

# **ANALYSES DE CLASSES LATENTES AVEC R**

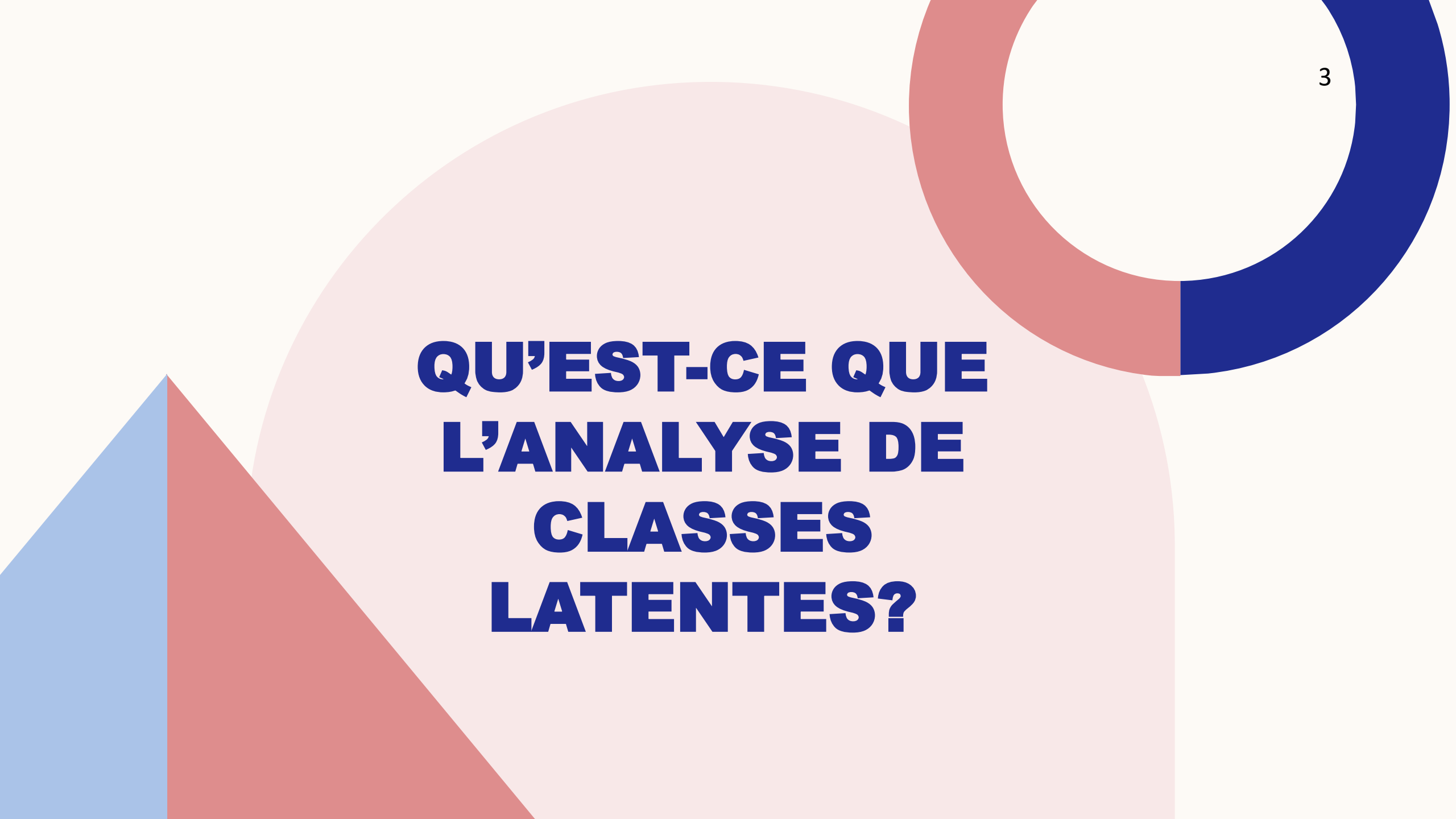
Marion Chatelois, B. Sc.

Agathe Bellemare-Lepage, B. Sc.

Pier-Olivier Caron, Ph. D.

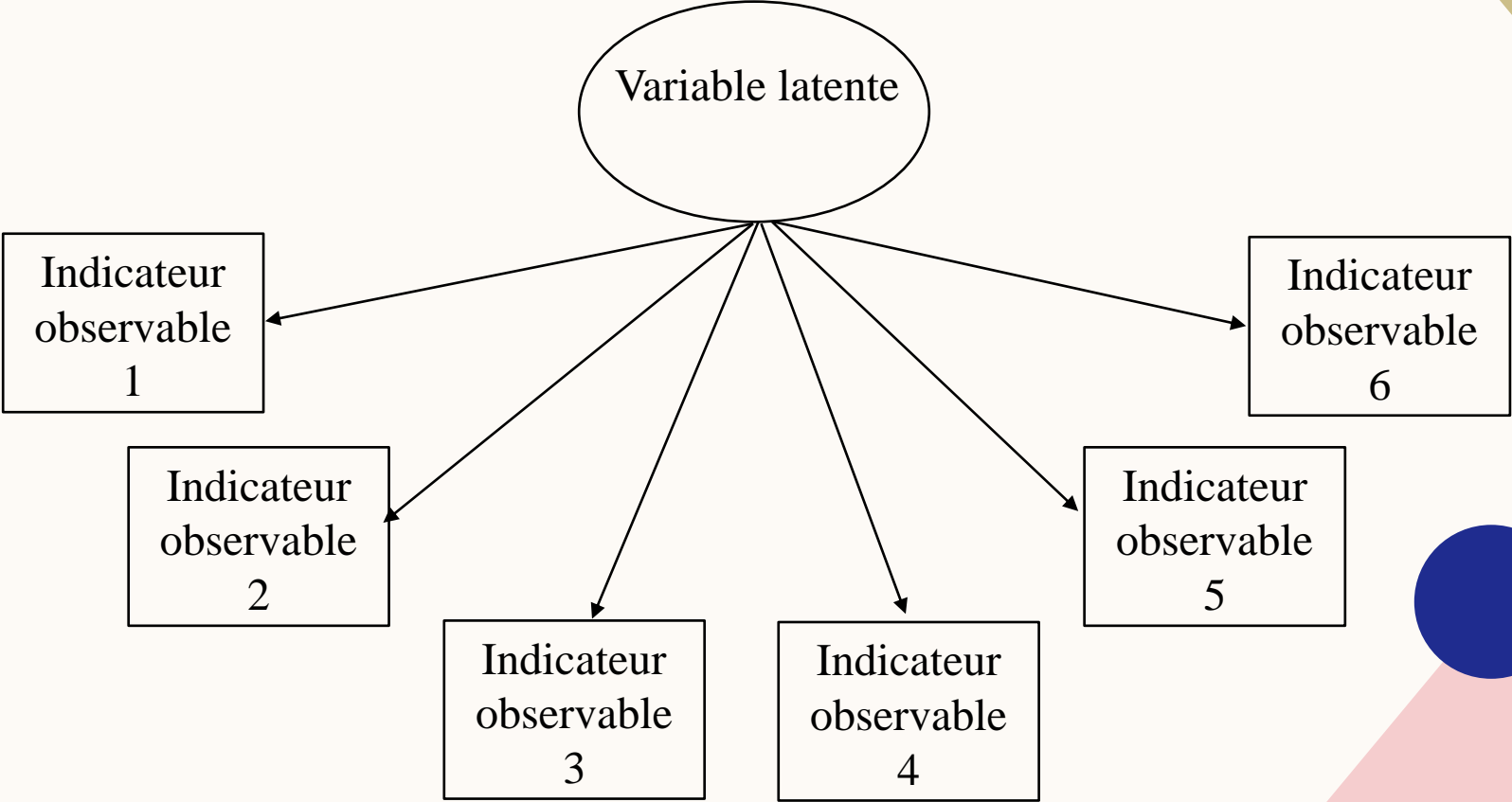
# R

- Logiciel accessible **gratuitement**
- Avantages en termes de **programmation** d'analyses, de **visualisation** des données, de **gestion** des variables et de l'environnement de travail
- **Rstudio** : interface graphique à utiliser conjointement



