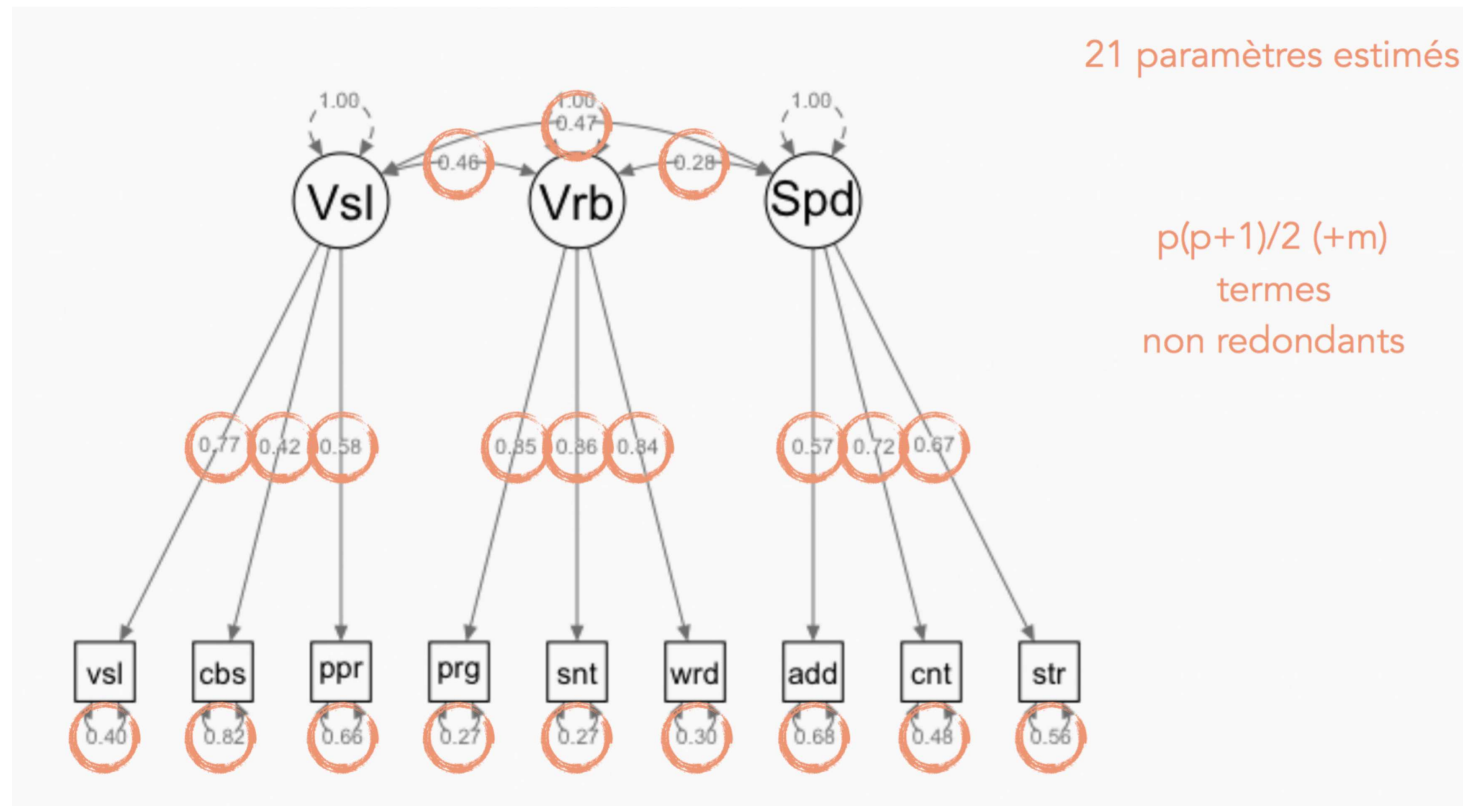
# Modèle de mesure et modèle structurel

Modèle structurel

Modèle de mesure

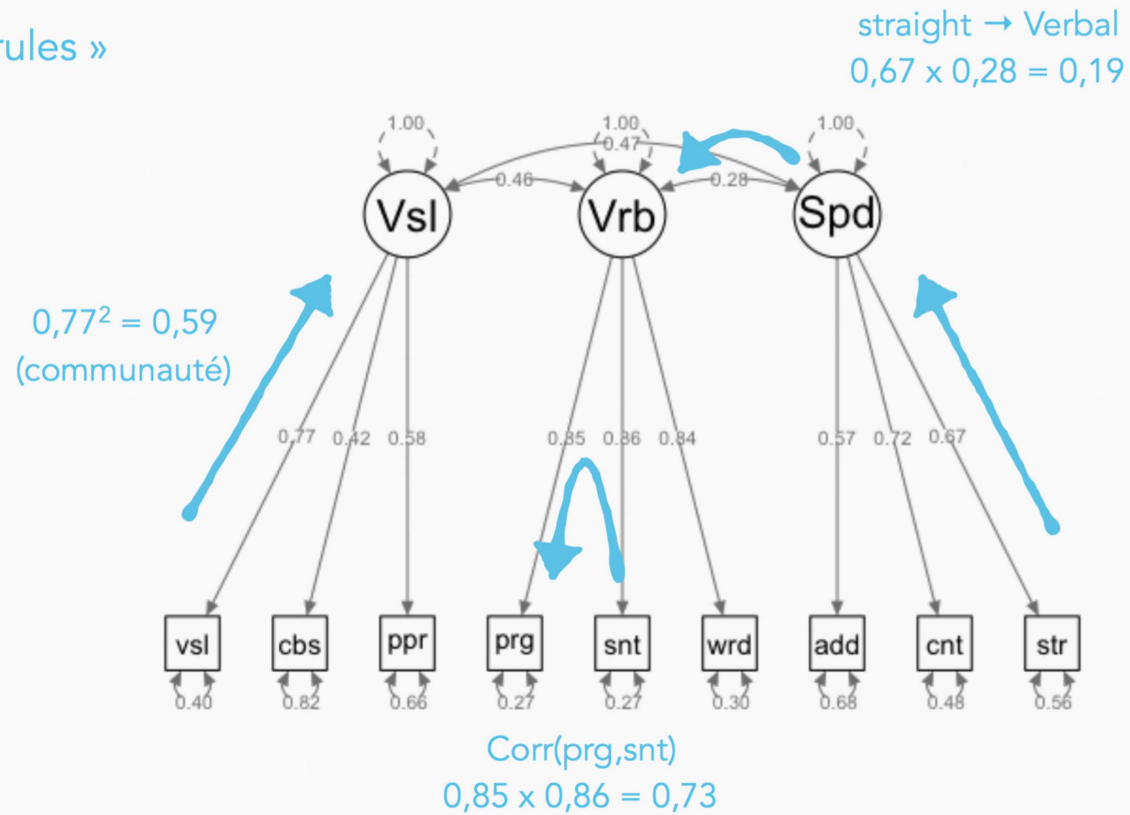


# Identification du modèle



# Interprétation des coefficients

« tracing rules »



# Syntaxe lavaan

Notre modèle réflexif de base (3 facteurs corrélés) s'écrit :

```
m <- 'Visual =~ visual + cubes + paper
      Verbal =~ paragrap + sentence + wordm
      Speed  =~ addition + counting + straight'
```

- =~ régression MV - LV
- ~~ covariance LV-LV, corrélation MV - MV
- ~ régression LV-LV
- == contraintes (identification ou autre)
- := construction d'un nouveau terme de modèle

## Modèles de covariance

L'avantage des modèles CFA est de permettre de quantifier l'écart entre le modèle et les données, de tester si cet écart est significativement différent de ce qu'on attendrait sous  $H_0$ , et de comparer des modèles emboîtés.

```
C <- cov(d)
N <- nrow(d)
r <- cfa(m, std.lv = TRUE, sample.cov = cov(d), sample.nobs = N)
resid(r)
```



La matrice de covariance utilisée par lavaan est calculée avec un dénominateur à  $N$  et non  $N - 1$  (méthode ML).

