ORIGINAL ARTICLE

Assessing the structure of the CAST (Cannabis Abuse Screening Test) in 13 European countries using multigroup analyses

Stéphane Legleye^{1,2} | Aida Eslami³ | Stéphanie Bougeard⁴

¹Ined, Paris, France

²CESP, Fac. de médecine - Univ. Paris-Sud,
Fac. de médecine - UVSQ, INSERM, Université
Paris-Saclay, 94805, Villejuif, France

³University of British Columbia, Canada

⁴Anses, Ploufragan-Plouzane, Ploufragan,
France

Correspondence

Stéphane Legleye, Ined, Survey and Sampling,
133 Bd Davout, Paris, 75020, France.
Email: stephane.legleye@ined.fr

Abstract

Our aims are to describe and explain the structure of the Cannabis Abuse Screening Test (CAST) across countries. Standard statistical analyses fail to describe and explain several variables simultaneously while taking account of the group structure of individuals. The 2011 European School Survey Project on Alcohol and other Drugs (ESPAD): 5204 last-year cannabis users aged 15–16 from 13 European countries. Multigroup principal component analysis (mgPCA) and multigroup partial least squares (mgPLS). MgPCA shows that the CAST has a two-dimensional structure (frequency of use/problems and non-recreational use/dependency symptoms). All the countries present a good concordance with the common structure, except Kosovo, Lichtenstein and Romania. MgPLS shows that three explanative variables (in a total of eight) are mainly related with the CAST (the frequencies of cannabis use in the last 12 months and in the last 30 days and the age at first cannabis use) while Kosovo, Lichtenstein and Romania also present specificities. The CAST structure appears stable in the 13 countries except for Kosovo, Lichtenstein and Romania that also show specific relationships between the CAST variables and their determinants.

KEYWORDS

cannabis problematic use, CAST, ESPAD survey, factorial structure, invariance, multigroup analyses, multigroup PCA, multigroup PLS

1 | INTRODUCTION

Many screening scales assessing cannabis-related problems have been developed and tested in recent years (Beck & Legleye, 2008) although few of them have been validated in Europe (Annaheim & Legleye, 2016; Piontek, Kraus, & Klempova, 2008). One of the most-used scale is the Cannabis Abuse Screening Test (CAST) (Legleye, Karila, Beck, & Reynaud, 2007). This test assesses various aspects of the cannabis consumption in the past 12 months. Originally designed for teenagers, this test has been adopted in 2007 in the European School Survey Project on Alcohol and other Drugs (ESPAD) (Hibell et al., 2009; Piontek, Kraus, & Pabst, 2009), a survey focussed on drug use across pupils of 36 European countries. Its psychometric properties have been assessed in representative samples of teenagers in France and Italy (Bastiani et al., 2013; Legleye, Piontek, & Kraus, 2011b; Legleye, Piontek, Kraus, Morand, & Falissard, 2013), among Hungarian students (Gyepesi et al., 2014) and among French adults (Legleye et al., 2015). Good internal, psychometric and screening properties were also found in small samples

of young adults in Spain (Cuenca-Royo et al., 2012; Fernandez-Artamendi, Fernández-Hermida, Muñoz-Fernández, Secades-Villa, & García-Fernández, 2012).

Although the CAST is widely used, its cross-cultural validity in the European context is not yet assessed. The patterns of cannabis use is known to be not uniform in Europe [European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), 2011]: UK or Czech Republic show very high prevalences in comparison with Scandinavian or southern countries such as Malta or Latvia. This is reflected through large differences in mean CAST scores across countries (Pabst, Kraus, & Piontek, 2012). Do these differences in means represent real differences in problematic use levels? To compare cultures, countries or subpopulations of different social groups through a standardized tool, researchers have to assess that the same concept is measured the same way everywhere (Matsumoto & Van de Vijver, 2011). This problem is not specific to cannabis: in alcohol research, it has been found that measurement tools endorse cultural and political preconceptions that can influence response, even in self-reported questionnaires (Room, 2006). This is even the

case with the Diagnostic Statistical Manual (DSM) and the International Classification of Diseases (ICD) whose purpose is to be cross-cultural gold standards (Rehm & Room, 2015; Room, 1998). The CAST was designed in France but the European countries differ greatly by their timing in the stage of the cannabis diffusion (Legleye et al., 2014), their socio-economic contexts, their legal systems and attitudes toward drug use, suggesting that perceptions of cannabis may vary accordingly (Piontek, Kraus, Bjarnason, Demetrovics, & Ramstedt, 2013; ter Bogt, Schmid, Gabhainn, Fotiou, & Vollebergh, 2006; ter Bogt et al., 2013).

In this article, our aims are two-fold. The first aim is to detect if the CAST has the same structure in all the countries under study. The second aim is to understand if the CAST is related with variables associated with the use and context of cannabis use in the same way in all the countries. Studying a phenomenon across countries (or more generally, groups) requires addressing the problem of multivariate datasets divided into groups of individuals (groups = countries). These data are found in the literature under different names: multilevel, hierarchical, clustered, nested or multi-group data. When a single outcome and a limited number of non-correlated explanatory variables are considered, this problem is usually solved with generalized estimating equations although they do not produce group-specific parameters (Fitzmaurice, Laird, & Ware, 2011) or with Generalized Linear Mixed Models to get population as well as group parameters (Goldstein, 2010; McCullagh & Nelder, 1989). But for more complex data with several outcomes and a large number of quasi-collinear explanatory variables, these standard methods fail. Principal component analysis (PCA) (Jolliffe, 2002) or partial least square (PLS) regression (Wold, 1966) are extensively-used tools to deal with many variables as they are based on dimensionality reduction. But they assume that individuals come from a single population and are independent with each other. This is not the case when individuals come from various known groups. The commonly-used solution consists in performing a multivariate approach such as PCA or PLS regression on each group separately. But this yields to a large number of parameters and entails a difficulty in interpreting the outcomes and in comparing results across the groups. Multigroup analyses were recently developed as generalizations of multivariate analyses to the case of multigroup setting (Eslami, 2013).

In this article, we choose to apply multigroup PCA (mgPCA) (Eslami, Qannari, Kohler, & Bougeard, 2013a) and multigroup PLS regression (mgPLS) (Eslami, Qannari, Kohler, & Bougeard, 2013b). mgPCA is applied to the first aim, studying the CAST structure and detecting its variations across countries. mgPLS is applied to the second one, explaining the CAST with explanatory variables and detecting whether the associations vary across countries. This article is organized in four sections. The CAST data and the two multigroup methods are introduced and presented in a practical way in the Material and Method section. Results from the multigroup methods applied to the CAST data are given in the Application section. The Discussion section focuses on several aspects of the originality of multigroup analyses applied to cross-cultural surveys and its comparison to other methods. Finally, conclusions and perspectives are given.