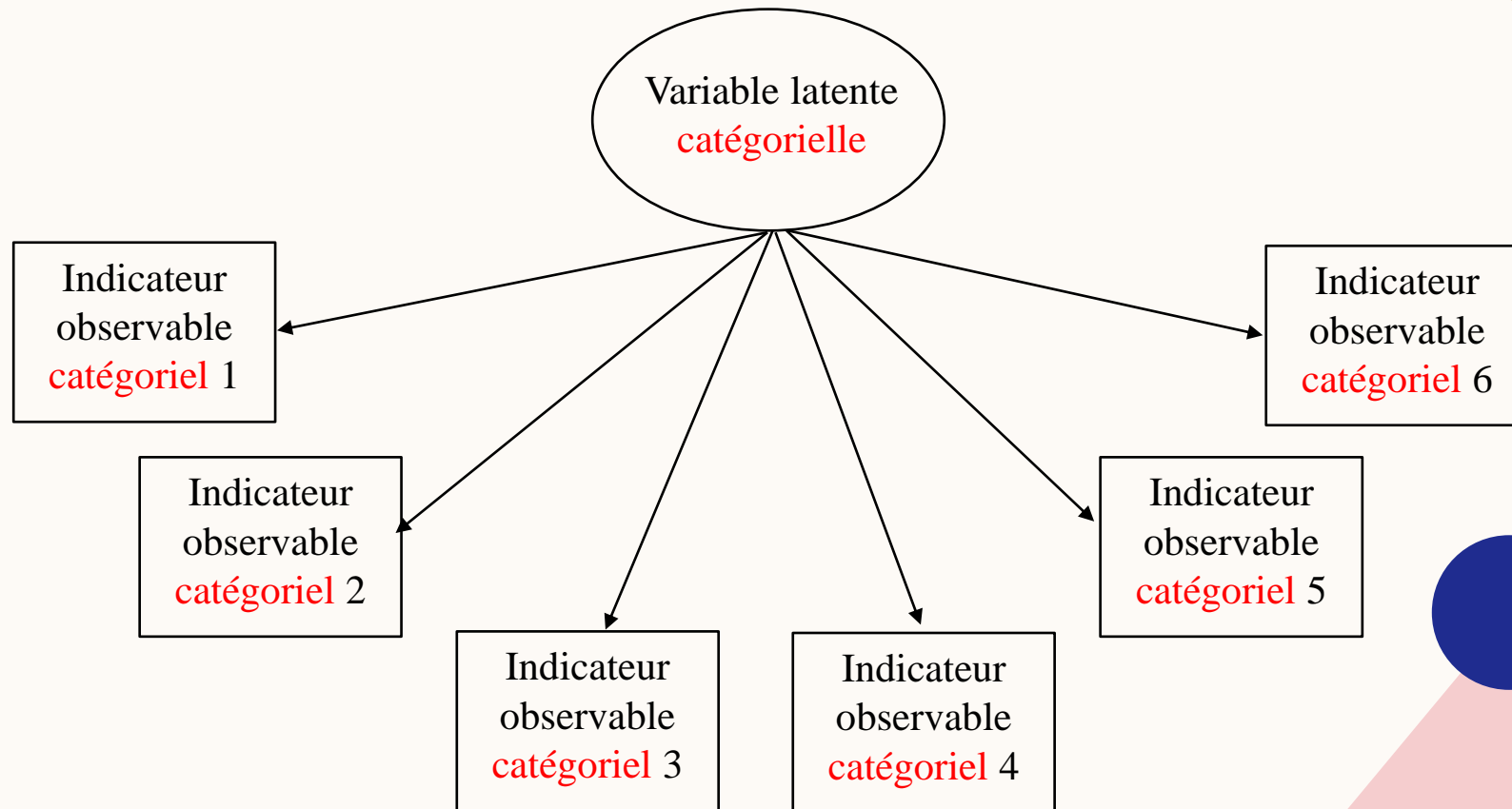
# **QU'EST-CE QUE L'ANALYSE DE CLASSES LATENTES?**

