The development of substance use in adolescence: results from comparing two longitudinal studies in England and Germany

Alex Sutherland* & Luca Mariotti*

1. Introduction.

Substance use (1) by young people is regarded as a problem in many Western countries. Even with rates of behaviour such as smoking in decline since the mid-1990s, recent evidence suggests that other forms of substance use such as drinking have increased in some countries (2). Studies examining the development of substance use typically compare data derived from multiple

* Dr. Sutherland is Research Methods Teaching Associate at the Joint Schools’ Social Science (JSSS) programme at Cambridge, UK.
* Sociologist, PhD candidate at the University of Bielefeld, Germany.
cohort cross-sectional studies, such as the European School Survey Project on Alcohol and Other Drugs (www.espad.org). Although very informative at the aggregate level, these studies fail to account for individual differences in the developmental process; we know how many people use substances every year, but we do not know for how many years every subject remains involved in substance use. Thanks to developments in longitudinal research we are now able to overcome these shortcomings and study how the frequencies of alcohol abuse and cannabis consumption evolve for each subject during adolescence. Within the theoretical and methodological framework of Developmental and Life-Course Criminology (DLC) (3), over the last twenty years, many new longitudinal studies concerned with the development of crime and problem behaviours have been carried out (4) also within the European context. In this paper, we compare alcohol and cannabis use using data from two ongoing longitudinal studies in England and Germany. The focus lies on both theoretical and methodological issues. In the first case, we approach the study of substance use from a developmental perspective, analysing individual variations in the amount of use in a period, adolescence, where young people are known to be more prone to experiment with risk taking behaviours (5). In the second case, we statistically compare developmental trends in the two cohorts, focusing on differences in both the frequency of use and the shape of the development. After highlighting existing research in this domain, the first part of the paper describes the studies and data used, including the steps taken in making the comparison. The second part of the paper details this comparison and the results from multiple-group multiple-cohort latent growth models, which focus on individual rather than aggregate level differences between the samples. The results show that the Peterborough sample display considerably earlier initiation and higher levels of use than their Duisburg counterparts.

2. Previous research.

This, of course, is not the first time comparisons have been made between adolescent substance use in different countries. Recently, there have been moves to standardize data collection on this subject across a number of European countries (or collect data if none exist). The results of this effort, for example the aforementioned ESPAD study, provide good quality cross-sectional data on 35 European countries, which includes ‘old’ and ‘new’ European states. Data from ESPAD suggest that there would be large differences between English and German samples for both alcohol and cannabis use. The results for ‘drunkenness in the last 30 days’ (6) and ‘lifetime cannabis use’ for 15/16 year olds differ notably. For alcohol, 8.5% of German adolescents compared with 14% of English adolescents reporting being drunk in the last 30 days. For lifetime cannabis use (ESPAD does not have a more recent measure), 27% of the German, compared to 38% of the English, sample reported having ever used cannabis by the time they reach 15/16 (7) (see figures 5 and 6 below). However, although some other studies have dealt with the development of drug and alcohol use using longitudinal data and statistical techniques (8), to our knowledge, there are none which have
attempted to statistically compare developmental trajectories in two European countries.

3. Preliminary questions for comparison.
Before making comparisons of (any) data, we thought it sensible to answer a number of questions which are set out below (figure 1). The first two sets of questions are applicable to comparisons for any kind of data and are sometimes overlooked when making comparisons between studies, particularly the issue of representativeness. Further, when studies are patently non-comparable, readers are sometimes required to imagine the equivalent of $2 + 2 = 5$. We hope that such stretches of the imagination are not required here.

- **Who (or what) are we comparing?**
  a. People, places, businesses, etc.?
  b. What data are we comparing – numerical (what kind?), text?
- **Study representativeness**
  a. Are the studies representative of their respective populations?
  b. Are those populations comparable (or is it a case of ‘apples and oranges’)?
- **Attrition**
  a. Does this affect one study more than another?
  b. Does this prevent comparison?

Figure 1: questions for data comparison

The third set of questions relates specifically to the comparison of longitudinal data. Attrition is the bane of such projects and can be problematic for studies on their own. When making comparisons between studies this problem is magnified and may prevent such comparisons being made. We will return to these questions later on in this article.