2 | MATERIAL AND METHODS

2.1 | Data

The ESPAD European survey is a pen and paper self-administered school survey which aims at collecting data on alcohol and other drugs following the same protocol in several countries every four years (www.espad.org). The CAST is an optional module of ESPAD since 2007 (Piontek et al., 2009) and was chosen by 13 countries in 2011 (Pabst et al., 2012). The 2011 database analysed consists of 5204 pupils aged 15–16 who reported having smoked cannabis in the last 12 months. They are originated from the following countries: Belgium ($n = 331$), Cyprus ($n = 177$), Czech Republic ($n = 1013$), France ($n = 723$), Germany ($n = 365$), Italy ($n = 617$), Kosovo ($n = 55$), Latvia ($n = 292$), Lichtenstein ($n = 52$), Poland ($n = 1113$), Romania ($n = 93$), Slovak Republic ($n = 246$) and Ukraine ($n = 127$). Data are available upon request.

2.2 | The CAST

The CAST is a screening tool aimed at identifying problematic cannabis use, that is patterns of cannabis use leading to negative consequences on a health or social level for the user himself or others (EMCDDA, 2009, p. 31). It consists of six questions (items) about the past 12 months: non-recreational use (CAST1 “Have you smoked cannabis before midday?”), CAST2 “Have you smoked cannabis when you were alone?”), memory disorders (CAST3 “Have you had memory problems when you smoked cannabis?”), reproaches from family or friends (CAST4 “Have friends or family members told you that you should reduce or stop your cannabis consumption?”), unsuccessful attempts to quit (CAST5 “Have you tried to reduce or stop your cannabis use without succeeding?”) and problems associated with cannabis consumption (CAST6 “Have you had problems because of your cannabis use (argument, fight, accident, poor results at school, ...)?”). All items are answered on a five-point scale (0 “never”, 1 “rarely”, 2 “from time to time”, 3 “fairly often”, 4 “very often”).

2.3 | Explanatory variables

Eight questions were selected to explain the cannabis consumption: Can12M and Can30D “frequency of cannabis use in the last 12 months and in the last 30 days” (0, 1–2, 3–5, 6–9, 10–19, 20–39, 40+ times), Alcohol12M “frequency of alcohol use in the last 12 months” (0, 1–2, 3–5, 6–9, 10–19, 20–39, 40+ times), DrunkL “frequency of drunkenness during life” (0, 1–2, 3–5, 6–9, 10–19, 20–39, 40+ times), Tobacco “number of tobacco cigarettes in the last 30 days” (0, < 1/week, < 1/day, 1–5/day, 6–10/day, 11–20/day, 21+/day), CanOnset “age at first cannabis use” (nine or less, 10, 11, 12, 13, 14, 15, 16+), CanFriends “proportion of friends smoking cannabis” (none, some, most, almost all, all) and CanRisks “perception of risks associated to regular cannabis use” (no risk, light risk, moderate risk, important risk).

2.4 | Multigroup principal component analysis (mgPCA)

The first aim is to describe the CAST dataset (5204 pupils $\times$ 6 variables), where individuals are sampled from 13 countries. The mgPCA (Eslami

et al., 2013a) takes into account this group structure and proposes index and graphical displays to interpret the common level that reflects the overall population, as well as the group specificities in comparison with the common structure. Consider a single dataset $\mathbf{X}$ involving P quantitative variables and N individuals a priori divided into M groups

$\mathbf{X}_m$ with $m = (1, \dots, M)$, containing N_m individuals such as : $\sum_{m=1}^M N_m = N$.

We assume that each dataset $\mathbf{X}_m$ of dimensions $(N_m \times P)$ is column-centred. The principle of mgPCA is to seek a common vector of loadings a to study the relationships between the P variables, these links being common to all the individuals. In addition, to better understand the group specificity in comparison with the common structure, the P variables may also be viewed through their group vector of loadings a_m . The graphical display in Figure 1 depicts all these elements.

Formally, this principle consists in seeking a vector of loading a common to all the groups tightly linked to the M vectors of group loadings $(a_1, \dots, a_M)$ so as to maximize the following criterion for the first

component ($h = 1$) (Eslami et al., 2013a) : $\sum_{m=1}^M \langle a_m^{(1)}, a^{(1)} \rangle^2$, where $\langle \cdot \rangle$

FIGURE 1 Graphical display of dataset $\mathbf{X}$ divided into $M = 3$ groups of individuals with its common loading a , group loadings $(a_1, \dots, a_M)$, common component t and group components $(t_1, \dots, t_M)$

stands for the scalar product, with the M group loadings $a_m^{(1)} = \mathbf{X}_m' t_m$ for $m = (1, \dots, M)$ and the norm constraints $\|t_m^{(1)}\| = \|a^{(1)}\| = 1$. The solution is unique and derived from a matrix eigenanalysis. The subsequent vectors of loadings are sought by considering the same maximization problem and adding orthogonality constraints. As for standard PCA, the choice of the optimal number of components h^{opt} to be interpreted is crucial. Two standard rules are simultaneously applied: the scree test based on the graph of eigenvalues against component number and the component interpretability. A detailed account of mgPCA is given in (Eslami, 2013).

2.5 | Multigroup partial least squares (mgPLS)

The second aim is to investigate the relationships between the multivariate explanatory dataset and the CAST dataset. The mgPLS (Eslami et al., 2013b) is based on a standard PLS method (Wold, 1966) but takes the group structure into account. It proposes indexes and graphical displays to interpret the common as well as the group levels. Consider an explanatory dataset $\mathbf{X}$ and a dependent one $\mathbf{Y}$ which respectively consist in the measurement of P and Q quantitative variables on the same N individuals partitioned into M groups a priori known. We assume that each dataset associated with group m for $m = (1, \dots, M)$ $\mathbf{X}_m$ of dimensions $(N_m \times P)$, respectively $\mathbf{Y}_m$ of dimensions $(N_m \times Q)$, is column-centred. Again, $\sum_{m=1}^M N_m = N$. The principle of mgPLS is to seek vectors of loadings a and b common to all the groups, respectively in explanatory ($\mathbf{X}$) and dependent ($\mathbf{Y}$) spaces, such as their associated components $t_m = \mathbf{X}_m a$ and $u_m = \mathbf{Y}_m b$ are tightly linked. The graphical display in Figure 2 depicts all these elements.