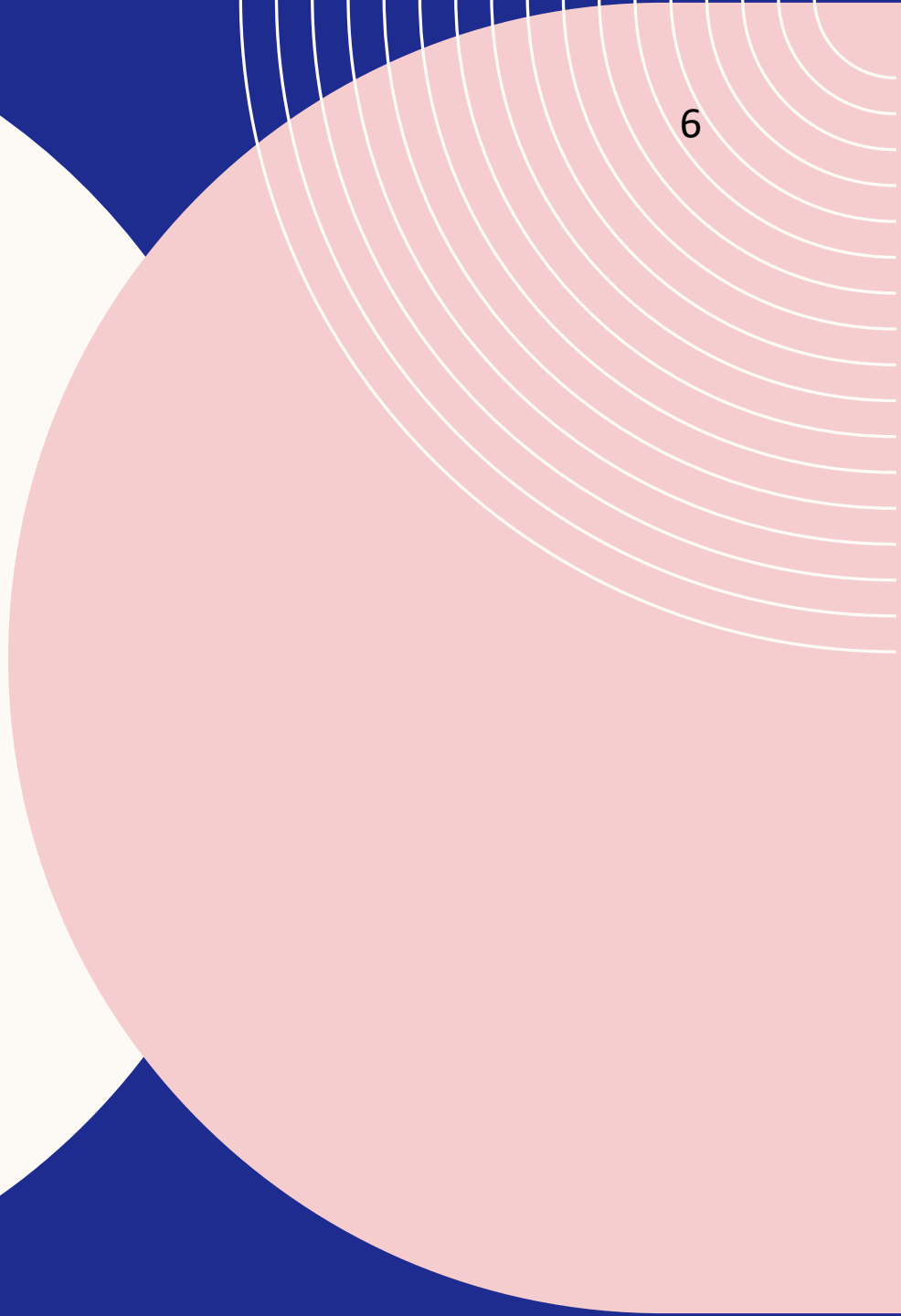
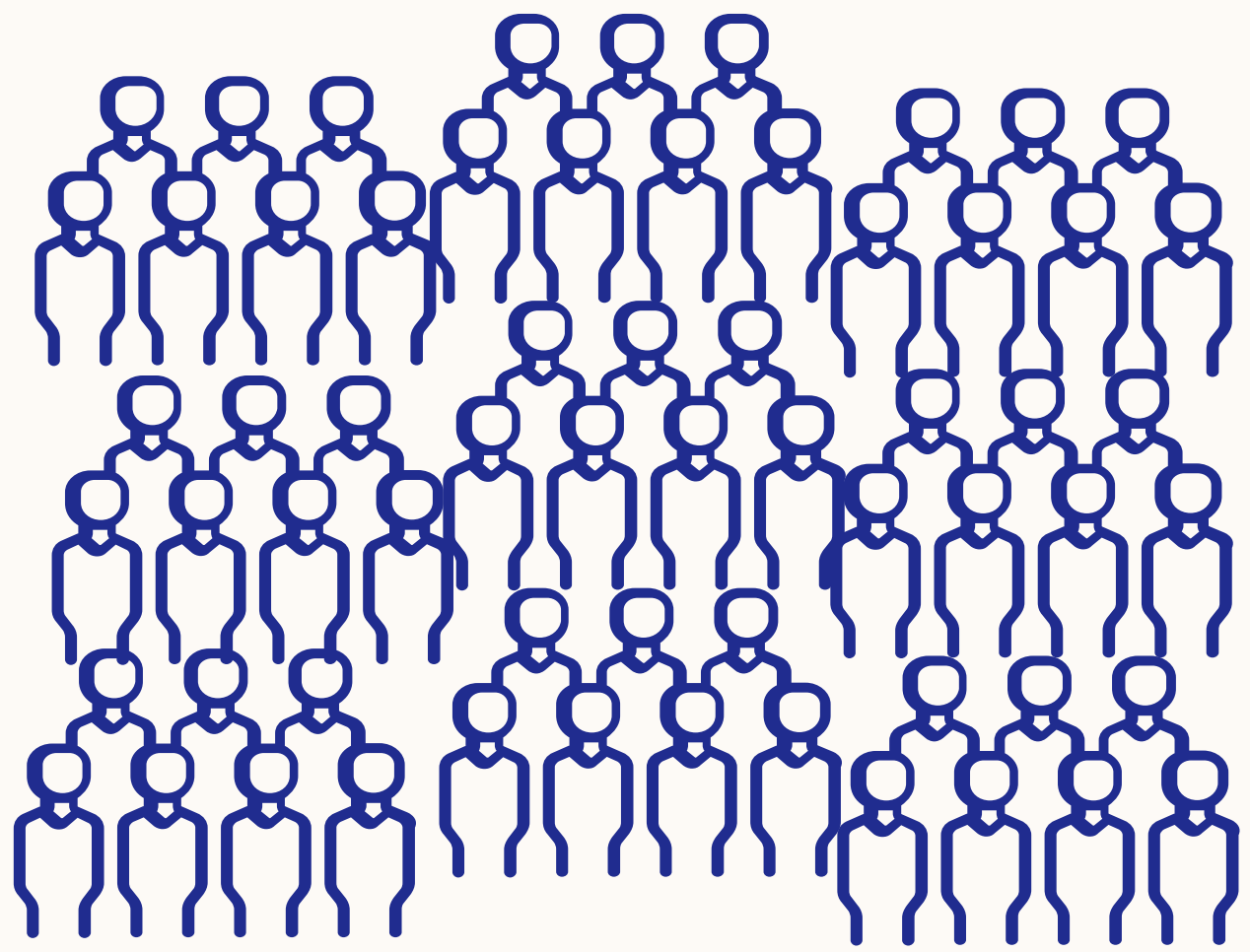
# FAMILLE DES ANALYSES DE STRUCTURES LATENTES

