

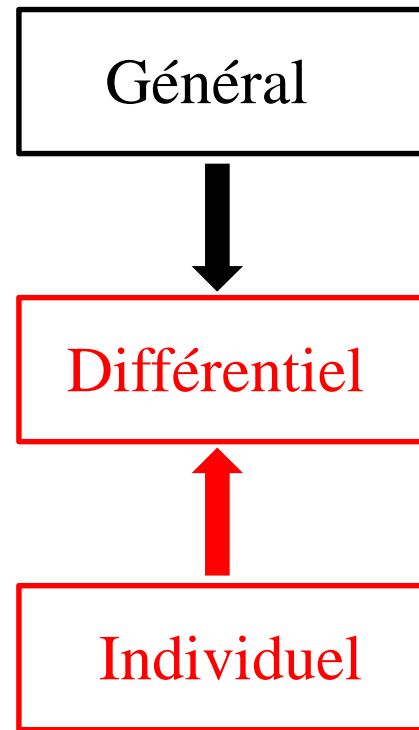
Utilisation des modèles de Markov cachés dans l'analyse individuelle des variations d'état de contrôle attentionnel à une tâche d'alternance

Jacques Juhel & Stéphane Deline

CRPCC/LPE

Université Européenne de Bretagne – Rennes 2

De l'individuel au différentiel



Regroupement d'individus ayant un profil de variabilité intra-individuelle « comparable »
(Molenaar, 2004)

Le paradigme d'alternance entre tâches dans l'étude des mécanismes de contrôle cognitif

► L'étude des coûts d'alternance

1) Condition mixte *vs* condition homogène: coûts globaux d'alternance ($TR_{A \text{ dans } ABAB} > TR_{A \text{ dans } AAAA}$).

Interprétation : coûts associés au maintien et à la coordination des séquences de tâches (par ex., Kray & Lindenberger, 2000).

2) Items alternés *vs* items répétés en condition mixte : coûts locaux d'alternance ($TR_{\text{alternés}} > Tr_{\text{répétés}}$).

Interprétation : coûts réfléchissant l'engagement des processus exécutifs de désactivation de la séquence antérieure non pertinente et d'activation de la séquence pertinente pour l'essai en cours (Monsell, 2003).

Le paradigme d'alternance entre tâches dans l'étude des mécanismes de contrôle cognitif

Effets liés à l'âge (méta-analyse de Wasylshyn *et al.*, 2011)

- ▶ Les coûts de mixage sont sensibles au processus de vieillissement

Interprétation : baisse d'efficacité liée à l'âge des processus de maintien et de coordination des séquences mentales (mémoire de travail).

- ▶ Les coûts locaux d'alternance **ne seraient pas sensibles** au processus de vieillissement

Interprétation : processus d'attention sélective préservés par l'âge.

mais voir par ex., Hartley, Kieley, & Slabach, 1990; Lien *et al.*, 2008; Meiran, Gotler, & Perlman, 2001; Reimers & Maylor, 2005.

Variations des TR

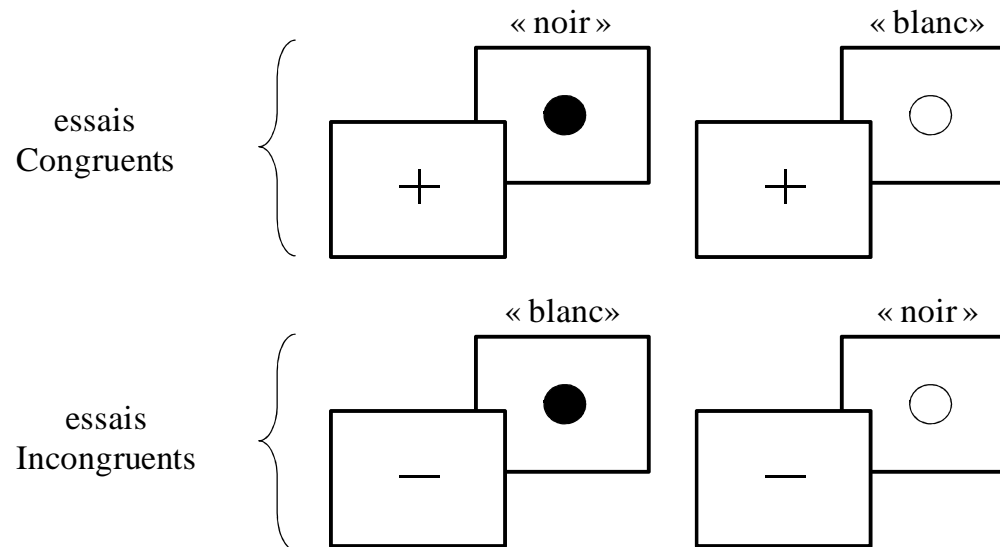
L'ensemble des réponses observées correspond au mode **typique** de fonctionnement de l'individu dans la situation proposée.

- ▶ Classiquement: suppression des TR aux essais échoués, des TR extrêmes (approche paramétrique ou non paramétrique, bornes arbitraires) et des TR « contaminés » par un échec précédent.

Par exemple :

« only correct trials that were preceded by at least two correct trials were included in the RT analysis. As outliers, we excluded RTs 4.5 standard deviations from the mean RT » (Lawo et al., 2012).

Exemple : tâche d'alternance entre essais de difficulté inégale



Série temporelle bivariée (TR et précision) : 208 essais I C I C I C I C...

