2014 PINGAULT et al. SUPPLEMENT Psychiatry Research Inattention Graduation

Online supplement

Table of contents

Further explanations on the new trajectory estimation approach

When clustering longitudinal data, two types of information are relevant: on one side, the evolution of trajectories; on the other side, the mean level of trajectories. To our knowledge, the existing partitioning methods do not distinguish between these two types of information. Therefore, the means of the individual trajectories strongly influence clustering and, thus, the resulting trajectories. As a consequence, the average levels of trajectories are strongly associated to the trajectories.

For a better understanding of these issues, we give the example of the clustering of student grades. Students whose mean is elevated often have repeated elevated grades across all assessment times and those with low means have repeated low grades. In this situation, classification algorithms group participants whose means are close without really taking into account the evolution of grades. The "good progressing" and "good regressing" will be classified in the same cluster, while the "weak progressing" and the "weak regressing" will be classified in another cluster (see Figure 1.a).

In a developmental perspective, the form of the trajectory (i.e. the evolution) conveys important information, independent of mean levels. In our example, it may be important to identify the clusters "Progressing vs. Regressing" than "Good vs. Weak" (see figure 1.b).

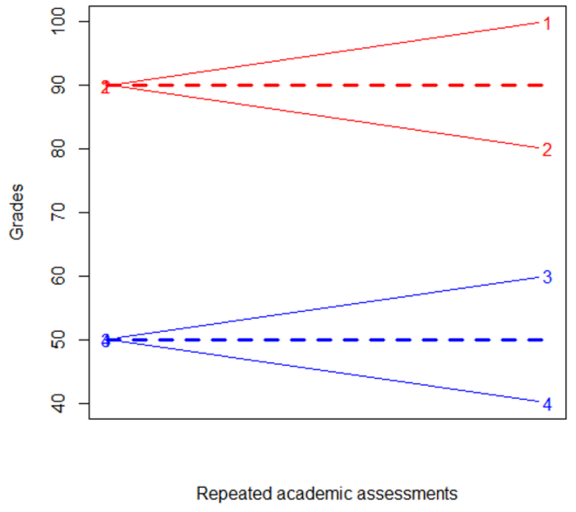


Fig1.a: Classical partition

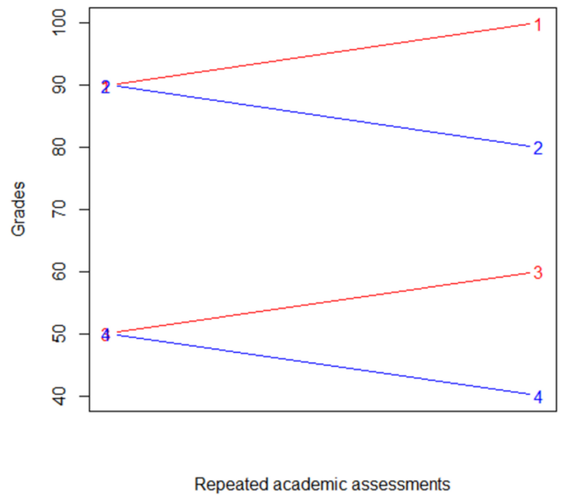
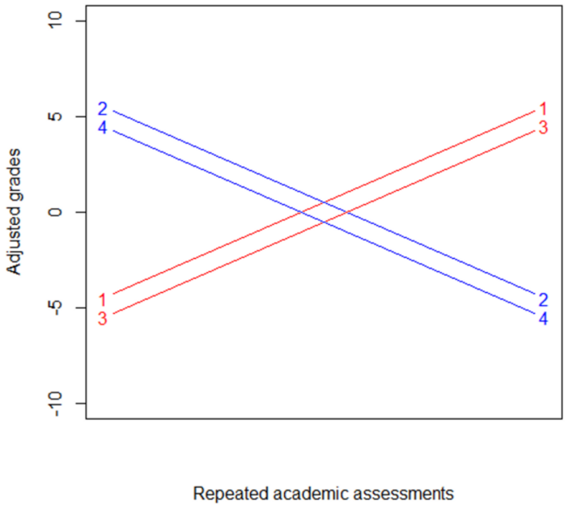


Fig 1.b: Alternative partition

Figure 1.a: classical algorithms will cluster students 1 and 2 in one group and students 3 and 4 in a second group. The resulting clusters will be “good students” (in red) and “weak students” (in blue). The resulting trajectories (dots) will be parallel and will not grasp the distinction between progressing and regressing students. Figure 1.b: it may be more relevant to cluster students 1 and 3 (red) in a “progressing” trajectory and students 2 and 4 in a “regressing” trajectory (blue).