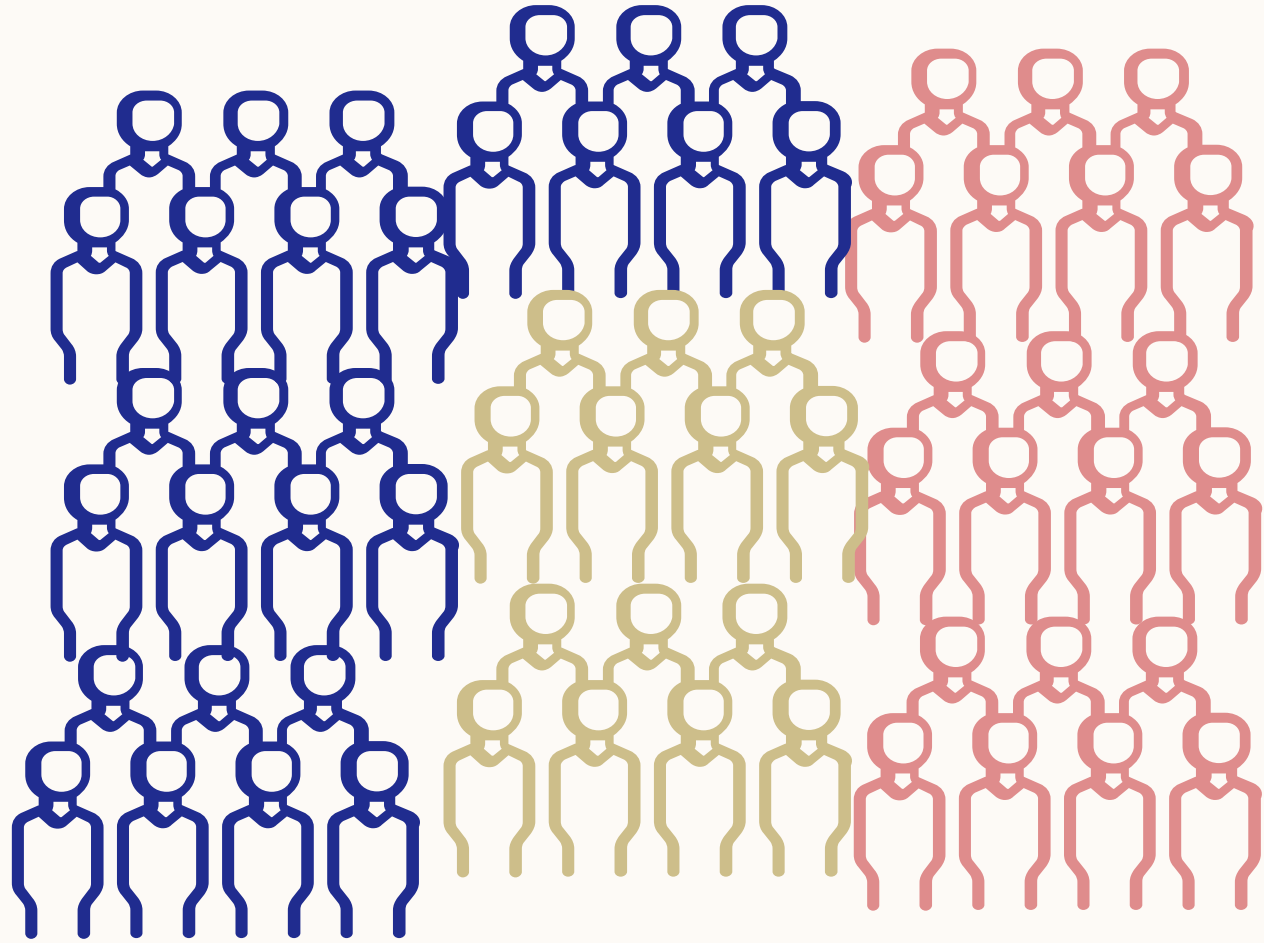


# ANALYSES DE CLASSES LATENTES

5

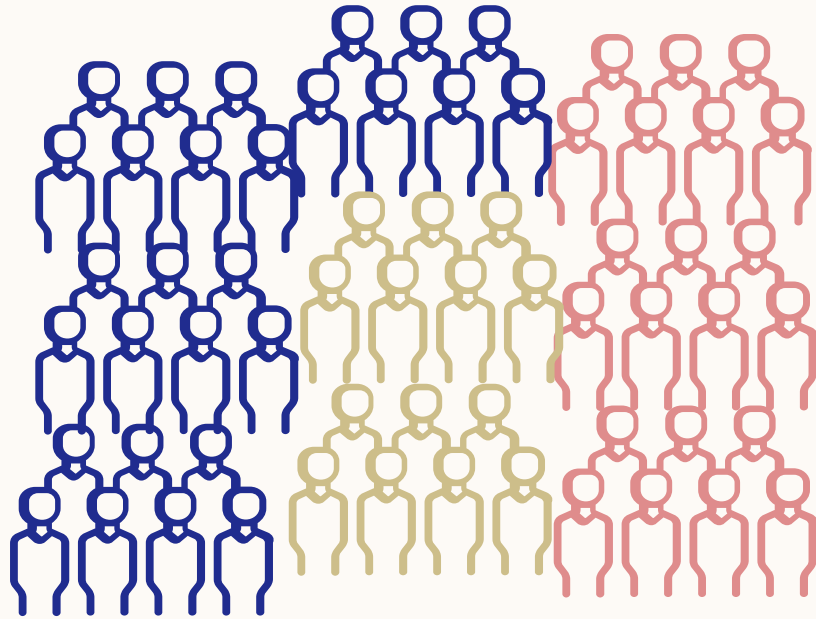






# IMPLICATIONS

- Personnalisation des interventions aux besoins et enjeux d'une classe spécifique



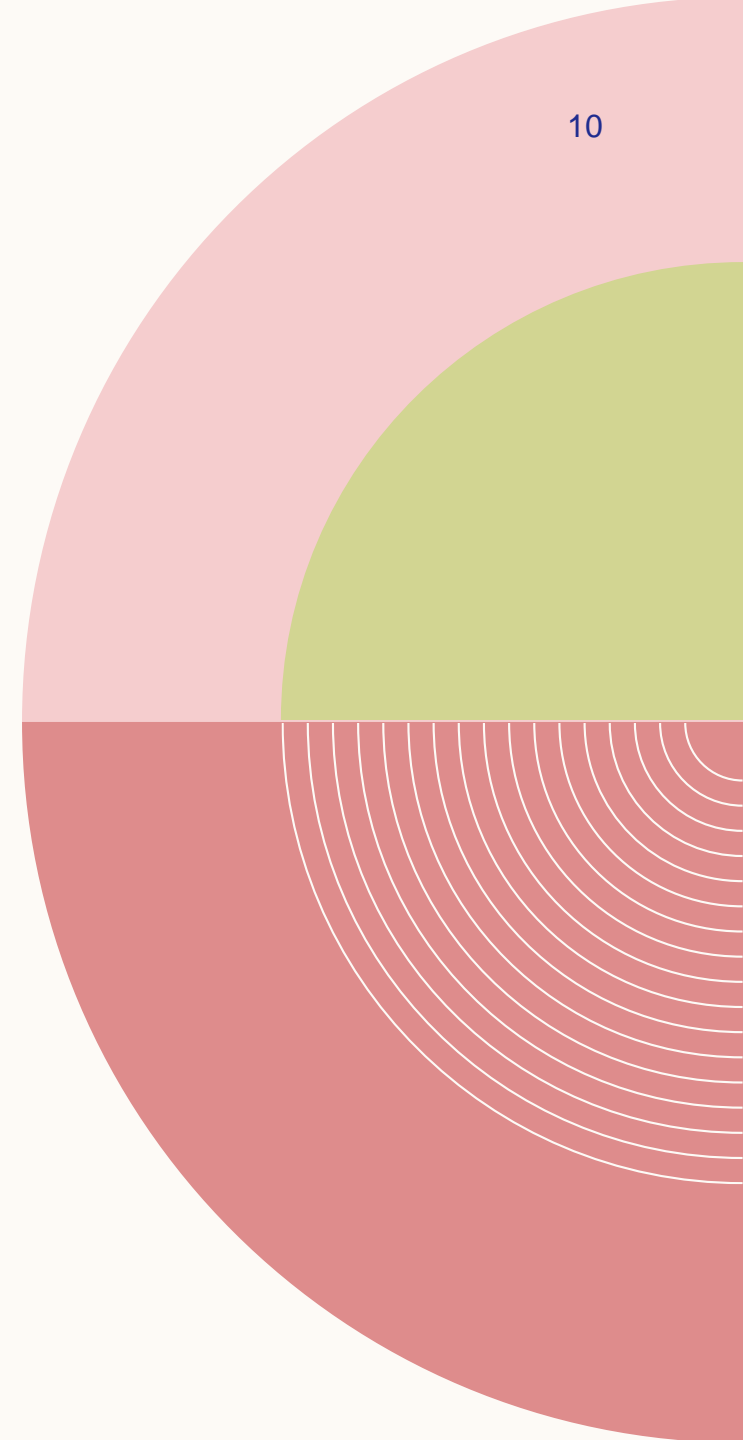
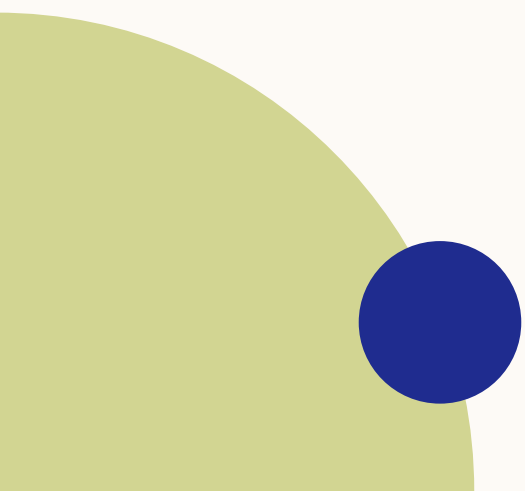


# **POSTULAT THÉORIQUE L'INDÉPENDANCE LOCALE**

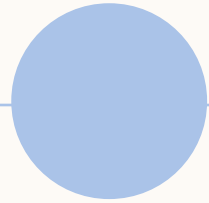
Les interrelations entre les indicateurs observables qui ne peuvent pas être attribuées à la chance sont expliquées par la variable latente.