Formally, this principle consists in the maximization of the following

criterion for the first component ($h = 1$): $\sum_{m=1}^M \text{cov}(\mathbf{X}_m a^{(1)}, \mathbf{Y}_m b^{(1)})$,

where cov stands for covariance, with the norm constraints $\|a^{(1)}\| = \|b^{(1)}\| = 1$. The solution is unique and derived from a matrix eigenanalysis. To assess the group specificity in comparison with the common structure, the vectors $a_m^{(1)}$ and $b_m^{(1)}$ are defined as the specific group vectors of loadings associated with $\mathbf{X}_m$ and $\mathbf{Y}_m$ retrieved from $a_m^{(1)} = \mathbf{X}_m' u_m^{(1)} / \|\mathbf{X}_m' u_m^{(1)}\|$ and $b_m^{(1)} = \mathbf{Y}_m' t_m^{(1)} / \|\mathbf{Y}_m' t_m^{(1)}\|$. As

FIGURE 2 Graphical display of two datasets $\mathbf{X}$ and $\mathbf{Y}$ divided into $M = 3$ groups of individuals with their common loadings a and b , group loadings $(a_1, \dots, a_M)$ and $(b_1, \dots, b_M)$, common components t and u and group components $(t_1, \dots, t_M)$ and $(u_1, \dots, u_M)$

standard PLS, mgPLS seeks regression coefficients between $\mathbf{X}$ and $\mathbf{Y}$ that are common to all the groups on the basis of the components. To investigate other directions, subsequent loadings and components are sought by considering the same maximization problem and using the deflated $\mathbf{X}$ with respect to the first component. As for standard PLS, the choice of the optimal number of components h^{opt} to be used is pivotal. Three rules are simultaneously used: the scree test, the percentage of explained variances of the dependent and the explanatory datasets and the component interpretability. Details can be found in (Eslami et al., 2013b; Eslami, Qannari, Kohler, & Bougeard, 2014a).

2.6 | Interpretation levels and similarity indexes

Multigroup analyses may be interpreted at the common and the group level for variables and individuals. From our experience, the interpretation at the individual level is of little use while the interpretation at the variable level is of paramount interest. In addition, the interpretation of the groups in regard with the common structure is interesting to explore the data structure. (i) At the common level, the main aim of multigroup analyses is to highlight the relationships between the variables which are common to all the groups. This can easily be achieved by the graphical displays of the common loadings, i.e. $(a^{(1)}, \dots, a^{(h)})$ for mgPCA and $(a^{(1)}, \dots, a^{(h)})$ and $(b^{(1)}, \dots, b^{(h)})$ for mgPLS. (ii) At the group level, we are interested in understanding whether the relationships between the variables under study are the same for all the groups. This can be achieved by the comparison of the graphical displays of the group and the common loadings. For mgPCA and mgPLS, for several dimensions under study $h = (1, \dots, h^{\text{opt}})$, h^{opt} being the optimal number of dimensions, an overall similarity index is given by $S_m^{(h^{\text{opt}})} = (1/h^{\text{opt}}) \sum_{h=1}^{h^{\text{opt}}} \left| \cos(a^{(h)}, a_m^{(h)}) \right|$ for each group $m = (1, \dots, M)$. It varies between zero (total dissimilarity between common and group loadings) and one (perfect agreement up to dimension h^{opt}). For mgPLS, these indices are also processed for the comparison of the dependent loadings b and $(b_1, \dots, b_M)$.

2.7 | Data pre-treatment

A crucial point in multivariate analysis is the question of the data pre-treatment, i.e. data centring and scaling. Data centring means that the variable mean is subtracted from each corresponding variable value; consequently, each variable has a zero mean. Data scaling means that each variable value is divided by its corresponding variable standard deviation; consequently, each variable have a standard deviation of one. In multigroup analysis, the centring question is not a real one for the user as data are by default centred within each group. But the scaling question is more complex as it may concern overall data as well as group ones. Indeed, four ways of scaling can be chosen: variables can be globally scaled (overall scaling) and/or scaled within groups (group scaling). Overall scaling implies that variables have the same weight in the analysis. Contrariwise, variable weights depend on their respective variances. Group scaling implies that groups have the same weight in the analysis. Contrariwise, group weights depend on their respective inertia. Depending on the data and on the goals of the analysis, the user has to choose between these four different ways of scaling. The

most standard choice (and the easiest to interpret) is to apply an overall as well as a group scaling. We applied this solution.

2.8 | Multigroup analyses in practice

In multigroup analyses, variables must be preferably continuous or ordered but no specific distribution assumption are required. It is commonly accepted that in PCA and mgPCA, the number of individuals should be higher than the number of variables but no such assumption is stated for PLS and mgPLS. From a statistical point of view and from our experience, any group number can be chosen, this choice being only driven by the data features. Statistical procedures for mgPCA and mgPLS and associated interpretation tools are freely available through the "multigroup" package in R (Eslami, Qannari, Bougeard, & Sanchez, 2015). The R code used for this study is available online.

3 | APPLICATION

3.1 | Data description and pre-treatments

Mean distributions and associated standard deviations of the CAST variables are given in Table 1.

France has the highest frequency of use mean in the last 12 months (15.2), whereas Cyprus lies in eighth position (8.9), Latvia in 10th position (7.0) and Kosovo in the last one (2.5). The CAST is commonly used to screen for problematic use: a score is computed for each individual by adding the values of the six CAST items, ranging between 0 and 24 (Legleye et al., 2015; Legleye et al., 2013). ESPAD data show large differences in mean CAST scores (Pabst et al., 2012): Cyprus had the highest mean (2.5) followed by Italy (1.6) and Latvia (1.0). Large discrepancies between frequency of use means and CAST scores means across countries are thus observed. As the six CAST variables have different variances, data are globally centred and scaled to give the same weight to all of them. Thereafter, in order to give the same weight to all the countries in the statistical analysis, all the variables are centred and scaled by country.

3.2 | The CAST structure

mgPCA suggest to retain two components as they explain 62.5% of the overall inertia (Dim1 = 46.4%, Dim2 = 16.1%). Figure 3 (top left plot) depicts the mgPCA common loadings for the first two components.

The two variables related to the non-recreational use (CAST1, CAST2) are linked and quite independent from "reproaches from family or friends" (CAST4) and unsuccessful quite attempts (CAST5). As the CAST variable loadings on the first component are similar – ranging between 0.32 (CAST5) and 0.45 (CAST2), this component is interpreted as the frequency of problems and use, that is the cumulation of all items. This is a classical result in PCA, especially here because all items were designed to depict one aspect of the same concept of problematic use. As the loadings on the second component oppose CAST1 and CAST2 (−0.49, −0.38) on the one side and CAST4 and CAST5 on the other (0.36, 0.68), it can be interpreted as an opposition of non-recreational use and dependency symptoms. To compare