## Indices de modification

Il est parfois nécessaire de vérifier s'il n'existe pas des déviations locales au modèle postulé (en particulier au niveau des résidus), en particulier ajouter des paramètres susceptibles d'améliorer la qualité globale d'ajustement du modèle.<sup>1</sup>

```
r <- cfa(m, data = d)
standardizedSolution(r)
resid(r)           ❶
modindices(r)      ❷
```

---

1. T BROWN. *Confirmatory factor analysis for applied research*. Guildford Press, New York, 2006.



```
m1 <- 'Visual =~ visual + cubes + paper
      Verbal =~ paragraf + sentence + wordm
      Speed  =~ addition + counting + straight'

m2 <- 'Visual =~ visual + cubes + paper
      Verbal =~ paragraf + sentence + wordm
      Speed  =~ addition + counting + straight
      addition ~~ counting' ❶

r1 <- cfa(m1, data = d)
r2 <- cfa(m2, data = d)
anova(r1, r2) ❷
```

## Analyse stratifiée sur plusieurs groupes

**Première approche :** estimer séparément les modèles dans les deux groupes définis par la variable `school`.

```
r3 <- cfa(m1, data = HS, group = "school")  
r3
```

**Deuxième approche :** contraindre les charges factorielles à être égales dans les deux groupes.

```
r4 <- cfa(m1, data = HS, group = "school",  
          group.equal = "loadings")  
r4
```

# Invariance de mesure

Pour garantir des comparaisons de groupes valides, il faut au préalable s'assurer que l'on mesure bien la même chose de la même manière.<sup>2</sup>

1. **Invariance configurale** : structure factorielle identique entre groupes
2. **Invariance faible** : égalité des charges factorielles entre groupes
3. **Invariance forte** : égalité des moyennes entre groupes
4. **Invariance stricte** : égalité des erreurs entre groupes
5. Invariance stricte + égalité des variances LV ; Invariance stricte + égalité des variances LV + égalité des moyennes LV ; Invariance partielle : conditions non vérifiées pour l'ensemble des MV

---

2. J. TERESI. « Different Approaches to Differential Item Functioning in Health Applications : Advantages, Disadvantages and Some Neglected Topics ». In : *Medical Care* 44 (2006), S152–S170.

# Invariance et comparaison de groupes<sup>3</sup>

	Type of Invariance	Constraints	Between-Groups Comparisons Allowed
1	Configural	Same model. No parameter constraints.	None
2	Weak	1 + all loadings constrained to be equal between groups (but can vary within a group). Latent (co)variances allowed to vary between groups.	Latent (co)variances [weak evidence]
3	Strong	2 + all intercepts are constrained to be equal between groups (but can vary within a group). Latent means allowed to vary between groups.	Latent means, latent (co)variances [strong evidence]
4	Strict	3 + error variances are constrained to be the same between groups (but can vary within a group).	

Models with larger numbers are nested within the models with smaller numbers.

---

3. AA BEAUJEAN. *Latent Variable Modeling Using R, A Step-by-Step Guide*. New York : Routledge, 2014.

# Le package semTools

R: Measurement Invariance Tests for Categorical Items ▾ Find in Topic

Theta parameterization is used to represent SEM for categorical items. That is, residual variances are modeled instead of the total variance of underlying normal variate for each item. Five models can be tested based on different constraints across groups.

1. Model 1: configural invariance. The same factor structure is imposed on all groups.
2. Model 2: weak invariance. The factor loadings are constrained to be equal across groups.
3. Model 3: strong invariance. The factor loadings and thresholds are constrained to be equal across groups.
4. Model 4: strict invariance. The factor loadings, thresholds and residual variances are constrained to be equal across groups. For categorical variables, all residual variances are fixed as 1.
5. Model 5: The factor loadings, thresholds, residual variances and means are constrained to be equal across groups.

However, if all items have two items (dichotomous), scalar invariance and weak invariance cannot be separated because thresholds need to be equal across groups for scale identification. Users can specify `strict` option to include the strict invariance model for the invariance testing. See the further details of scale identification and different parameterization in Millsap and Yun-Tein (2004).

## Illustration

```
library(semTools)
measurementInvariance(model = m, data = d,
                      group = "school") ❶
measurementInvariance(model = m, data = d,
                      group = "school", strict = TRUE) ❷
```

La comparaison des modèles emboîtés 1 à 4 (❶) est suffisante quand on s'intéresse au modèle de mesure et à la comparaison de scores d'un questionnaire.