# SITUATION HYPOTHÉTIQUE

10

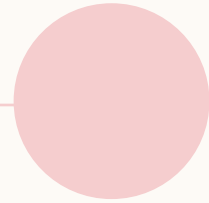


# LA VIOLENCE DANS LES RELATIONS AMOUREUSES



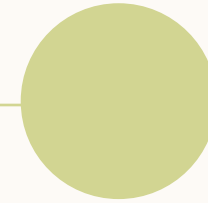
## VIOLENCE PHYSIQUE

Intention de faire mal à son partenaire



## VIOLENCE PSYCHOLOGIQUE

Intention de faire mal à son partenaire émotionnellement



## VIOLENCE SEXUELLE

Forcer un partenaire à poser des actes sexuels sans son consentement



## **VIOLENCE RÉACTIVE**

En réaction à une  
provocation perçue



## **VIOLENCE PROACTIVE**

Comportement prémédité ou  
stratégique dans le but d'en  
arriver à ses fins

# OBJECTIF

Identifier les classes de **perpétrateurs de violence dans les relations amoureuses** à l'adolescence, qui se distinguent quant aux **formes de violence utilisées** et à la perpétration **réactive ou proactive**

# SIX INDICATEURS

Violence physique réactive

Violence physique proactive

Violence psychologique réactive

Violence psychologique proactive

Violence sexuelle réactive

Violence sexuelle proactive

# ÉLÉMENTS À CONSIDÉRER

- **Taille** échantillon : entre 300 et 1000
- Choix des **indicateurs** : au moins 3, mais un plus grand nombre est bénéfique
- **Dichotomisation** : selon le domaine de recherche concerné

# DÉMONSTRATION

Tutoriel sur la plateforme R avec poLCA



# DÉROULEMENT TYPIQUE

1. Tester séquentiellement des modèles avec un nombre croissant de classes latentes
2. Comparer les modèles
3. Cesser d'estimer des classes additionnelles quand il n'y a plus d'amélioration
4. Sélectionner le modèle permettant le meilleur ajustement

## Analyses de classes latentes avec poLCA

```
jd <- read.csv2("violencerelamour.csv")
```

```
install.packages("poLCA")  
library(poLCA)
```

## Formule utilisée dans l'analyse

```
f1 <- as.formula(cbind(rea_sex,  
rea_psy, rea_phys, pro_phys, pro_sex,  
pro_psy) ~ 1)
```

## Modèle de classes latentes

```
poLCA(formula, data, nclass)
```

## Analyser les données

```
LCA1 <- poLCA(f1, data = jd, nclass = 1)
LCA2 <- poLCA(f1, data = jd, nclass = 2)
LCA3 <- poLCA(f1, data = jd, nclass = 3)
```

## Analyser les données

```
LCA4 <- poLCA(f1, data=jd, nclass=4)
```

```
ALERT : iterations finished, MAXIMUM  
LIKELIHOOD NOT FOUND
```

```
LCA4 <- poLCA(f1, data=jd, nclass=4,  
maxiter = 10000)
```

## Exemple de « sortie »

23

```
LCA3 <- poLCA(f1, data = jd, nclass = 3)
```

Conditional item response (column) probabilities, by outcome variable, for each class (row)

```
$rea_sex
```

```
Pr(1) Pr(2)
```

```
class 1: 0.1650 0.8350
```

```
class 2: 0.9591 0.0409
```

```
class 3: 0.2173 0.7827
```

```
$pro_phys
```

```
Pr(1) Pr(2)
```

```
class 1: 0.2390 0.7610
```

```
class 2: 0.8663 0.1337
```

```
class 3: 0.9848 0.0152
```

```
$rea_psy
```

```
Pr(1) Pr(2)
```

```
class 1: 0.1069 0.8931
```

```
class 2: 0.9070 0.0930
```

```
class 3: 0.1833 0.8167
```

```
$pro_sex
```

```
Pr(1) Pr(2)
```

```
class 1: 0.3089 0.6911
```

```
class 2: 0.9616 0.0384
```

```
class 3: 0.9876 0.0124
```

```
$rea_phys
```

```
Pr(1) Pr(2)
```

```
class 1: 0.1943 0.8057
```

```
class 2: 0.9166 0.0834
```

```
class 3: 0.2151 0.7849
```

```
$pro_psy
```

```
Pr(1) Pr(2)
```

```
class 1: 0.1861 0.8139
```

```
class 2: 0.8800 0.1200
```

```
class 3: 0.9595 0.0405
```

## Exemple de « sortie »

24

Estimated class population shares

0.4388 0.1367 0.4245

Predicted class memberships (by modal posterior prob.)

0.4312 0.1262 0.4425

=====  
Fit for 3 latent classes:  
=====

number of observations: 800

number of estimated parameters: 20

residual degrees of freedom: 43

maximum log-likelihood: -2195.821

AIC(3): 4431.642

BIC(3): 4525.334

$G^2(3)$ : 51.86892 (Likelihood ratio/deviance statistic)

$X^2(3)$ : 43.6554 (Chi-square goodness of fit)



## Entropie relative

```
poLCA.relentropy <- function(model) {  
  pik <- as.data.frame(model$posterior)  
  E <- 1 - (sum(-1 * pik * log(pik), na.rm = TRUE) /  
            (nrow(pik) * log(ncol(pik)))) )  
  return(E)  
}
```

```
poLCA.relentropy(LCA1)  
poLCA.relentropy(LCA2)  
poLCA.relentropy(LCA3)  
poLCA.relentropy(LCA4)
```