TABLE 1 Means and standard deviations of the frequency of cannabis use and of the CAST in 13 European countries

country	N	Means										Standard deviations											
		Frequency of use	CAST1	CAST2	CAST3	CAST4	CAST5	CAST6	CAST score	Frequency of use	CAST1	CAST2	CAST3	CAST4	CAST5	CAST6	CAST score						
Belgium	331	13.9	1.8	1.5	1.4	1.4	1.3	1.4	2.9	16.9	1.1	0.9	0.9	1.0	0.9	4.0							
Cyprus	177	8.9	2.1	2.2	2.2	2.1	2.1	1.9	6.6	13.5	1.4	1.4	1.4	1.5	1.5	7.1							
Czech R	1013	11.5	1.7	1.5	1.6	1.5	1.5	1.3	3.0	15.8	1.0	0.9	1.0	1.0	1.0	3.9							
France	723	15.2	1.8	1.4	1.5	1.3	1.3	1.2	2.5	17.2	1.1	0.9	0.9	0.8	0.9	3.4							
Germany	365	10.2	1.7	1.4	1.6	1.5	1.4	1.4	2.8	14.8	1.1	0.9	1.0	1.0	1.0	3.9							
Italy	617	14.1	2.0	1.8	1.6	1.6	1.6	1.4	4.0	16.9	1.2	1.2	1.1	1.2	1.2	4.4							
Kosovo	55	2.5	1.3	1.4	1.4	1.4	1.5	1.3	2.2	4.5	0.7	0.9	0.9	0.9	1.1	4.7							
Latvia	292	7.0	1.6	1.3	1.4	1.3	1.3	1.2	2.0	12.0	0.9	0.6	0.7	0.8	0.6	2.9							
Liecht	52	11.1	1.6	1.4	1.4	1.6	1.6	1.3	2.9	15.6	1.1	1.0	1.0	1.2	1.3	3.6							
Poland	1113	8.6	1.8	1.3	1.5	1.5	1.5	1.2	2.7	13.2	1.0	0.8	0.9	1.0	1.2	3.7							
Romania	93	3.4	1.6	1.5	1.4	1.7	1.5	1.3	2.9	6.7	1.0	0.8	0.8	1.4	1.1	3.8							
Slovak R	246	12.8	1.8	1.5	1.6	1.6	1.5	1.4	3.4	17.2	1.1	0.9	1.0	1.0	1.1	4.0							
Ukraine	127	6.3	1.6	1.3	1.5	1.6	1.7	1.4	3.1	10.2	0.9	0.8	1.0	1.0	1.2	4.0							
Global		11.1	1.8	1.5	1.5	1.5	1.5	1.3		15.4	1.1	1.0	1.0	1.0	1.1								
Maximum		Fra	Cyp	Cyp	Cyp	Cyp	Cyp	Cyp	Cyp	17.2	Cyp	Cyp	Cyp	Cyp	Cyp	Cyp							
Minimum		Kos	Kos	Lat	Kos	Fra	Fra	Lat	Lat	4.5	Kos	Lat	Lat	Fra	Lat	Lat							
Max/Min		6.0	1.6	1.8	1.6	1.6	1.6	1.6	3.3	3.9	2.1	2.3	1.9	1.8	1.8	2.4							

Note: Frequency of use = frequency of cannabis use in the last 12 months based on median values of the following ranges of frequencies (50 being retained for the upper category): 1–2, 3–5, 6–9, 10–19, 20–39, 40+. The CAST score is the sum of the six items of the CAST (range 0–24).

FIGURE 3 Graphical displays of mgPCA loadings: common loading plot and group loadings of Kosovo, Lichtenstein and Romania

each country with this common structure, similarity indexes are depicted in Figure 4 for these first two dimensions.

It shows that the CAST has a robust and common structure in most of the countries as similarity indexes are close to one. More precisely on the first dimension, all the countries are coherent with the common structure, the most deviant countries being Liechtenstein and Romania with a similarity index above 0.99. On the second dimension, Kosovo differs the most with a similarity index above 0.90. To understand these minor differences, the group loadings of these three countries are also illustrated in Figure 3. For Kosovo, the first component has a large inertia (79.0%, whereas only 6.3% for the second one) meaning that the CAST is highly unidimensional for this country. In addition, some relationships between CAST variables are

specific to this country: CAST2 (smoking when alone) and CAST6 (problems due to consumption) are highly correlated, whereas the correlation between CAST1 and CAST2 is lower. For Liechtenstein, the specific features mainly come from the strong links between CAST4 (reproaches) and CAST5 (unsuccessful quite attempts), CAST3 (memory problems) and CAST6 (problems), and finally CAST1 and CAST2. The specificity of Romania comes from the high correlations between CAST1, CAST2 and CAST3 and the poor contribution of CAST6.

These three deviating countries are those with the smallest samples. Excluding them does not change the eigenvalues and the common structure but leads to much higher similarity indexes: they are above 0.996 (Cyprus) on the first dimension and above 0.975 (Ukraine) on dimension 2.

FIGURE 4 Similarity index between common and group loadings from mgPCA

3.3 | Study of the relationships between the CAST and the context variables

The mgPLS aims at explaining the CAST with some elements of the context of drug use, both measured at the individual level. It may also find out determinants of CAST that may vary according to countries and may contribute to explain differences in CAST structures.

The six CAST variables represent now the dataset to explain (Y), whereas the eight explanatory variables (see Material and Methods) represent the explanatory one (X). Two dimensions are retained as they explain 48.7% of the total inertia (Dim1 = 37.1%, Dim2 = 11.6%). They account for 49.5% of the inertia of the explanatory dataset and 15.9% on the dependent one. The common loadings plot is given in Figure 5 (top left plot).

The six CAST variables are correlated with each other the same way as in mgPCA. The main result is that they are positively correlated with the frequencies of cannabis use in the last 12 months (Can12M) as well as in the last 30 days (Can30D) and negatively correlated with the age at first cannabis use (CanOnset). The links between the frequency of alcohol use (Alcohol12M) and the six CAST variables are mildly negative. It follows that the problematical use of cannabis (identified by high CAST scores) is mostly linked to the frequency of cannabis use and negatively but weakly linked to the frequency of alcohol use and to the precocity of onset. The proportion of friends

who smoke cannabis regularly (CanFriends) and the perception of the risks associated with cannabis use (CanRisks) only play a marginal role. The common relationships between explanatory and dependent variables can also be viewed through the mgPLS common regression coefficients between dependent and explanatory variables depicted in Table 2.