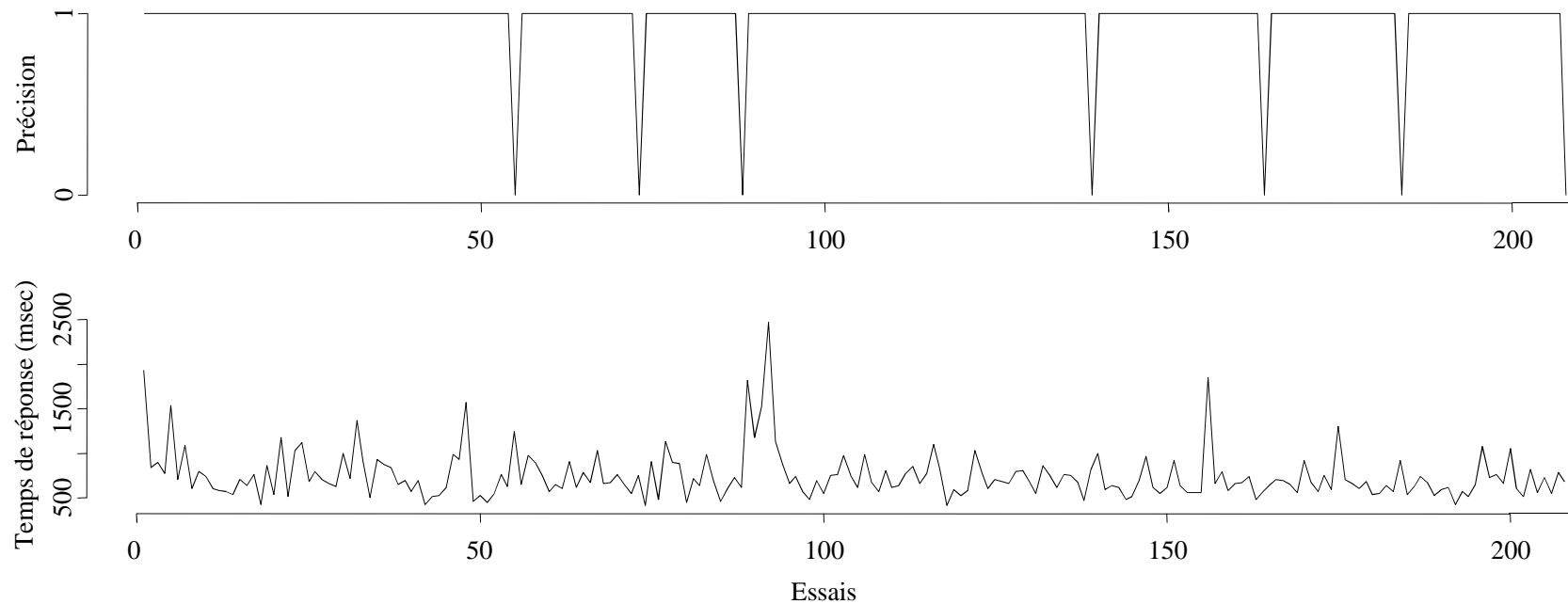
Effet du matériel (I vs C) confondu avec la condition d'alternance.

Y a-t-il un effet d'interaction entre l'âge et la variable manipulée « item C/I alterné » ?

Exemple : tâche d'alternance entre essais de difficulté inégale

23 participants âgés de 20 à 83 ans

Série temporelle bivariée : id 13



Modèles linéaires mixtes appliqués aux ln(TR)

Procédure classique: 12,52 % d'observations manquantes.

Paramètres (REML)	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5
Effets fixes					
Constante	6,776 (0,081)	6,723 (0,081)	6,723 (0,081)	6,060 (0,144)	6,109 (0,151)
Item		0,107 (0,009)	0,108 (0,019)	0,108 (0,019)	0,058 (0,049)
Age				0,014 (0,003)	0,013 (0,003)
Item x Age					0,001 (0,001)
Effets aléatoires					
Variance de la constante	0,1512	0,1513	0,1508	0,0813	0,0817
Variance de la pente Item			0,0069	0,0069	0,0068
<i>Variance résiduelle</i>	0,0853	0,0824	0,0808	0,0808	0,0808
BIC	1735	1608	1575	1577	1596

Anova(m3,m4)

$\chi^2=17,5$, ddl=1, $p<.001$

► Effets du type d'item, de l'âge mais pas d'effet d'interaction entre l'âge et le type d'item présenté en alternance.

Variations des TR

L'ensemble des réponses observées correspond au mode **typique** de fonctionnement de l'individu dans la situation proposée.

- ▶ Classiquement: suppression des TR aux essais échoués, des TR extrêmes (approche paramétrique ou non paramétrique, bornes arbitraires) et des TR « contaminés » par un échec précédent.

« only correct trials that were preceded by at least two correct trials were included in the RT analysis. As outliers, we excluded RTs 4.5 standard deviations from the mean RT » (Lawo et al., 2012).

- ▶ Toutes les latences sont « vraies ».

Les variations des TR réfléchissent des variations du niveau de contrôle attentionnel

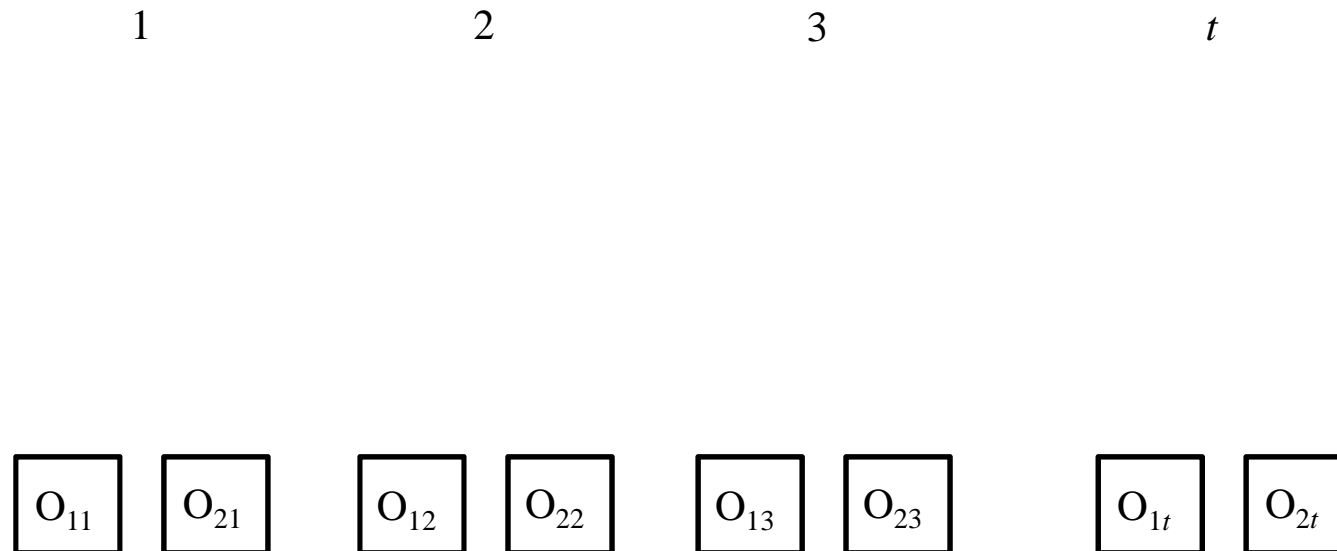
- ▶ Chercher d'abord à identifier **au niveau individuel** une variable cachée « état de contrôle attentionnel » correspondant à l'un et/ou l'autre des modes :
 - soutenu ou **optimal** : réponses rapides et précises;
 - relaxé : réponses plus lentes et moins précises.