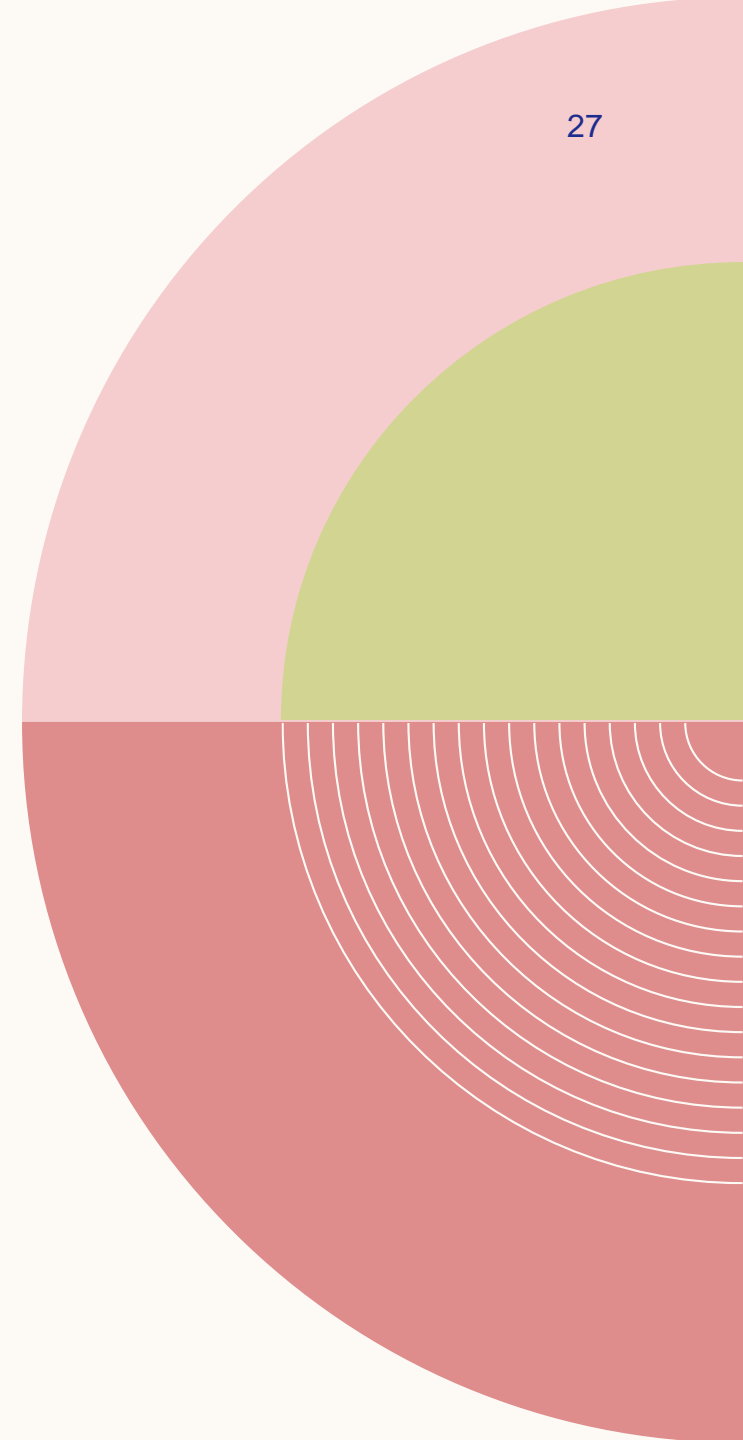
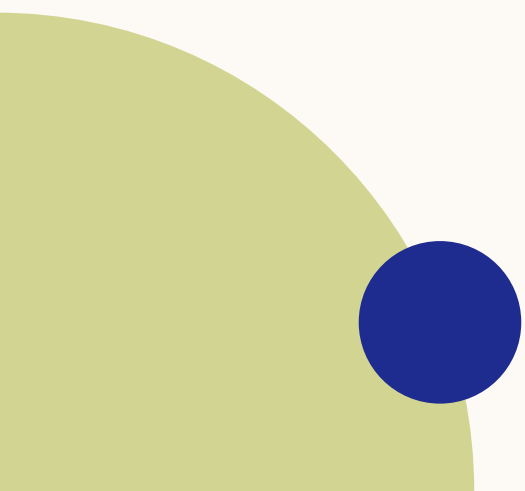
## Test de Lo-Mendell-Rubin

```
install.packages("tidyLPA")  
library(tidyLPA)
```

```
calc_lrt(n = LCA2$N,  
         null_ll = LCA2$l1lik,  
         null_param = LCA2$npar,  
         null_classes = length(LCA2$P),  
         alt_ll = LCA3$l1lik,  
         alt_param = LCA3$npar,  
         alt_classes = length(LCA3$P))
```

# INTERPRÉTATION

27



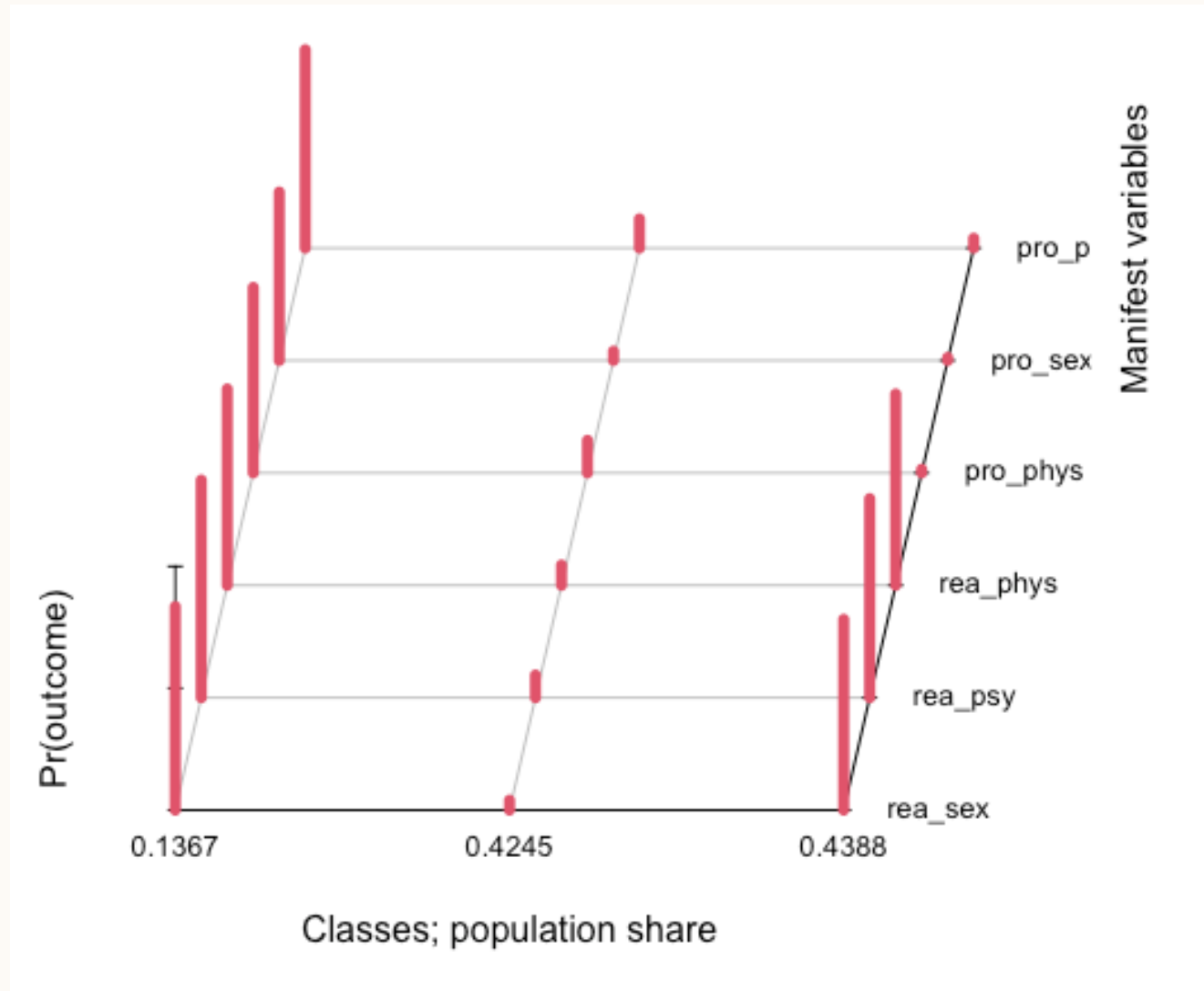
**TABLEAU 1**  
**INDICES D'AJUSTEMENT DU MODÈLE ET STATISTIQUE DIAGNOSTIQUE**

Modèles testés	Maximum log-likelihood	AIC	BIC	LMR*	Entropie relative	Degrés de liberté
1 classe	-2687.896	5387.793	5415.9	N/A	N/A	57
2 classes	-2365.285	4756.57	4817.47	614.576***	0.802721	50
3 classes	-2195.821	<b>4431.642</b>	<b>4525.334</b>	<b>322.830***</b>	<b>0.8396914</b>	43
4 classes	-2191.063	4436.127	4562.611	9.064	0.7452709	36