One way to achieve this result is to work on trajectories *adjusted on their means*: for individual i, let y_{ij} be the value of its trajectory at time j with $1 \leq j \leq t$ (t being the number of repeated assessments). The sequence of measures $(y_{i1}, y_{i2}, ..., y_{it})$ is the trajectory of individual i and is noted $y_{i.}$. *To adjust a trajectory* $y_{i.}$ *on its mean* is to transform it into a trajectory $y'_{i.}$ by subtracting its $mean(y_{i.})$ to each of its values : $y'_{ij} = y_{ij} - mean(yi)$. Adjusted trajectories of our previous example (student grades) are shown in Figure 2.

The adjusted trajectories all have a mean of zero and, thus, are bound to cross. Partitioning such an adjusted population will focus on the group of participants with similar trends regardless of their level (those who progress together, those who regress together) which is our purpose.



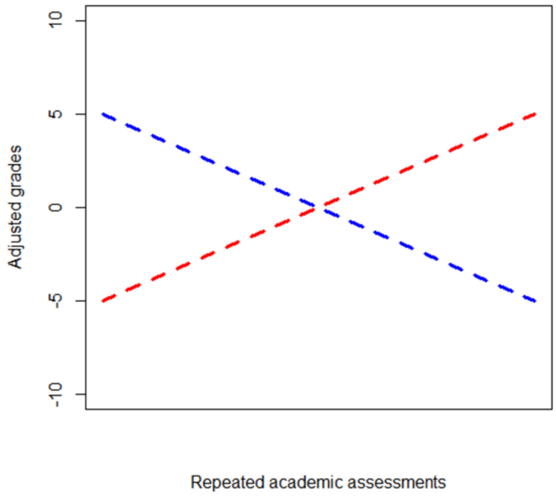


Figure 2.a : Adjusted trajectories   Figure 2.b : Adjusted resulting trajectories

Figure 2.a presents students' trajectories after adjustment. Students 1 and 3 will be clustered in one group; students 2 and 4 will be clustered in another group. Therefore, two different typical trajectories (progressing in red and regressing in blue) are detected. They are presented in figure 2.b.

Adjusting on the mean before partitioning enables a focus on the form of the trajectories. A consequence of this approach is that the resulting groups are less associated to the mean levels than in the case of a classical clustering. However, the two variables – resulting trajectories and mean levels – will not necessarily be completely independent. For example, let us consider the case where 40 good students progress and 20 good students regress while 10 weak students progress and 30 weak students regress. Classical partitioning techniques will identify two groups, "good" and "weak". The link between the groups and the average will be very high (Table 1.a). Partitioning after adjustment on the mean would identify two groups, "progressing" and "regressing." A total of 40 good students and 20 weak students will be in the group “progressing”, while 10 good students and 30 weak students will be in the group “Regressing”. There is still an association between the groups and the means, but this association is weaker than in the classical case (Table 1.b).

		Clusters	
		Good	**Weak**
Mean	**Good**	60	0
	Weak	0	40

Table 1.a

		Clusters	
		Progressing	**Regressing**
Mean	**Good**	40	20
	Weak	10	30

Table 1.b

Table 1: association between the mean trajectory and the groups obtained after partitioning. In the classical case (1.a, left), the association between variables is very strong. After adjusting on the mean (1.b, right), it is less strong, but may still exist.

From a statistical point of view, the adjustment on the mean before clustering allow to include simultaneously the mean of the trajectories and the clusters in the same model, which is not possible in the classical case.

The same phenomenon, described in Table 1, is observed for our data on inattention. Table 2 presents the mean levels of inattention across the 7 years for the adjusted trajectories. As can be seen, the mean is significantly higher for the increasing as well as the fluctuating trajectories. In particular, the difference in averaged inattention levels between the increasing and stable trajectories is equal to 0.79 standard deviation (the standard deviation for inattention across the seven years being equal to 1.88, the mean being 2.26). However, the association between trajectories and mean levels is much stronger in the case of classical trajectories. For ease of comparison, we selected a three trajectory solution on the same sample, estimated with the normal Kml algorithm, resulting in three parallel trajectories: low, intermediate, high. The difference between the low and the high trajectories in the classical case is this time equal to 2.4 standard deviations of inattention. This important reduction in the association between trajectories and mean level can be seen in the F value of the one-way anova in each case.