- ▶ Etudier **dans un second temps** les différences liées à l'âge et les différences entre individus sous l'angle :
 - qualitatif : par ex., état stable *vs* état instable;
 - quantitatif : sur la base d'une forme ou d'une autre de validation du caractère « comparable » des observations individuelles.

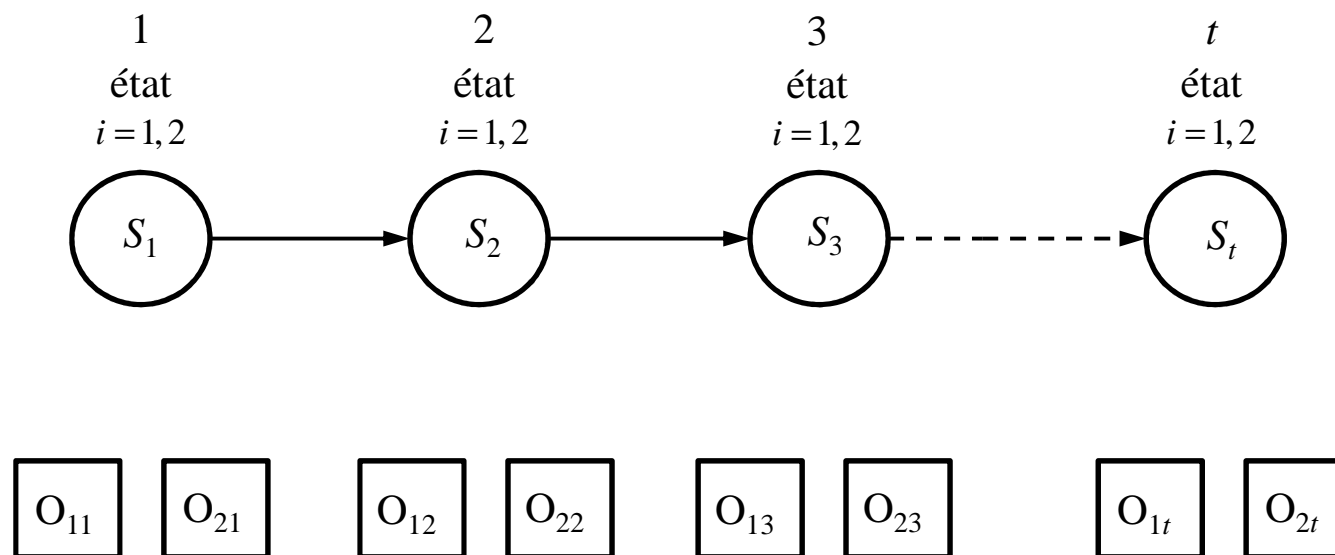
Identifier l'état d'un processus latent avec le Modèle de Markov Caché (HMM)

- ▶ Chaque observation est émise par un état - parmi k états d'un processus latent - selon une certaine loi de probabilité.
- ▶ Les observations sont réparties dans k différentes composantes d'un mélange de distributions.
- ▶ Les dépendances temporelles entre observations sont dues aux dépendances temporelles entre états (indépendance locale conditionnelle à l'état).
- ▶ L'état au temps t du processus latent ne dépend que de son état au temps $t-1$ (évolution markovienne des états).

Modèle de Markov Caché (HMM) : 2 observations à chaque pas de temps



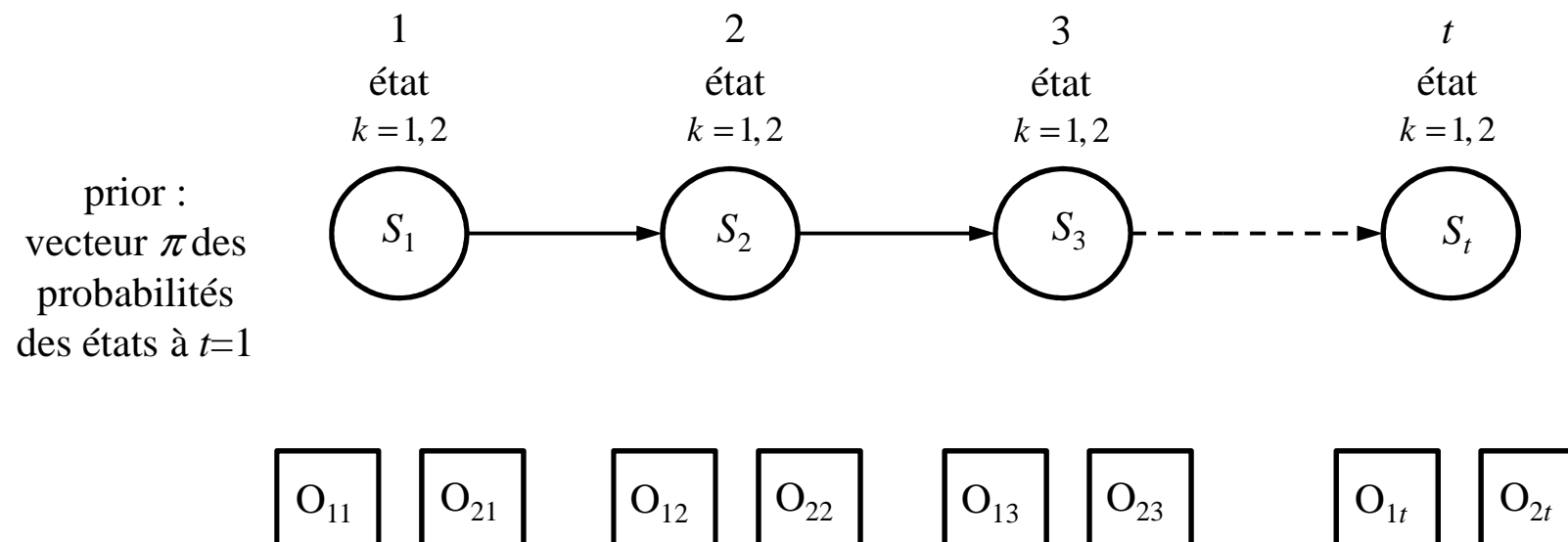
Modèle de Markov Caché (HMM) : 2 observations et 2 états possibles à chaque pas de temps