Note. \*Lo-Mendell-Rubin ad-hoc adjusted likelihood ratio test. \*\*\* p<.001

FIGURE 2

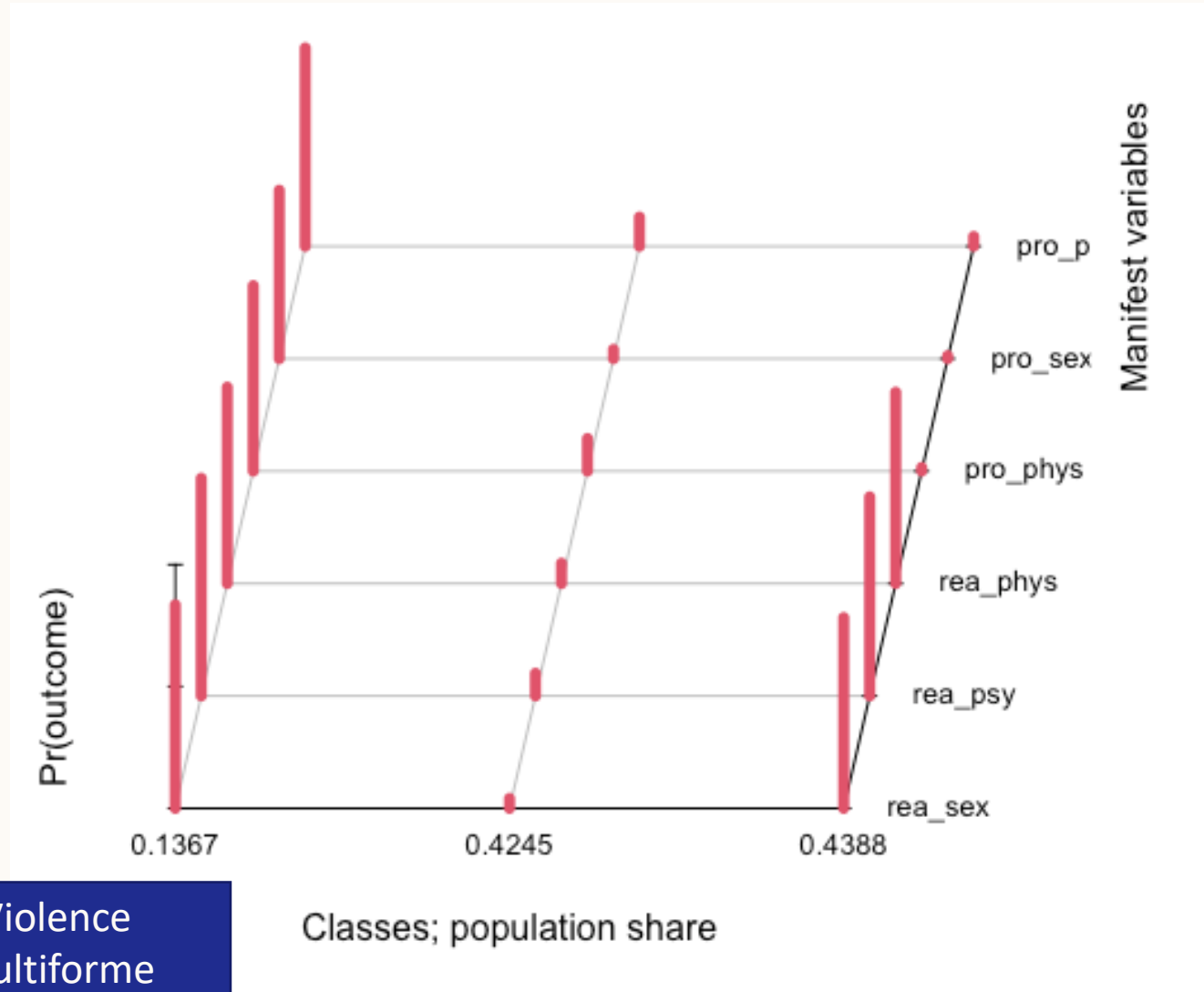
*SCHÉMAS DE RÉPONSES POUR LE MODÈLE À TROIS CLASSES LATENTES*



Plot (LCA3)

FIGURE 2

*SCHÉMAS DE RÉPONSES POUR LE MODÈLE À TROIS CLASSES LATENTES*

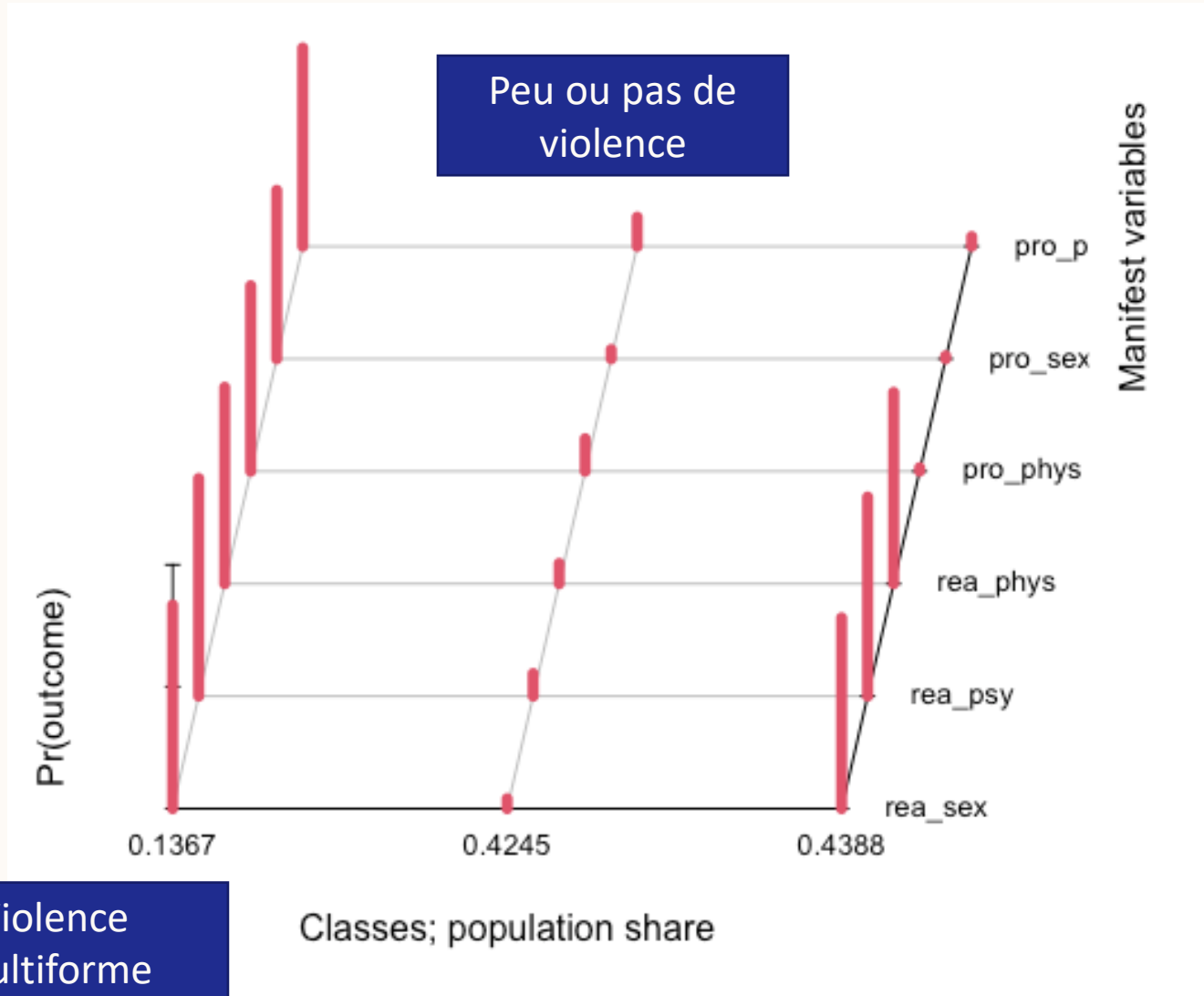


Plot (LCA3)