Table 2 confirms that the frequencies of cannabis use in the last 12 months (Can12M) and in the last 30 days (Can30D) are the two main explanatory variables. They present similar coefficients for each CAST variables with the highest values for CAST1 and CAST2 (above 0.2) and the lowest ones for CAST5 (around 0.05). This latter result suggests that “unsuccessful quit attempts” is somewhat apart in the CAST scale and is not well-explained by the frequency of use. The link between the age at onset (CanOnset) and the CAST variables is negative, reflecting the influence of the precocity of onset on the development of cannabis use, especially with the frequency of use (CAST1 and CAST2), memory problems (CAST3) and problems (CAST6). It is worth noting that the other context variables have a reduced impact on the CAST: the link with alcohol consumption (Alcohol12M) is low and negative for all CAST variables while the links with drunkenness (DrunkL) are low but positive. Also surprisingly, the proportion of friends that smoke cannabis (CanFriends) and the perception of the risk of regular cannabis smoking (CanRisks) show weak relationships with the CAST variables even with CAST1 (smoking cannabis before midday) and

FIGURE 5 Graphical displays of mgPLS loadings: common loading plot and group loadings of Kosovo, Lichtenstein and Romania. In red, the CAST variables. In black, the explanatory variables. CanOnset, age at cannabis first use (experimentation); CanRisks, perception of cannabis-related risks; Can12M, frequency of cannabis use in the last 12 months; Can30D, frequency of cannabis use in the last 30 days; CanFriends, proportion of friends that smoke cannabis regularly; Alcohol12M, frequency of alcohol use in the last 12 months; DrunkL, frequency of drunkenness episodes in life

TABLE 2 Regression coefficients of mgPLS of the CAST variables (Y) on consumption context (X) for an optimal model with three dimensions

	CAST1	CAST2	CAST3	CAST4	CAST5	CAST6
Can12M: Frequency of cannabis 12 months	0.22	0.20	0.14	0.12	0.06	0.12
Can30D: Frequency of cannabis 30 days	0.22	0.20	0.12	0.10	0.04	0.09
CanOnset: Age at first cannabis use	-0.13	-0.13	-0.10	-0.08	-0.06	-0.09
Tobacco: Tobacco use	0.07	0.06	0.08	0.08	0.08	0.10
Alcohol12M: Frequency of alcohol 12 months	-0.03	-0.06	-0.04	-0.05	-0.04	-0.04
DrunkL: Frequency of drunkenness during life	0.05	0.03	0.05	0.04	0.04	0.06
CanFriends: Proportion of friends smoking cannabis	0.09	0.06	0.04	0.03	0.00	0.03
CanRisks: Perception of cannabis-related risks	-0.10	-0.07	0.00	0.02	0.07	0.04

Note: In bold type the coefficients above 0.10 in absolute value.

CAST2 (smoking when alone). Nevertheless, these latter links are negative, showing that the teenagers who smoke cannabis before midday or alone perceive the regular use of cannabis as less risky than the others. All these results suggest that the problematical use of cannabis as measured by the CAST is a special pattern of use which is not clearly linked to licit drugs and with the influence of peers' consumption. Only some explanatory variables are strongly linked with the CAST, explaining why the total variance of **Y** explained by **X** is low.

The country specificity in regard with the common structure is given in Figure 6 for the explanatory space (left plot) and the dependent space (right plot).

In accordance with the latter mgPCA results (**Y** dataset), all the countries under study have similar relationships between the CAST variables in regard with the common structure with small differences for Kosovo (index =0.89) and Romania (0.68) on the first component and Kosovo (0.78), Romania (0.76) and Liechtenstein (0.50) on the second one. The relationships between the explanatory variables are similar with the common structure for the first component but differ for the second one for some countries: Italy (index =0.81), Slovak Republic (0.79), France (0.77), Cyprus (0.75), Poland (0.75) show good to moderate fit, whereas Ukraine (0.66), Latvia (0.58), Romania (0.51), Liechtenstein (0.38) and especially Kosovo (0.25) show poor fit.

To understand these differences, the group loadings of the three most deviating countries are given in Figure 5. For Kosovo (Figure 5, top right plot), it is worth noting that CAST6 (problems) and CAST4 (reproaches from the entourage) are highly correlated, whereas CAST5

(difficulty to reduce or stop) stands apart. Furthermore, CAST variables are positively linked to all the explanatory variables, even with drug use, except with the age at cannabis onset (CanOnset). CanOnset is linked with CAST5. In Kosovo, cannabis is smoked by people who perceive cannabis as dangerous (CanRisks), while it is the opposite in average. It follows that the problematical use of cannabis is associated with the frequency of cannabis use in the last 30 days and in the last 12 months (Can30D and Can12M) but also with the perception of high risk related to cannabis use (CanRisks) and recent onset (CanOnset) contrariwise to the common results. Kosovo is the country with the lowest CAST score and the lowest frequency of use (Table 1): if the proportion of frequent users is low, most users may consider that the regular use is risky, which is a common denial and coping mechanism in addiction (Becker, 1963; Peretti-Watel et al., 2007).

For Liechtenstein (Figure 5, bottom left plot), the proportion of explained variance in the first factorial plan is also relatively lower than in average (35.0% and 9.9% in the first and second dimension versus respectively 37.3% and 11.6%). The CAST variables are less correlated with each other than in the common structure. They are also poorly associated with alcohol use (Alcohol12M), drunkenness in life (DrunkL), age at cannabis onset (CanOnset) and the perception of cannabis harms (CanRisks). Liechtenstein is characterized by a relatively high CAST score average and a high frequency of use (Table 1). Finally for Romania (Figure 5, bottom right plot), the CAST variables are also less correlated with each other than in the common structure and are mostly linked to the frequency of use in the last 12 months (Can12M)

Note: Germany and Czech Republic are depicted as a single point "Ger / Cze" in the Y space.

FIGURE 6 Similarity index between common and group loadings from mgPLS in the explanatory space (**X**) and in the dependent one (**Y**) for the first two dimensions. Note: Germany and Czech Republic are depicted as a single point "Ger/Cze" in the Y space

and to a lesser extent to the proportion of friends who smoke cannabis (CanFriends). The link with drunkenness in life (DrunkL) is weak for all the CAST variables except for CAST3. The proportion of friends who smoke regularly cannabis (CanFriends) and the perception cannabis harms (CanRisks) play a small role, while the link with the precocity of onset is small but positive. In Romania, problematic cannabis use aspects assessed by the CAST are primarily linked to the frequency of use and are related with teenagers who do not report many drunkenness episodes in life (DrunkL) and a late cannabis onset (CanOnset). This may reflect a late timing in the diffusion of cannabis, which is supported by the fact that Romania has a very low frequency of use and a rather high CAST score average (Table 1).

As in mgPCA, a replication of the analysis excluding Romania, Liechtenstein and Kosovo that are the most deviating countries does not change the common structure but leads to much higher similarity indexes: in the X space, they are above 0.84 (Cyprus) on the first dimension and above 0.59 on the second dimension (Latvia), while the corresponding values in the Y space are 0.94 (Cyprus) on the first dimension and 0.82 (Latvia) on the second dimension.

4 | DISCUSSION

4.1 | Summary of the findings

In this article, we applied two innovative methods, mgPCA and mgPLS, to study the CAST structure and its associations with potential explanatory variables to the 2011 ESPAD data, a school survey where individuals are sampled from 13 groups (countries). In these analyses, the group effect is structuring but non-relevant although differences between each group and the common structure (for all groups) are computed. These methods are based on maximization criteria; they seek common and group loading weights which may be useful for enhancing the interpretation of the outcomes and in graphical displays. The proposed methods and their associated interpretation tools are developed in the "multigroup" package of the free software R (Eslami et al., 2015).