## Dimensions *versus* catégories

L'analyse factorielle suppose que les variables manifestes et latentes sont continues. En pratique, en psychiatrie ou en qualité de vie, les mesures auto-rapportées reposent sur des items dichotomiques (oui/non) ou polytomiques (type Likert).

Cas des données catégorielles : modèle de type IRT ou FA sur matrice de corrélation appropriée (IFA, package psych), choix d'un estimateur approprié (WLSM(V)/DWLS et paramétrisation theta<sup>4</sup>).

Le package lavaan permet de traiter le cas des variables catégorielles et fournit des modèles à seuil.<sup>5</sup>

---

4. B.O. MUTHÉN. « Goodness of fit with categorical and other nonnormal variables ». In : *Testing structural equation Models*. Sous la dir. de K.A. BOLLEN et J.S. LONG. Newbury Park, CA : Sage, 1993, p. 205–234.

5. BEAUJEAN, *Latent Variable Modeling Using R, A Step-by-Step Guide*, chap. 6.

## Exemple<sup>6</sup>

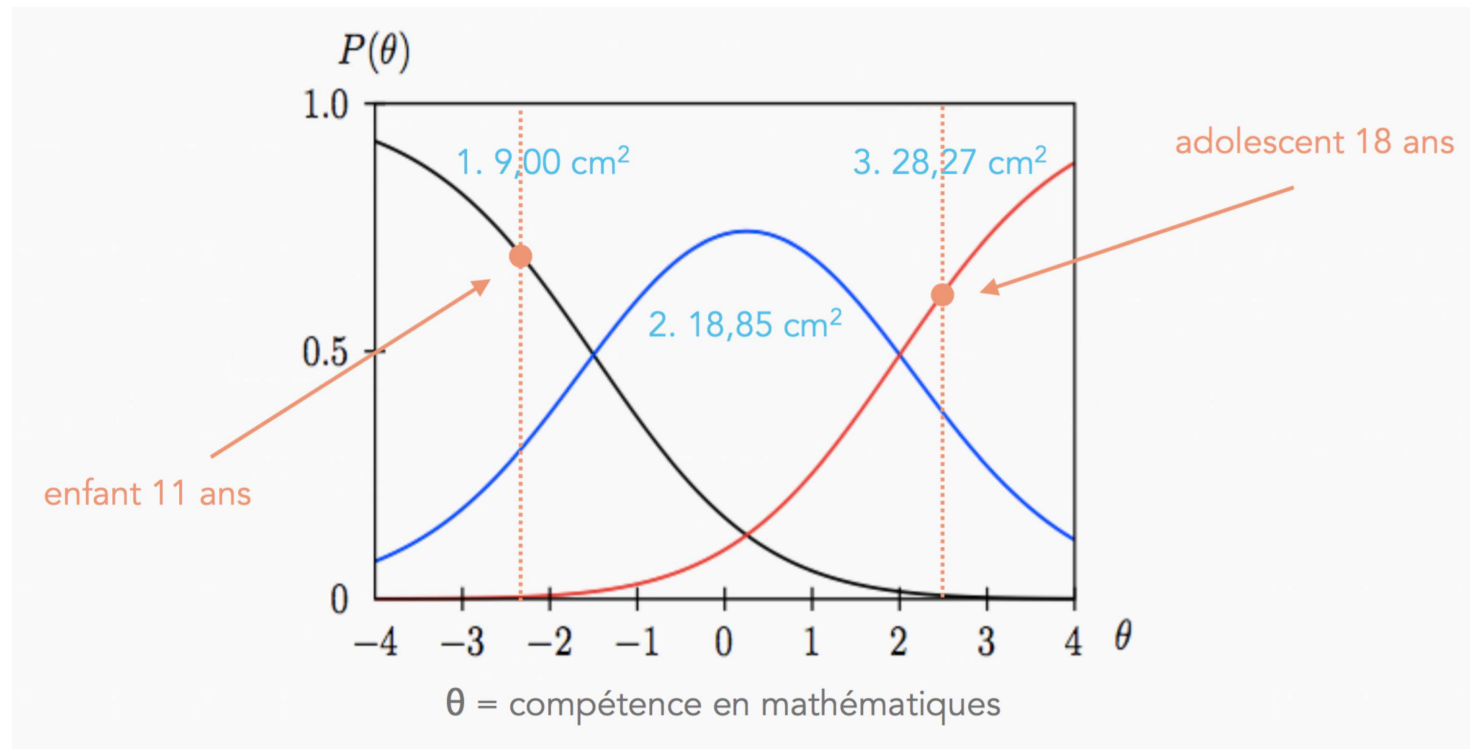
Quelle est la surface d'un cercle ayant un rayon de 3 cm ?