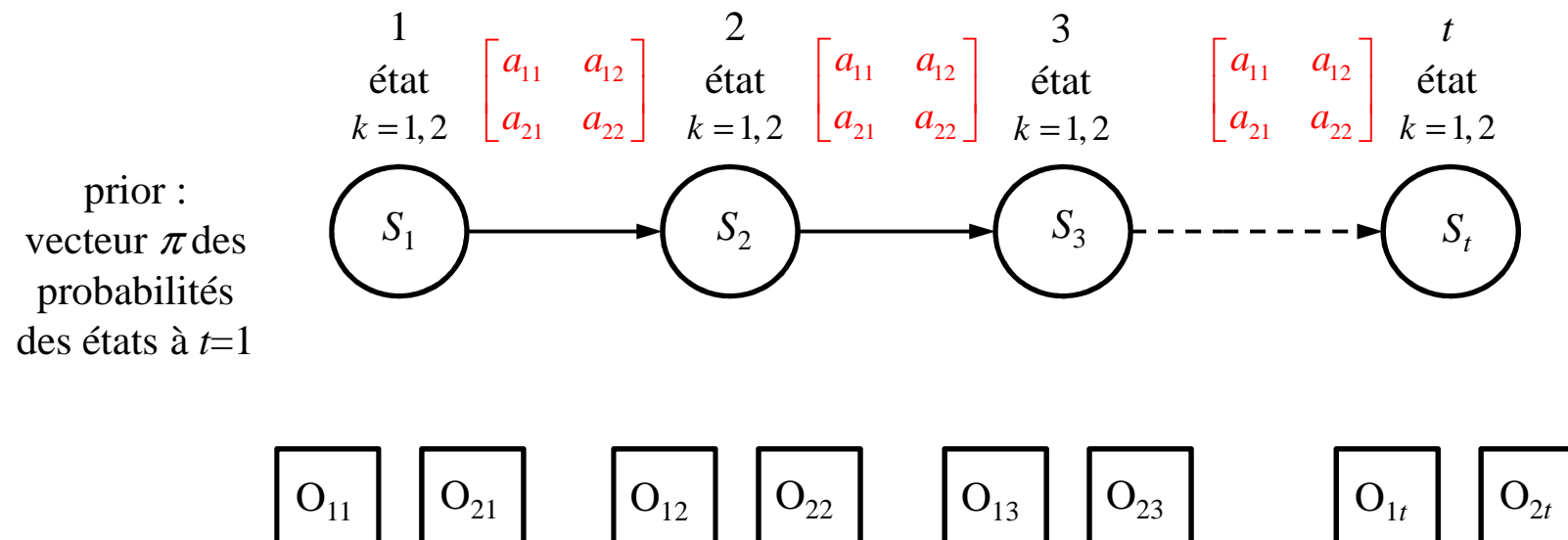
Hypothèse de 1^{er} ordre: $p(S_t | S_1, S_2, \dots, S_{t-1}) = p(S_t | S_{t-1})$

Chaîne de Markov: le passé et le futur sont conditionnellement indépendants lorsqu'on conditionne par rapport au présent.

Modèle de Markov Caché (HMM) : 2 observations et 2 états possibles à chaque pas de temps