Applying mgPCA, we conclude that for most of the countries under study, the CAST variables are related with each other on the first component (frequency of problems and use) but bear some differences on the second one (opposition of non-recreational use and dependency symptoms): this overall structure appears stable across the 13 countries with some minor exceptions for Kosovo, Liechtenstein and Romania. With mgPLS, we conclude that three main variables are related with CAST: the frequencies of cannabis use in the last 12 months and in the last 30 days, in a positive way, and the age at first cannabis use in a negative way. The number of friends smoking cannabis and the perception of cannabis risks play only a secondary role. Again, Kosovo, Liechtenstein and Romania present different relationships between explanatory variables and CAST, especially for the age at first cannabis use, these secondary variables and with alcohol use. In these three countries, the link between the CAST and the precocity of onset is very weak and even positive in Romania. All these findings suggest that the patterns of psychoactive drugs use differ from the average in these countries and that

the place of cannabis as a psychoactive product among alcohol and tobacco is specific.

A replication of the analyses without Kosovo, Liechtenstein and Romania suggest that the results concerning the CAST structure are robust but that Latvia is still standing apart when considering the link between the explanatory variables and the CAST. A more thorough investigation of the diffusion of cannabis in these countries would be necessary.

4.2 | Comparison with previous studies assessing the CAST structure

Previous studies of the CAST using confirmatory factor analysis (CFA) were conducted in separate countries. They found a uni-dimensional structure or a bi-dimensional structure (CAST1 and CAST2 on the first factor, the other items on the second one), the two being eventually highly correlated (Pearson correlation coefficient above 0.7) (Bastiani et al., 2013; Cuenca-Royo, Sánchez-Niubó, Torrens, Suelves, & Domingo-Salvany, 2013; Fernandez-Artamendi et al., 2012; Legleye et al., 2015; Legleye, Piontek et al., 2011b; Legleye et al., 2013), except in one study where it was above 0.5 (Cuenca-Royo et al., 2012). First, we have to notice that in these publications, the number of dimensions seems to depend on sample size and heterogeneity as two dimensions are more common in large samples with large age ranks. With the ESPAD survey, we have a very large sample size and heterogeneous populations, which is thus clearly in favour of two dimensions. Second, aims of the CFA are to reduce the dimensionality and to derive a simplified model of a set of variables according to both theory and observation. In all publications mentioned earlier, cross-loadings of the items were not considered, that is another standard simplification, and factors were correlated. In contrast, in PCA as in mgPCA, the dimensions derive only from the data: variables load on every principal components (there are cross-loadings). All components are computed but they are not correlated by construction: the reduction in dimension is made only by the interpretation of a limited number of components. CFA and PCA (on which is based mgPCA) are different methods: as a consequence, we cannot compare the principal components to the factors in CFA although we may, to some extent, compare the components to the factors produced by exploratory factor analysis (EFA) because EFA considers cross-loadings (although correlations between EFA factors depend on the rotation) than to those of the CFA.

However, the structural information provided by CFA and mgPCA are coherent. In mgPCA, the first component is the frequency of use and problems (that is the accumulation of all items) which is clearly dominant, while the second is the opposition between non-recreational use and problems. In CFA, the first factor is the frequency of non-recreational use and the second is the frequency of problems. As these factors are highly correlated, their common part (which is important given the level of correlation) reconstructs the first component which is clearly dominant, whereas their opposition is captured by the second orthogonal component, which is clearly secondary. This is at least the case for the countries that are close to the common structure, and that studied the CAST with CFA namely France and Italy (Spain did not conduct ESPAD) (Bastiani et al., 2013; Legleye et al., 2015; Legleye et al., 2013).

4.3 | Multigroup analyses for the case of unbalanced groups or unbalanced dimensions across groups

Depending on the data pre-treatments, multigroup analyses handle unbalanced groups in different ways. As variables are not necessarily scaled within groups, the group weight may depend of the group inertia and is usually higher for a small group. When variables are scaled within groups (our choice in this study), all the groups have the same weight in the analysis. Therefore, the scaling question must be carefully answered in regard with the data feature and the goal of the statistical analysis. Multigroup analyses being driven by seeking common components to all the groups, the common structure as well as the group structures have all the same dimensions. However, groups may be summarized by different number of components. This point must be studied further.

4.4 | Alternative comparable methods

It is worth noting that mgPCA gives the same common loadings as Dual Multiple Factorial Analysis (DMFA) (Lê & Pagès, 2010) implemented in the popular FactoMineR R package. But mgPCA has the advantages of (i) processing group loadings as well as common loadings associated with a useful similarity index and (ii) being based on a straightforward maximization criterion which can be extended to more complex data such as two block and multiblock data (Eslami, 2013). mgPCA can also be compared to the Simultaneous Component Analysis (SCA) (Kiers & Ten Berge, 1994). It has been proven that the well-known version of SCA (SCA-P) leads to the same results as mgPCA (Eslami, 2013).

4.5 | Invariance analysis

Measurement invariance is usually tested with multigroup CFA (Bartholomew, Knott, & Moustaki, 2011; Matsumoto & Van de Vijver, 2011). Configural (similarity of the factors across groups), metric (identity of the loadings across groups) and scalar (identity of the intercepts across groups) invariances are successively checked. Theoretically, only scalar invariance (i.e. measurement invariance) allows comparing means across groups. CFA relies on distributional assumptions of variables and on regression modelling. One advantage is that it uses statistical testing but from a practical point of view, it is complex and scalar invariance is rarely achieved: the number of tests and parameter constraint changes are usually high before assessing to what extent the scale is not fully invariant across the groups (Muthén & Asparouhov, 2012), especially when the numbers of groups or variables are large (Van de Schoot, Schmidt, & De Beuckelaer, 2015). These methods rely on an abundance of fit indexes and are sensitive to sample size and to misspecifications (Saris, Satorra, & van der Veld, 2009). In addition they need quite large samples to be reliable, as many parameters are estimated; finally, the conclusions are sensitive to the choice of the parameters that are considered fixed only because of identifiability constraints and also to the choice of a country of reference. Nevertheless, they can handle binary, ordinary and multinomial variables. The proposed mgPCA and mgPLS are not originally designed to check for invariance measurement: they intend to provide a reliable picture of the common relationships between variables in multigroup settings.

They do not provide statistical tests and the input variables are all treated as continuous; they are not sensitive to the choice of identifiability constraints; they assess differences to the common structure instead of a group of reference. As a result, our judgement about the robustness of the structure of the CAST in the ESPAD survey is not a claim of "true" configural or metric invariance but is a strong indication that they may hold: it is a preliminary and methodological work before testing invariance. However, mgPLS shows hints to understand differences between countries that are observed in mgPCA, which is not straightforward in multigroup CFA.