In addition, it should be noted that the mean levels of inattention are higher for the increasing but also the fluctuating trajectories, compared to the stable trajectory. However, only the increasing trajectory makes a significant multivariate contribution to failure to graduate from high school (OR: 1.76, see manuscript) whereas the fluctuating trajectory makes no contribution (OR: 1.02). This is coherent with the interpretation that the effect of the increasing trajectory is due to developmental considerations rather than higher mean levels of inattention (in addition to the fact that we controlled in the model for mean levels).

Table 2: Association between Trajectories and Mean Levels of Inattention

Adjusted trajectories		Classical trajectories	
Group	Mean	Group	Mean
Increasing	3.08	High	5.19
Fluctuating	2.77	Intermediate	2.82
Stable	1.60	Low	0.69
Anova: F value = **147** (p <.001)		Anova: F value = **5034** (p <.001)	

Information on the *Family Socioeconomic Adversity Index*

The index was based on information collected at the start of the study and was created by averaging the following indices: 1) family structure (intact or not intact), 2) parents' levels of education, 3) parents' occupational status (Blishen et al., 1987) and 4) parents' age at birth of the first child. Families at or below the 30th percentile for each of these indices (or a non intact family) were coded as having 1 adversity point. The final score ranges from 0 to 1.

Information on missing data

The information regarding high-school graduation was extracted from official records and present for all participants. All participants had at least one measurement of inattention and were kept in the trajectory estimation (of note is that 98.1% of participants had 3 or more assessments available). A total of 104 participants (5.2%) had missing data on the adversity index so that the final model was estimated on the remaining participants, i.e. 1896.

Choice of the number of trajectories and sensitivity analyses

The choice of the number of trajectories is one of the more controversial aspects in trajectory estimation (Bauer, 2007). Among criteria used in the literature are: (1) convergence issues as the number of trajectories rises; (2) taxometric theory (e.g. ADHD subtypes); (3) previous research; (4) the relative fit of the model; (5) usefulness of classes.

(1) Contrary to parametric approaches, non-parametric approaches as the one used in the present study do not encounter convergence issues.

(2) In our case, the absence of developmental taxometric theories regarding the development of inattention prevented precise hypotheses concerning the number of trajectories.

(3) We know from previous research that inattention tends to be stable or rise (Larsson et al., 2011; Willcutt et al., 2012), which leads to a greater likelihood of rising inattention for at least a subgroup of children (contrary to what is reported in the literature for other behaviours like physical aggression (Nagin and Tremblay, 1999; Broidy et al., 2003).

(4) The fit of the model is a relevant criterion in our case. However, most of the literature on developmental trajectories has relied solely on the Bayesian Information Criterion (BIC), which tends to overestimate the number of trajectories (Nagin, 2005). When several fit indices are used, contradictions usually arise (Larsson et al., 2011). We used two indices, Calinski-Harabatz (Genolini and Falissard, 2010, 2011) and the BIC. We also examined the global average of post-probabilities (the post-probability being the probability for each participant to effectively belong to the trajectory he/her was classified in). As in other studies(Larsson et al., 2011), the fit indices were contradictory, with Calinski-Harabatz indicating 2 trajectories as the best solution, while the BIC continued to improve up to 5 trajectories (See Table 3 below). Of note is that the improvement in the BIC between the 4 and 5 trajectory solutions was minimal. Therefore, based on the two statistical criteria, a solution between 2 and 4 seemed indicated. The average post-probability was.85 for the 2 trajectory solution and .79 for the 3 trajectory solution. It continued to decrease thereafter with .76 for 4 trajectories.

Table 3: Fit indices & global post-probability for each model

Trajectories	2	3	4	5	6
Calinski-Harabatz	417.3	358.0	308.3	282.1	253.4
BIC	-41233.1	-41165.4	-41111.3	-41009.9	-41024.9
Post-probability	0.848	0.789	0.760	0.737	0.733

Note. A higher value for the Calinski-Harabatz criterion indicates a better fit. A value closer to zero for the BIC indicates a better fit.

(5) The usefulness of classes was defined in van Lier et al. (2007) (p.603) as “the subjective interpretation of the developmental course of the trajectories and the number of children in each class”. The two and three trajectory solutions gave interpretable results with a stable trajectory and a rising one in both cases, and, in the case of the three trajectories solution, a fluctuating one. Solutions with further trajectories added to the noise with crossing trajectories without a definite developmental course.