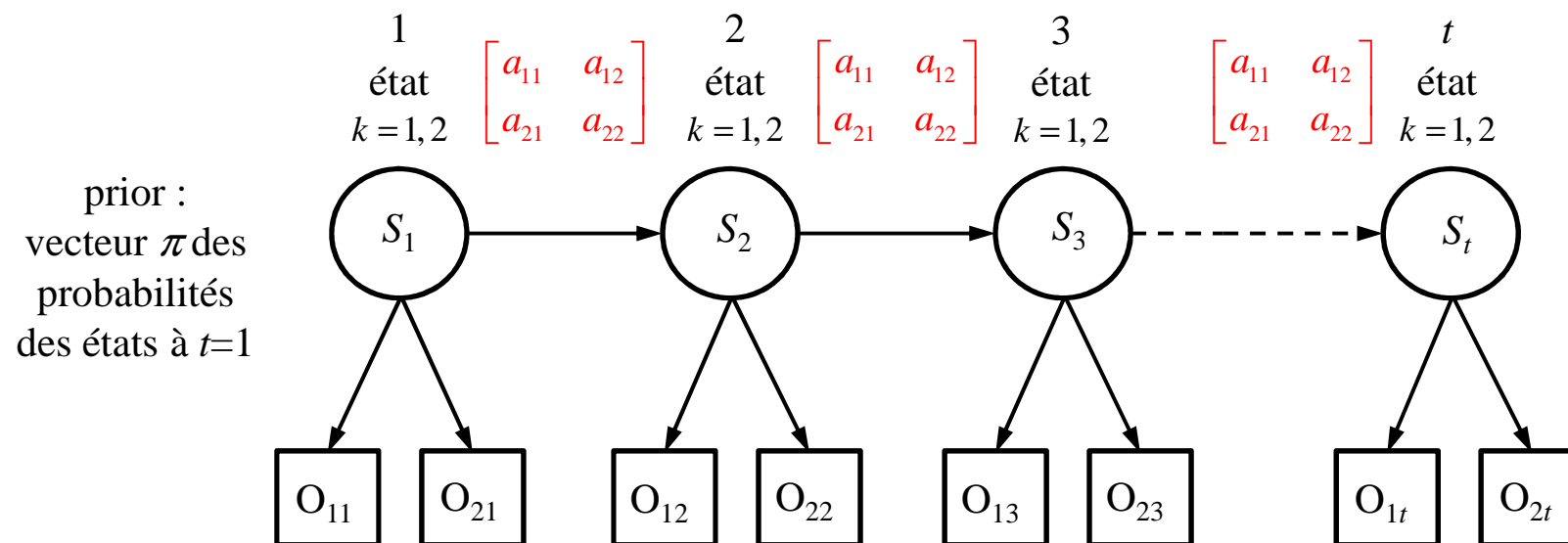


Modèle de Markov Caché (HMM) : 2 observations et 2 états possibles à chaque pas de temps



Chaîne de Markov: la probabilité de passage d'un état à un autre ne varie pas au cours du temps (HMM « homogène »)

Modèle de Markov Caché (HMM) : 2 observations et 2 états possibles à chaque pas de temps



Modèle de réponse : densités conditionnelles des observations associées à l'état latent k (gaussienne, multinomiale, etc.)

Modèle de Markov Caché (HMM)

HMM \approx modèle de mélange de lois de probabilité pour séries temporelles individuelles.

- ▶ Espace d'états fini S : la variable d'état S_t peut prendre de 1 à k valeurs possibles.
- ▶ Modèle de transition A : matrice de probabilités de transition a_{ij} de l'état S_i à l'état S_j (pour chaque S_i à l'instant t , $\sum_j a_{ij} = 1$).
- ▶ Modèle(s) de réponse B_t : distributions des observations sachant l'état latent.
- ▶ Ensemble des probabilités initiales des états π_t , avec $\sum_{i=1}^k \pi_i = 1$.

Modèle de Markov Caché (HMM)

► Estimation des paramètres

Quel est le HMM le plus probable d'avoir donné lieu à l'ensemble des observations $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_t$?

Maximiser la vraisemblance de $P(\mathbf{O}_t | S_t, \mathbf{A}, \mathbf{B})$: problème à données incomplètes résolu par optimisation numérique

Modèle de Markov Caché (HMM)

► Algorithme de restauration-maximisation

Initialisation de l'algorithme ($t=1$);

1) étape de restauration : remplacement des valeurs manquantes* sachant la loi des observations pour les vraies valeurs – inconnues – des paramètres du modèle;

2) étape de maximisation : estimation des paramètres par maximisation de la vraisemblance complétée sachant les observations et les états.

* Espérance conditionnelle pour l'algorithme *avant-arrière* de Baum (Espérance-Maximisation), valeur la plus probable pour l'algorithme de Viterbi (Maximum A Posteriori), etc.

HMM appliqué aux données précédentes

- ▶ *Dependent Mixture Model* (depmixS4; Visser & Speekenbrink, 2008, 2010)

Modèle défini par le nombre d'états, l'état initial, le modèle de transition et les fonctions de distribution des réponses (gaussienne pour le log(TR), binomiale pour la précision).

Sélection **pour chaque participant** du meilleur modèle (plus petit BIC) parmi les modèles suivants:

$m_{1,i}$: 1 état latent, HMM homogène

$m_{2,i}$: 2 états latents, HMM homogène

$m_{3,i}$: 2 états latents, HMM homogène, type d'item \rightarrow log(TR)

$m_{4,i}$: 2 états latents, type d'item \rightarrow transition

$m_{5,i}$: 2 états latents, type d'item \rightarrow log(TR), type d'item \rightarrow transition

Participant 13 : HMM homogène

id 13						
HMM	états	prior	log(TR)	PR	transition	BIC
m ₁	1	état 1	~1	~1		160
m ₂	2	état 1	~1	~1		147
m ₃	2	état 1	~1 + item	~1		151,2
m ₄	2	état 1	~1	~1	~item	157,6
m ₅	2	état 1	~1 + item	~1	~item	160,1