In alcohol research, it has been found that differences in culture may impair cross-national comparisons of alcohol use disorders, even when standardized tests are used (Room, 2006) because the patterns of use may differ so much that the established thresholds may need adaptations. The CAST is usually used to screen people with moderate or high risk of problematic use: it may be the case that beyond the global structure of the CAST variables that we found, country specific thresholds have to be used to obtain measurement invariance with the CAST as a screening instrument. This will be the purpose of a further study using multigroup CFA.

4.6 | Limitations

In this analysis, we only consider some explanatory variables to explain the CAST variables, all being measured at the individual level. Socio-economic variables considered at the individual level could also play an important role; unfortunately, the most relevant variables (highest education level of each of the parents) are optional and were not available for all countries. In addition, multigroup methods cannot consider contextual aggregated data at the country level as all the variables are country-centred. It is highly probable that these variables, e.g. country wealth, unemployment rates, education level, legislations towards drug use, play an important role as it was found for alcohol use in the ESPAD data (Legleye, Morand, & Garnier, 2011a). For instance, Kosovo and Liechtenstein are small and specific countries, the first one being a recent independent state that endured war and social troubles in the past years, whereas the second one is one of the wealthiest countries in the world with less than 40,000 inhabitants. Romania is a recent and relatively poor member of the European Union with an important agricultural sector. The sample sizes for these countries are also very small (respectively 52, 55 and 93 subjects) and their exclusion led to much higher similarities. Nevertheless, a previous publication on the 2007 ESPAD data showed that the most influential variables explaining cannabis perception were cannabis use and use by friends, far beyond country-level variables (Piontek et al., 2013).

5 | CONCLUSION AND PERSPECTIVES

The CAST structure appears rather stable across the 13 studied countries despite some discrepancies in three countries (Kosovo, Liechtenstein and Romania), that are linked to differences in the relation between the explanatory variables and the CAST variables. mgPCA and mgPLS may be used in epidemiology or sociology studies involving a priori groups or populations. The next step will be to study its

measurement invariance. On a methodological point of view, the mgPCA and mgPLS may be used in studies involving a priori groups or populations. For the case of surveys with a more complex block structure (e.g. socio-economic, drug use and health indicators, all measured on the same individuals) as well as a group one (e.g. individuals from different countries), multiblock and multigroup analysis can be applied (Eslami, Qannari, Kohler, & Bougeard, 2014b; Sabatier & Vivien, 2008).

DECLARATION OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

REFERENCES

- Annaheim, B., & Legleye, S. (2016). Screening for cannabis-related problems: Biology, diagnosis, treatment and pharmacology. In V. R. Preedy (Ed.), *The Comprehensive Handbook of Cannabis Related Pathology*. Amsterdam: Elsevier.
- Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*. New York: Wiley.
- Bastiani, L., Siciliano, V., Curzio, O., Luppi, C., Gori, M., Grassi, M., & Molinaro, S. (2013). Optimal scaling of the CAST and of SDS scale in a national sample of adolescents. *Addictive Behaviors*, 38(4), 2060–2067. doi:10.1016/j.addbeh.2012.12.016
- Beck, F., & Legleye, S. (2008). Measuring cannabis related problems and dependence at the population level. In S. Rödner Sznitman, B. Olsson, & R. Room (Eds.), *A Cannabis Reader: Global Issues and Local Experiences*. (pp. 29–57). Lisbon: European Monitoring Center on Drug and Drug Addiction (EMCDDA).
- Becker, H. S. (1963). *Outsiders: Studies in the Sociology of Deviance*. New York: The Free Press of Glencoe.
- Cuenca-Royo, A. M., Sánchez-Niubó, A., Forero, C. G., Torrens, M., Suelves, J. M., & Domingo-Salvany, A. (2012). Psychometric properties of the CAST and SDS scales in young adult cannabis users. *Addictive Behaviors*, 37(6), 709–715.
- Cuenca-Royo, A. M., Sánchez-Niubó, A., Torrens, M., Suelves, J. M., & Domingo-Salvany, A. (2013). [The Cannabis Abuse Screening Test (CAST) for psychiatric disorders diagnosis in young cannabis users] *Adicciones*, 25(1), 87–88.
- European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (2009). *A Cannabis Reader: Global Issues and Local Experiences. Volume 1: Cannabis in the Past*. Lisbon: EMCDDA.
- European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (2011). *2011 Annual Report on the State of the Drugs Problem in Europe*. (pp. 105). Lisbon: EMCDDA.
- Eslami, A. (2013). *Analyses factorielles de données structurées en groupes d'individus*. Rennes: Application en biologie.
- Eslami, A., Qannari, E. M., Bougeard, S., & Sanchez, G. (2015). Package 'multigroup'. CRAN, Comprehensive R Archive Network for the R programming language.
- Eslami, A., Qannari, E. M., Kohler, A., & Bougeard, S. (2013a). General overview of methods of analysis of multi-group datasets. *Revue des Nouvelles Technologies de l'Information*, 25, 108–123.
- Eslami, A., Qannari, E. M., Kohler, A., & Bougeard, S. (2013b). Multi-group PLS regression: Application to epidemiology. In G. Russolillo (Ed.), *New Perspectives in Partial Least Squares and Related Methods*. (pp. 243–255). Berlin: Springer-verlag.
- Eslami, A., Qannari, E. M., Kohler, A., & Bougeard, S. (2014a). Algorithms for multi-group PLS. *Journal of Chemometrics*, 28(3), 192–201.
- Eslami, A., Qannari, E. M., Kohler, A., & Bougeard, S. (2014b). Multivariate data analysis of multigroup datasets. *Application to sensory analysis. Chemometrics and Intelligent Laboratory Systems*, 133, 63–69. doi:10.1016/j.chemolab.2014.01.016
- Fernandez-Artamendi, S., Fernández-Hermida, J. R., Muñiz-Fernández, J., Secades-Villa, R., & García-Fernández, G. (2012). Screening of cannabis-related problems among youth: the CPQ-A-S and CAST questionnaires. *Substance Abuse Treatment Prevention and Policy*, 7(13), 2–9.
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (Eds) (2011). *Applied Longitudinal Analysis*. Hoboken, NJ: John Wiley & Sons.
- Goldstein, H. (2010). *Multilevel Statistical Models*. Hoboken, NJ: John Wiley & Sons.
- Gyepesi, A., Urban, R., Farkas, J., Kraus, L., Piontek, D., Paksi, B., ... Demetrovics, Z. (2014). Psychometric properties of the Cannabis Abuse Screening Test in Hungarian samples of adolescents and young adults. *European Addiction Research*, 20(3), 119–128. doi:10.1159/000353238
- Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., & Kraus, L. (2009). *The 2007 ESPAD Report – Substance Use Among Students in 35 European Countries*. (pp. 408). Stockholm: The Swedish Council for Information on Alcohol and Other Drugs (CAN).
- Jolliffe, I. T. (2002). *Principal Component Analysis*. New York: Springer-verlag.
- Kiers, H. A. L., & Ten Berge, J. M. F. (1994). Hierarchical relations between methods for simultaneous component analysis and a technique for rotation to a simple simultaneous structure. *British Journal of Mathematical and Statistical Psychology*, 47(1), 109–126. doi:10.1111/j.2044-8317.1994.tb01027.x
- Lê, S., & Pagès, J. (2010). DMFA: Dual hierarchical multiple factor analysis. *Communications in Statistics – Theory and Methods*, 39(3), 483–492.
- Legleye, S., Guignard, R., Richard, J. B., Ludwig, K., Pabst, A., & Beck, F. (2015). Properties of the Cannabis Abuse Screening Test (CAST) in the general population. *International Journal of Methods in Psychiatric Research*, 24(2), 170–183. doi:10.1002/mpr.1465
- Legleye, S., Karila, L., Beck, F., & Reynaud, M. (2007). Validation of the CAST, a general population Cannabis abuse screening test. *Journal of Substance Use*, 12(4), 233–242.
- Legleye, S., Morand, E., & Garnier, B. (2011a). Influence de la parité économique et politique sur la différence genrée de consommation d'alcool chez les adolescents européens [Influence of economic and political gender inequalities on gender alcohol use in european adolescents]. In Insee (Ed.), *Journées de méthodologie statistique*. Paris: Insee.
- Legleye, S., Piontek, D., & Kraus, L. (2011b). Psychometric properties of the Cannabis Abuse Screening Test (CAST) in a French sample of adolescents. *Drug & Alcohol Dependence*, 113(2–3), 229–235. doi:10.1016/j.drugalcdep.2010.08.011
- Legleye, S., Piontek, D., Kraus, L., Morand, E., & Falissard, B. (2013). A validation of the Cannabis Abuse Screening Test (CAST) using a latent class analysis of the DSM-IV among adolescents. *International Journal of Methods in Psychiatric Research*, 22(1), 16–26. doi:10.1002/mpr.1378
- Legleye, S., Piontek, D., Pampel, F., Goffette, C., Khlat, M., & Kraus, L. (2014). Is there a cannabis epidemic model? Evidence from France, Germany and USA. *International Journal of Drug Policy*, 25(6), 1103–1112. doi:10.1016/j.drugpo.2014.07.002
- Matsumoto, D., & Van de Vijver, F. J. (Eds) (2011). *Cross-cultural Research Methods in Psychology*. New York: Cambridge University Press.
- McCullagh, P., & Nelder, J. (1989). *Generalized Linear Models*. London: Chapman & Hall.
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychology Methods*, 17, 313–335. doi:10.1037/a0026802
- Pabst, A., Kraus, L., & Piontek, D. (2012). The Cannabis Abuse Screening Test (CAST): Examining the prevalence of cannabis-related problems among adolescents in 13 European countries. In B. Hibell, U. Guttormsson, A. Salme, O. Balakireva, T. Bjarnason, A. Kokkevi, & L. Kraus (Eds.), *The 2011 ESPAD Report Substance Use Among Students in*