Finally, as mentioned above, the statistical criteria indicated a range from 2 to 4 trajectories as adequate solutions. However, the global post-probability was already well below

.80 for the 4 trajectory solution. Furthermore, the 3 trajectory solution appeared as the most useful solution as it yielded an interpretable result, in coherence with previous research, i.e. a steady increasing trajectory accompanied by a stable and a fluctuating trajectory. We therefore decided to select this solution. In addition, in order to test whether the results were sensitive to the chosen number of trajectories, we repeated the analyses with a two trajectory solutions: the rising trajectory also predicted graduation failure independently of mean levels of inattention, adversity and sex.

Some authors have also considered the construct validation i.e. the associations between the trajectories and other variables such as predictors or outcomes (Bauer, 2007; van Lier et al., 2007). In the present study, the predictive value of the trajectories regarding high-school graduation can be interpreted as an evidence of construct validity.

Additional information on the trajectories

Table 4. Additional information on the trajectories.

	Stable	Fluctuating	Rising
N (%)	1003 (50.1)	514 (25.7)	483 (24.1)
Males (%)	47.0	46.7	60.0
Graduation failure (%)	23.2	34.6	48.9
Graduation failure in males (%)	32.3	42.9	53.8
Graduation failure in females (%)	15.2	27.4	41.5
Adversity index score (M & SD)	0.25 (0.24)	0.30 (0.25)	0.31 (0.26)

Note. Graduation failure percentages are not adjusted for confounding variables in Table 4. Adjusted failure rates and risk ratios are provided in the manuscript.

Table 5. Absolute levels of inattention at each age in each trajectories

		6 years	7 years	8 years	9 years	10 years	11 years	12 years
Stable	Mean	2.34	1.86	1.45	1.09	1.26	1.47	1.32
	SD	2.48	2.34	2.02	1.73	1.88	2.08	2.02
Fluctuating	Mean	1.38	3.61	4.02	3.39	3.86	2.10	1.50
	SD	1.78	2.52	2.42	2.41	2.39	2.09	1.79
Rising	Mean	1.28	2.06	2.38	3.09	3.87	4.33	4.51
	SD	1.63	2.21	2.22	2.55	2.49	2.30	2.34

Note. The three trajectories (stable, fluctuating, rising) were computed based on adjusted means as detailed in the first section of this document. The figure presented in the manuscript shows the adjusted trajectories with the adjusted inattention scores. For each of these trajectories, the observed levels of inattention (prior to adjustment) are provided in Table 5. It shows, for example, that the level of inattention at 6 years is lower for the rising trajectory than for the two others but higher at 12 years that at any other time point for any trajectory.

References (online supplement)

Bauer, D.J., 2007. Observations on the Use of Growth Mixture Models in Psychological Research. Multivariate Behavioral Research 42, 757–786.

Blishen, B.R., Carroll, W.K., Moore, C., 1987. The 1982 socioeconomic index for occupations in Canada. Canadian Review of Sociology and Anthropology 24, 465–488.

Broidy, L.M., Nagin, D.S., Tremblay, R.E., Bates, J.E., Brame, B., Dodge, K.A., Fergusson, D., Horwood, J.L., Loeber, R., Laird, R., Lynam, D.R., Moffitt, T.E., Pettit, G.S., Vitaro, F., 2003. Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. Developmental Psychology 39, 222–245.

Genolini, C., Falissard, B., 2010. KmL: K-means for longitudinal data. Computational Statistics 25, 317–328.

Genolini, C., Falissard, B., 2011. KmL: A package to cluster longitudinal data. Computer Methods and Programs in Biomedicine 104, e112–121.

Larsson, H., Dilshad, R., Lichtenstein, P., Barker, E.D., 2011. Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. Journal of Child Psychology and Psychiatry 52, 954–963.

Nagin, D., 2005. Group-Based Modeling of Development. Harvard University Press.

Nagin, D.S., Tremblay, R.E., 1999. Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development 70, 1181–1196.

Van Lier, P.A.C., van der Ende, J., Koot, H.M., Verhulst, F.C., 2007. Which better predicts conduct problems? The relationship of trajectories of conduct problems with ODD and ADHD symptoms from childhood into adolescence. Journal of Child Psychology and Psychiatry 48, 601–608.

Willcutt, E.G., Nigg, J.T., Pennington, B.F., Solanto, M.V., Rohde, L.A., Tannock, R., Loo, S.K., Carlson, C.L., McBurnett, K., Lahey, B.B., 2012. Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. Journal of Abnormal Psychology 121, 991–1010.