► Modèle 2

État 1 optimal : log(TR)=6.52, SD=0.23, PR=0.97%

État 2 « relaxé » : log(TR)=7.27, SD=0.29, PR=0.97%

état1 état2

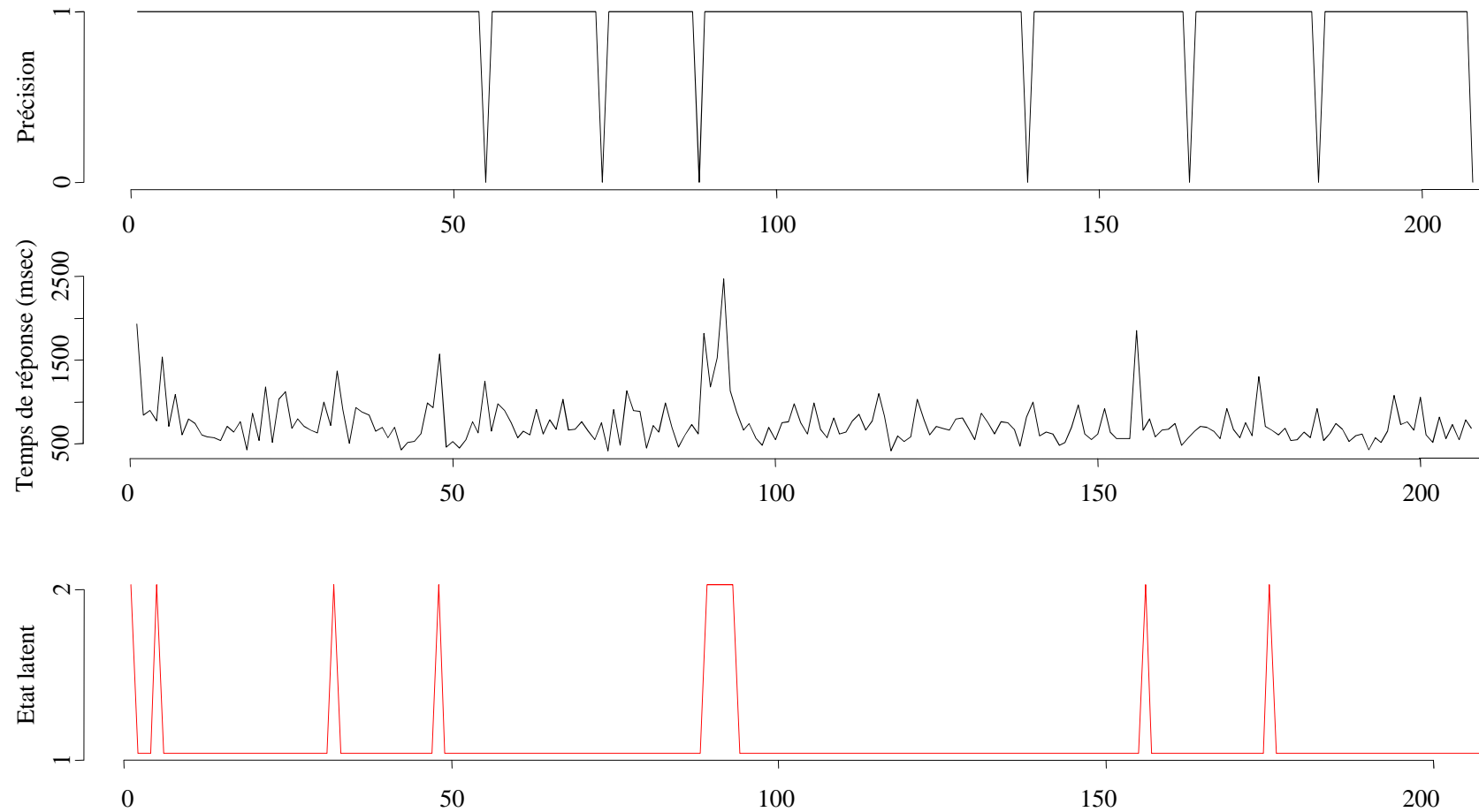
0.96	0.04
0.63	0.37

Participant 13 : HMM homogène

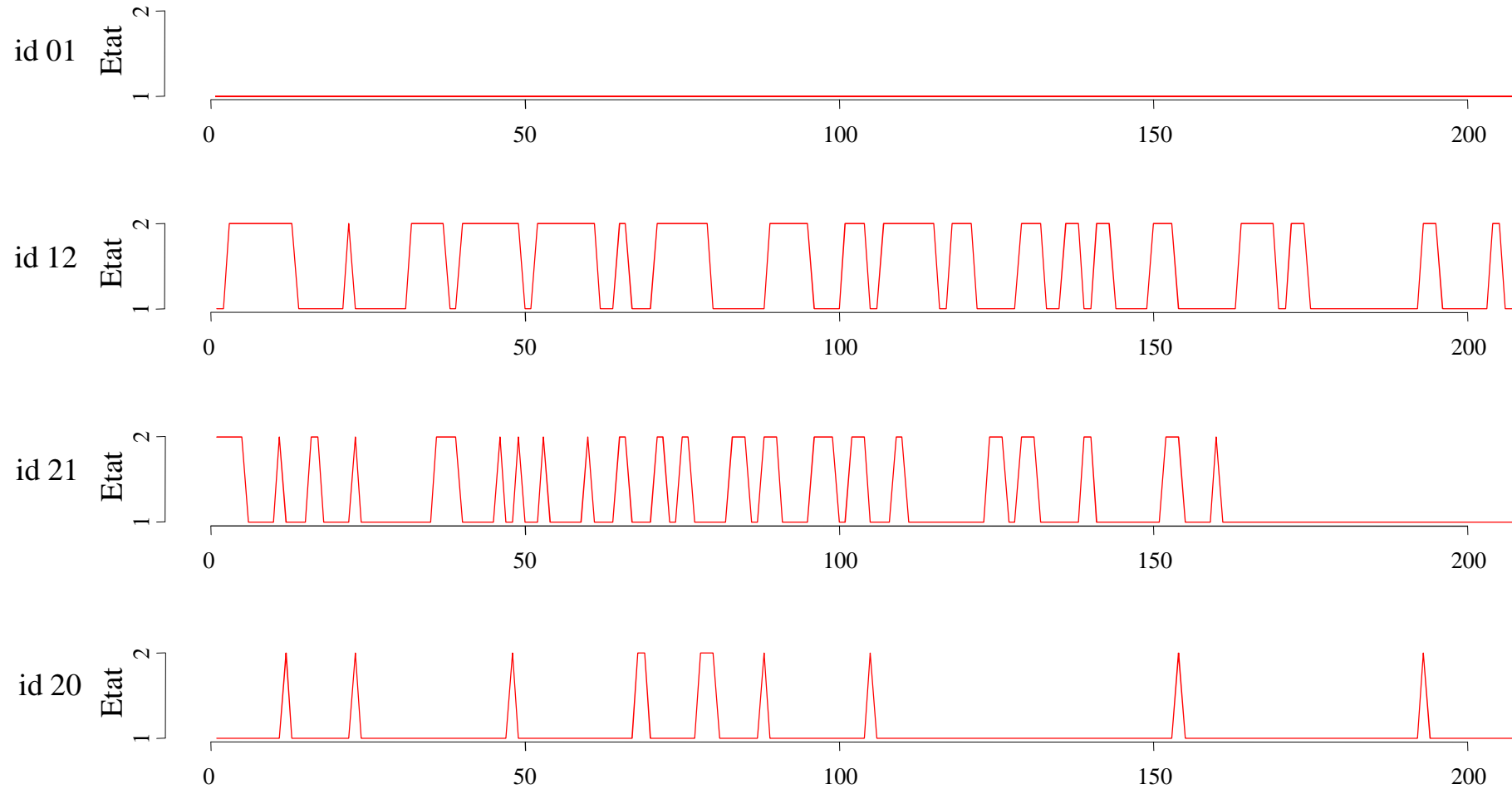
- Probabilités d'être dans l'état S_i ($i=1,2$) au temps t sachant la séquence d'observations \mathbf{O} et le modèle.