- 36 European Countries. (pp. 158–163). Stockholm, Sweden: The Swedish Council for Information on Alcohol and Other Drugs (CAN).
- Peretti-Watel, P., Constance, J., Guilbert, P., Gautier, A., Beck, F., & Moatti, J. (2007). Smoking too few cigarettes to be at risk? Smokers' perception of risk and risk denial, a French survey. *Tobacco Control*, 16, 351–356.
- Piontek, D., Kraus, L., Bjarnason, T., Demetrovics, Z., & Ramstedt, M. (2013). Individual and country-level effects of cannabis-related perceptions on cannabis use. A multilevel study among adolescents in 32 European countries. *Journal of Adolescent Health*, 52(4), 473–479. doi:10.1016/j.jadohealth.2012.07.010
- Piontek, D., Kraus, L., & Klempova, D. (2008). Short scales to assess cannabis-related problems: A review of psychometric properties. *Substance Abuse Treatment, Prevention, and Policy*, 3(1), 25.
- Piontek, D., Kraus, L., & Pabst, A. (2009). The ESPAD cannabis module. Cannabis-related problems among adolescents in 17 ESPAD countries. In B. Hibell, S. Ahlström, O. Balakireva, T. Bjarnason, A. Kokkevi, & L. Kraus (Eds.), *The 2007 ESPAD report Substance Use Among Students in 35 European Countries*. (pp. 164–170). Stockholm: The Swedish Council for Information on Alcohol and other Drugs (CAN).
- Rehm, J., & Room, R. (2015). Cultural specificity in alcohol use disorders. *Lancet*. doi:10.1016/S0140-6736(15)00123-3
- Room, R. (1998). Alcohol and drug disorders in the International Classification of Diseases: a shifting kaleidoscope. *Drug and Alcohol Review*, 17(3), 305–317. doi:10.1080/09595239800187141
- Room, R. (2006). Taking account of cultural and societal influences on substance use diagnoses and criteria. *Addiction*, 101(S1), 31–39.
- Sabatier, R., & Vivien, M. (2008). A new linear method for analyzing four-way multiblock tables: STATIS-4. *Journal of Chemometrics*, 22(6), 399–407.
- Saris, W., Satorra, A., & van der Veld, W. M. (2009). Testing Structural Equation Models or detection of misspecifications? *Structural Equation Modeling*, 16, 561–582. doi:10.1080/10705510903203433
- ter Bogt, T., Schmid, H., Gabhainn, S., Fotiou, A., & Vollebergh, W. (2006). Economic and cultural correlates of cannabis use among mid-adolescents in 31 countries. *Addiction*, 101, 241–251.
- ter Bogt, T. F. M., de Looze, M., Molcho, M., Godeau, E., Hublet, A., Kokkevi, A., ... Pickett, W. (2013). Do societal wealth, family affluence and gender account for trends in adolescent cannabis use? A 30 country cross-national study. *Addiction*, 109, 273–283. doi:10.1111/add.12373
- Van de Schoot, R., Schmidt, P., & De Beuckelaer, A. (Eds.) (2015). *Measurement invariance*. Lausanne: Frontiers Media. doi:10.3389/978-2-88919-650-0
- Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In P. R. Krishnaiah (Ed.), *Multivariate Analysis*. (pp. 391–420). New York: Academic Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Legleye S, Eslami A, Bougeard S. Assessing the structure of the CAST (Cannabis Abuse Screening Test) in 13 European countries using multigroup analyses. *Int J Methods Psychiatr Res*. 2017;26:e1552. <https://doi.org/10.1002/mpr.1552>