Violence  
multiforme

FIGURE 2

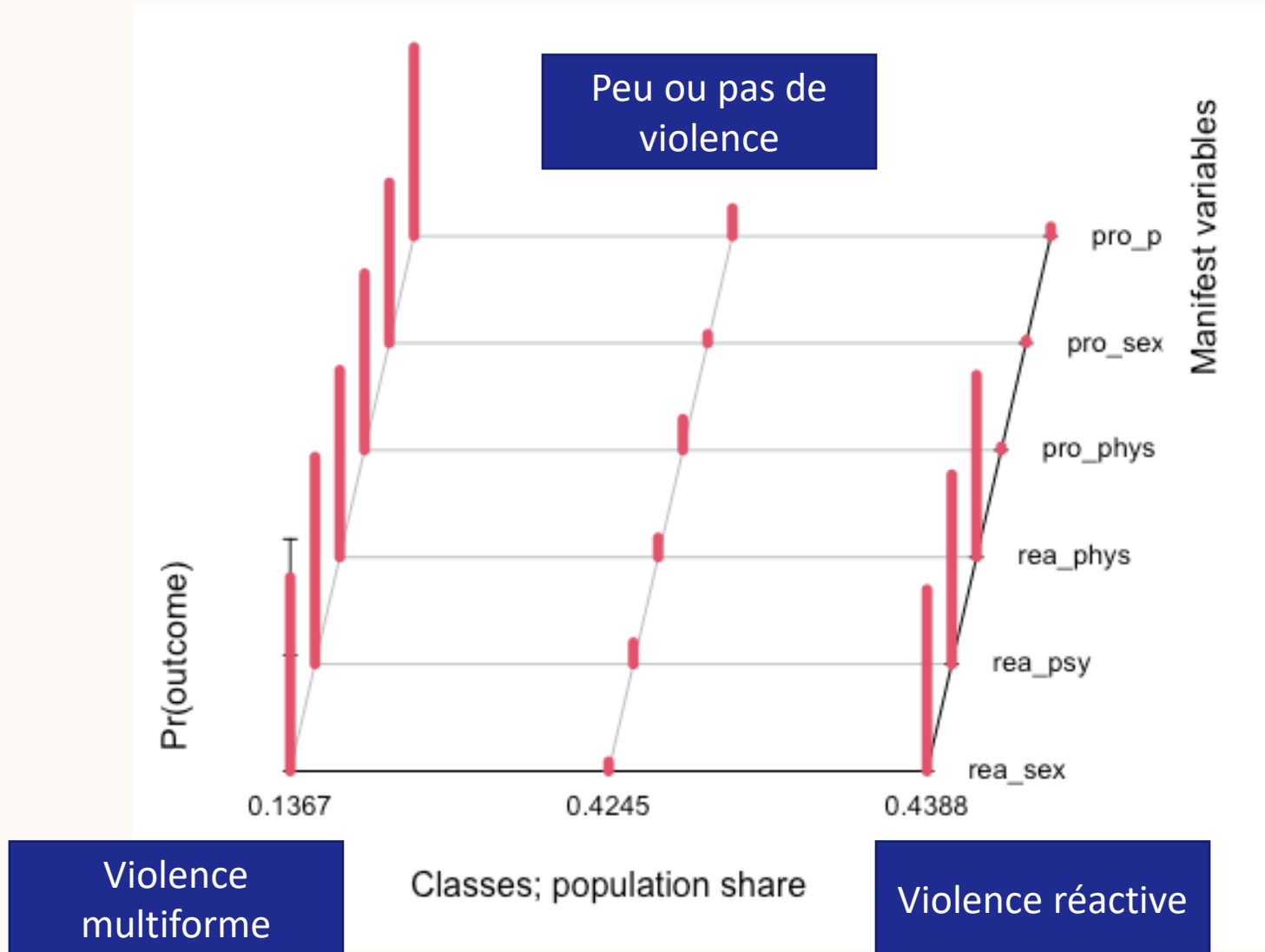
*SCHÉMAS DE RÉPONSES POUR LE MODÈLE À TROIS CLASSES LATENTES*



Plot (LCA3)

FIGURE 2

*SCHÉMAS DE RÉPONSES POUR LE MODÈLE À TROIS CLASSES LATENTES*



Plot (LCA3)



# Exemple de discussion

- Rapporter les classes à nouveau
- Faire des liens avec les études antérieures

# CONCLUSION

# RÉFÉRENCES

- Bushman, B. J. et Anderson, C. A. (2001). Media violence and the American public: Scientific facts versus media misinformation. *American Psychologist*, 56(6-7), 477. <https://doi-org/10.1037/0003-066X.56.6-7.477>
- Caron, P.-O. (en écriture). *Méthodes quantitatives avec R*. <https://mqr.teluq.ca>
- Centers for Disease Control and Prevention (2017). *Preventing Teen Dating Violence*. Atlanta: Centers for Disease Control and Prevention. [www.cdc.gov/](http://www.cdc.gov/).
- Karnowski, V. (2017). Latent class analysis. Dans Matthes, J. (dir.), *The International Encyclopedia of Communication Research Methods* (1-10). <https://doi-org/10.1002/9781118901731.iecrm0130>
- Nylund-Gibson, K. et Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, 4(4), 440. <https://doi-org/10.1037/tps0000176>
- Porcu, M. et Giambona, F. (2017). Introduction to latent class analysis with applications. *The Journal of Early Adolescence*, 37(1), 129-158. <https://doi-org/10.1177/027243161664845>
- Ramirez, J. M. et Andreu, J. M. (2006). Aggression, and some related psychological constructs (anger, hostility, and impulsivity) Some comments from a research project. *Neuroscience & Biobehavioral Reviews*, 30(3), 276-291. <https://doi.org/10.1016/j.neubiorev.2005.04.015>

# SCRIPT R CRÉATION JD

```
# Installer (au besoin) et appeler le package poLCA
install.packages("poLCA")
library(poLCA)
# Déclarer une valeur de départ, nommée « graine » ou seed en
anglais,
pour s'assurer de toujours obtenir les mêmes valeurs

set.seed(42)

# Déterminer vers quelle solution l'ACL converge, ainsi que les
probabilités de réponses à chaque item des six indicateurs, pour
chacune des trois classes prévues, à l'aide de la fonction probs

# Créer un modèle à 3 classes composé de 6 indicateurs dichotomiques.
Ainsi, il y a 6 matrices (indicateurs), de 3 lignes (classes) et 2
colonnes (nombre d'options par indicateur)
```

# SCRIPT R CRÉATION JD

```
probs <- list(matrix(c(.95, .05,  
                      .99, .01,  
                      .30, .70),  
                    ncol = 2, byrow = TRUE),  
             matrix(c(.95, .05,  
                      .20, .80,  
                      .15, .85),  
                    ncol = 2, byrow = TRUE),  
             matrix(c(.90, .10,  
                      .20, .80,  
                      .20, .80),  
                    ncol = 2, byrow = TRUE),
```

# SCRIPT R CRÉATION JD

```
matrix(c(.90, .10,  
        .98, .02,  
        .23, .77),  
       ncol = 2, byrow = TRUE),  
matrix(c(.90, .10,  
        .20, .80,  
        .10, .90),  
       ncol = 2, byrow = TRUE),  
matrix(c(.90, .10,  
        .97, .03,  
        .20, .80),  
       ncol = 2, byrow = TRUE))
```

# SCRIPT R CRÉATION JD

```
# Utiliser la fonction simdata() pour créer un jeu de données fictif, ici contenant 800 participants
```

```
simdata <- poLCA.simdata(N = 800, probs)
```

```
# Extraire le jeu de la liste de sortie de simdata()
```

```
jd <- simdata$dat
```

```
# Renommer les variables
```

```
colnames(jd) <- c("pro_sex", "rea_sex", "rea_phys",  
"pro_phys", "rea_psy", "pro_psy")
```