item	Observations		Posterior		état
	PR13	TR13	π_1	π_2	
...
83	1	980	0,955	0,045	1
84	1	691	0,999	0,001	1
85	1	461	1,000	0,000	1
86	1	601	1,000	0,000	1
87	1	727	0,998	0,002	1
88	0	611	1,000	0,000	1
89	1	1822	0,004	0,996	2
90	1	1179	0,129	0,871	2
91	1	1522	0,004	0,996	2
92	1	2475	0,000	1,000	2
93	1	1133	0,198	0,802	2
94	1	878	0,830	0,170	1
95	1	664	0,998	0,002	1
96	1	740	0,998	0,002	1
97	1	565	1,000	0,000	1
98	1	476	1,000	0,000	1
...

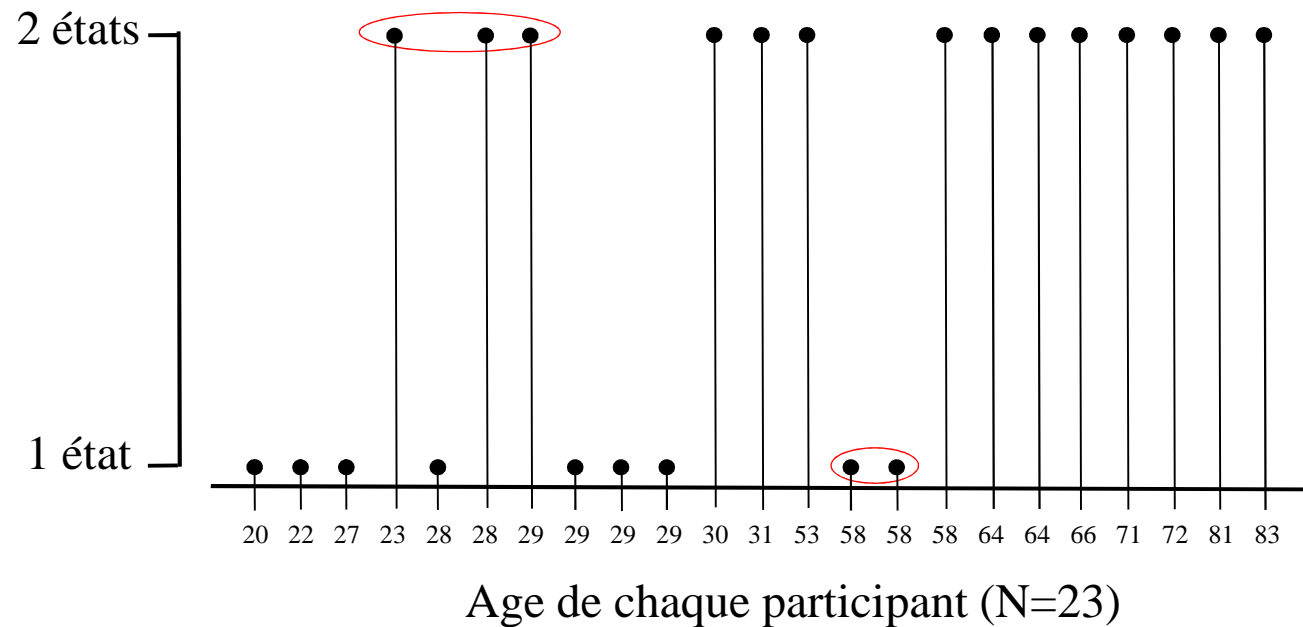
Participant 13 : HMM homogène



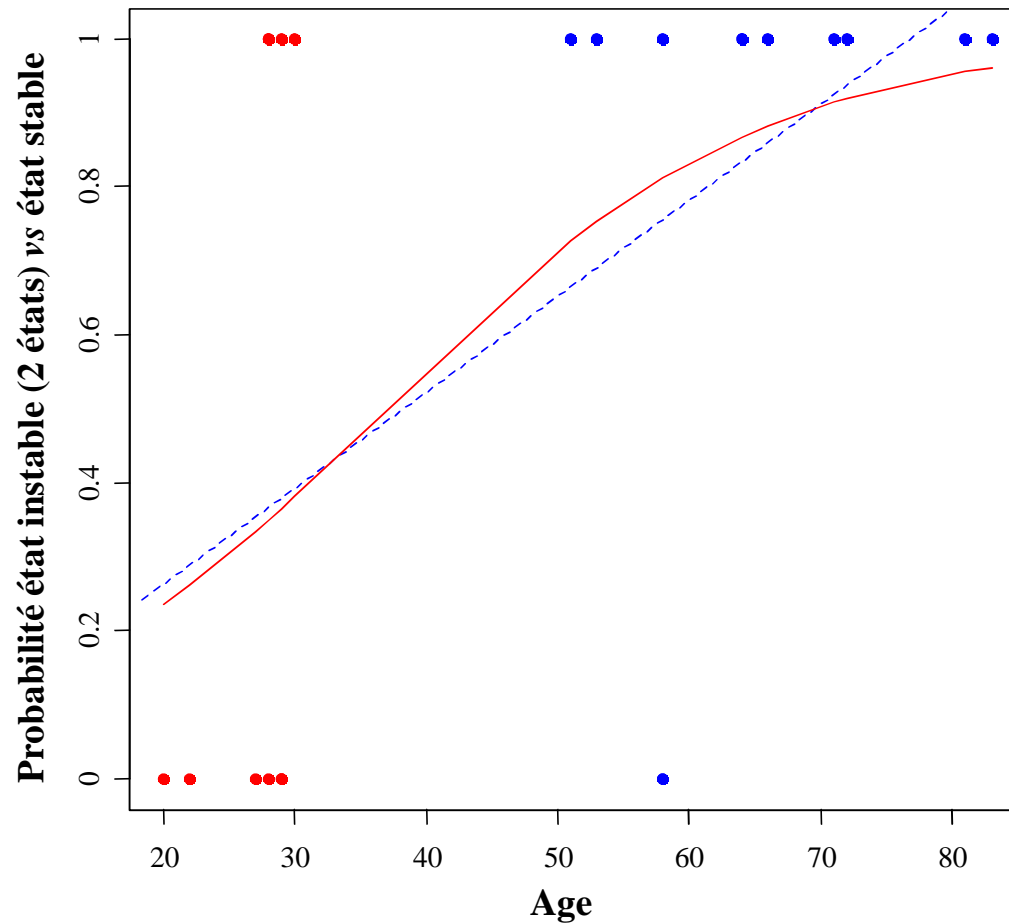
Autres participants



Stabilité / instabilité de l'état latent de contrôle attentionnel



Régression logistique de l'état (stable/instable) sur l'âge



$$\beta = 0.069 \quad \text{S.E.} = 0.030 \quad p < .05$$

Age: +1 odds=1.07

Age: +10 odds=1.99

Résultats individuels (extraits)

		Modèle de mesure	Modèle de transition
id 02 (22 ans)	état stable	$\ln(TR) = 6.16$ (ET = 0.23) PR = 0.98	
id 15 (58 ans)	état stable	$\ln(TR) = 6.68$ (ET = 0.26) PR = 0.99	
id 07 (29 ans)	état 1 ↓ ↑ état 2	$\ln(TR) = 6.21$ (ET = 0.14), $\beta_{I \text{ vs } C} = 0.15$ PR = 0.99 $\ln(TR) = 6.68$ (ET = 0.32), $\beta_{I \text{ vs } C} = 0.14$ PR = 0.83	état 1 état 2 état 1 $\begin{bmatrix} 0.14 & 0.86 \\ 0.81 & 0.19 \end{bmatrix}$ état 2
id 21 (72 ans)	état 1 ↓ ↑ état 2	$\ln(TR) = 6.52$ (ET = 0.19), $\beta_{I \text{ vs } C} = 0.27$ PR = 0.98 $\ln(TR) = 6.99$ (ET = 0.31), $\beta_{I \text{ vs } C} = 0.31$ PR = 0.98	état 1 état 2 état 1 $\begin{bmatrix} 0.17 & 0.83 \\ 0.59 & 0.41 \end{bmatrix}$ état 2

Modèles linéaires mixtes appliqués aux ln(TR)

Observations correspondant à un état optimal: 20,32 % d'observations manquantes.

Paramètres (REML)	Mod. 1	Mod. 2	Mod. 3	Mod. 4	Mod. 5
Effets fixes					
Constante	6,709 (0,084)	6,654 (0,084)	6,653 (0,086)	6,054 (0,166)	6,085 (0,175)
Item		0,110 (0,009)	0,113 (0,019)	0,113 (0,019)	0,089 ((0,049)
Age				0,013 (0,003)	0,012 (0,003)
Item x Age					0,001 (0,001)
Effets aléatoires					
Variance de la constante	0,1626	0,1626	0,1681	0,1099	0,1104
Variance de la pente Item			0,0068	0,0068	0,0072
<i>Variance résiduelle</i>	0,0654	0,0624	0,0606	0,0606	0,0606
BIC	586	422	368	374	394

Anova(m3,m4)

$\chi^2=12,7$, ddl=1, p<.001

► Résultats comparables à ceux obtenus avec la procédure classique.

Fiabilité des observations retenues

Estimation dans le cadre des modèles linéaires mixtes du coefficient de corrélation intra-classe (ICC) et de l'intervalle de confiance bootstrappé (basé sur la normalité des données).