- (a) 9,00 cm<sup>2</sup>
- (b) 18,85 cm<sup>2</sup>
- (c) 28,27 cm<sup>2</sup>

---

6. I PARTCHEV. *A visual guide to item response theory*. Friedrich-Schiller-Universität Jena. 2004.





# Estimateurs WLS

**Table A.2** Robust Estimators Available in lavaan.

Estimator Variant	Description
<i>Maximum Likelihood (ML)</i>	
MLM	Estimation with robust standard errors and Satorra-Bentler scaled test statistic. For complete data only.
MLMVS	Estimation with robust standard errors and a mean- and variance-adjusted test statistic (Satterthwaite approach). For complete data only.
MLMV	Estimation with robust standard errors and a mean- and variance-adjusted test statistic (scale-shifted approach). For complete data only.
MLF	Estimation with standard errors based on the first-order derivatives and a conventional test statistic. For both complete and incomplete data.
MLR	Estimation with robust (Huber-White) standard errors and a scaled test statistic that is asymptotically equal to the Yuan-Bentler test statistic. For both complete and incomplete data.
<i>Least-Squares (DWLS/ULS)</i>	
WLSM <sup>a,b</sup>	Weighted least squares estimation with robust standard errors and a mean-adjusted test statistic. For complete data only.
WLSMVS <sup>a,b</sup>	Weighted least squares estimation with robust standard errors and a mean- and variance-adjusted test statistic (Satterthwaite approach). For complete data only.
WLSMV <sup>a,b</sup>	Weighted least squares estimation with robust standard errors and a mean- and variance-adjusted test statistic (scale-shifted approach). For complete data only.
ULSM	Unweighted least squares estimation with robust standard errors and a mean-adjusted test statistic. For complete data only.
ULSMVS	Unweighted least squares estimation with robust standard errors and a mean- and variance-adjusted test statistic (Satterthwaite approach). For complete data only.
ULSMV	Unweighted least squares estimation with robust standard errors and a mean- and variance-adjusted test statistic (scale-shifted approach). For complete data only.

<sup>a</sup> For the robust weighted least squares variants (WLSM, WLSMVS, WLSMV), lavaan uses the diagonal of the weight matrix for estimation, but uses the full weight matrix to correct the standard errors and to compute the test statistic.

<sup>b</sup> As of version 05.15 in lavaan, categorical data can have missing data using missing="pairwise" argument.

## Illustration : ACP sur matrice de corrélation polychorique

Echelle HADS à 2 dimensions ( $N = 201$  patients) :

```
load("HADS.RData")  
dep <- c(1,3,4,5,9,13,14)  
anx <- seq(1,14)[-c(1,3,4,5,9,13,14)]  
library(polycor)  
ddep <- as.data.frame(lapply(data[,dep], as.ordered)) ❶  
C <- hetcor(ddep) ❷  
principal(C$correlations, rotate = "none")
```

## Illustration : CFA sur données catégorielles

**Deux options** : travailler avec la matrice de corrélations polychoriques ou utiliser les fonctionnalités de lavaan.

Premier cas de figure :

```
m <- 'Dep ~ Y1 + Y3 + Y4 + Y5 + Y9 + Y13 + Y14'
r <- cfa(model = m, sample.cov = C$correlations,
          sample.nobs = nrow(ddep), std.lv = TRUE)
parameterEstimates(r, ci = FALSE, standardized = TRUE)
```

Deuxième cas de figure (suppose que les variables sont des facteurs avec des niveaux ordonnés) :

```
r <- cfa(model = m, data = ddep, std.lv = TRUE)
```

Fichier de données et scripts R disponibles à l'adresse suivante :  
<https://bitbucket.org/chlallanne/eespe11>

– Typeset with FoilTEX (version 2), Revision f4328f7