- ▶ Procédure d'exclusion classique des TR extrêmes et contaminés

$$\text{ICC} = 0,522 \quad \text{biais} = 0,010 \quad \text{S.E.} = 0,015$$

$$\text{Intervalle de confiance 95\%} \quad [0,479 - 0,543]$$

- ▶ Procédure de sélection des TR « émis dans un état optimal »

$$\text{ICC} = 0,658 \quad \text{biais} = 0,006 \quad \text{S.E.} = 0,009$$

$$\text{Intervalle de confiance 95\%} \quad [0,633 - 0,672]$$

En conclusion

- ▶ Intérêt de HMM dans une approche « ascendante » : du niveau individuel au niveau différentiel (Molenaar, 2004).

Modélisation paramétrique très flexible (introduction de covariables dynamiques, nombreux modèles de réponse possibles GAMLSS; Stasinopoulos *et al.*, 2006) permettant :

- d'identifier des différences qualitatives dans la variabilité intra-individuelle.
- de renseigner la validité du regroupement de certaines réponses (même état) ou de certains individus (mêmes caractéristiques temporelles du processus latent).

En conclusion

- ▶ Approche applicable à de nombreuses situations expérimentales chronométrées mais nécessitant un nombre important d'essais.

Par exemple, sur un plus large échantillon : situation d'alternance permettant de distinguer le coût local d'alternance du coût de difficulté (I **II** C **C** C **III**).

Références

Molenaar, P.C.M. (2004). A manifesto on Psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2, 201-218.

Visser, I., Raijmakers, M.E.J., & Van der Maas, H.L.J. (2009). Hidden Markov Models for individual time series. In J. Valsiner, P.C.M. Molenaar, M. Lyra, & N. Chaudhary (Eds.), *Dynamic Process Methodology in the Social and Developmental Sciences* (pp. 269-289). Springer-Verlag, New York.

Visser, I., & Speekenbrink, M. (2010). DepmixS4: An R package for Hidden Markov Models. *Journal of Statistical Software*, 36(7),1-21.

Visser, I., Jansen, B., & Speekenbrink, M. (2010). A framework for discrete change. In P.C. Molenaar & K.M. Newell (Eds.), *Individual pathways of change: Statistical models for analyzing learning and development* (pp. 109-123). Washington, DC: American Psychological Association.

Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. (2011). Aging and task switching: A meta-analysis. *Psychology and Aging*, 26, 15-20.

West, R., & Travers, S. (2008). Differential effects of aging on processes underlying task switching. *Brain and Cognition*, 68, 67-80.