4. Study data.
Data for this paper come from two ongoing longitudinal studies which are focused on the study of crime. The first study is the ESRC funded Peterborough Adolescent and Young Adult Development Study (hereafter PADS+). The second is Kriminalität in der modernen Stadt (‘Crime in the Modern City’ [CRIMoC]), which is funded by the German National Science Foundation (DFG).

- **Peterborough Adolescent Development Study (PADS+)**
  PADS+ is a classic panel study, following a cohort of the same individuals over time for a given period. The study is specifically designed to test a new general theory of crime, Wikström’s Situational Action Theory (9). PADS+ is comprised of a one-third random sample of all young people in Peterborough, a medium sized city in the county of Cambridgeshire, England. PADS+ has been running since 2003 when data collection began with interviews of all parents whose children were involved in the study. Each year, participants are interviewed in groups and individually, and asked to report their criminal and substance use behaviour in the previous year and their current attitudes (along with many other measures). The initial sample for PADS was 716 young people. So far, the study has collected five
waves of data from young people, with a retention rate of 97% over the five waves. For the first four waves of data from PADS+, which this paper uses, the retention rate was 98%. The starting age of the sample is 11/12, which for most young people in England and Wales is the age of transition from primary school into secondary school. There is a 50:50 ratio of males to females in PADS+ (which one might expect from a random sample). Unlike some longitudinal studies of young people undertaken in England (10), PADS+ is representative of all young people in Peterborough and the East of England – not just school-attendees or those who happen to be at school on a given day (11). This was achieved by the study team following up all individuals who were not attending school (such as truants, delinquents or those ill on days the research team visited), interviewing them in libraries or other public buildings in order to retain the sample (further details of the methodology can be found on the study website www.pads.ac.uk).

• ‘Crime in the Modern City’ (CRiMoC)

CRiMoC is also a study concerned with crime as a social problem, and has been running annually since 2002. The study is more sociological in nature, but contains within it factors from numerous criminological theories. CRiMoC is a panel study which uses a cross-sectional data collection method. The study tracks one cohort of individuals over time, but rather than focusing on a sub-sample of this group (as PADS+ does), surveys all members every year during the study. In essence, the study is a population study of school-children in Duisburg, Germany (12). As with PADS+, participants were asked to self-report on their behaviour in the previous year. Unlike PADS+, there is some sample attrition primarily owing to difficulties with data collection requirements. The study was not allowed to collect names of individuals, and instead used multiple anonymous references to construct identifiers (13). In some instances, this meant that young people were unable to recall details from previous waves, resulting in unmatched questionnaires. The data used for this paper come from 1,552 individuals who returned data in the first five waves of the study (there is now a sixth but this is not utilized here). The start age for the CRiMoC study was 12/13, and the male/female split is 40:60. Overall however, the data are broadly representative of school-attending young people in Duisburg (14). A summary comparing the two studies is given below in table one; ethnicity is not compared because of differences between the ways in which the studies collected this data (discussed in more detail below).

<table>
<thead>
<tr>
<th>PADS+ (England)</th>
<th>CRiMoC (Germany)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designed as a ‘classic’ panel study.</td>
<td>1. Panel study with cross-sectional collection method.</td>
</tr>
<tr>
<td>Random 1/3 sample of all YP in Peterborough.</td>
<td>2. Population study of schools in Duisburg – derived panel data.</td>
</tr>
<tr>
<td>Little sample attrition (98% retention rate).</td>
<td>3. Attrition due mostly to confidentiality problems.</td>
</tr>
<tr>
<td>N=700 (four waves panel).</td>
<td>4. N=1,552 (five waves panel).</td>
</tr>
<tr>
<td>50:50 Male:Female.</td>
<td>6. 40:60 Male:Female</td>
</tr>
</tbody>
</table>
5. Differences between contexts.
Aside from the differences between the studies themselves, there are a number of marked differences between the contexts the studies take place in (summarized in table two below). Some of the more pertinent ones are discussed here and there may well be other factors which are not mentioned that could have a bearing on the level of substance use by a population. Obviously, this list is not exhaustive and other differences in context might have implications for the results of this comparison (15), but are not the focus of the paper. The major difference between school contexts is that the German school system is tiered – from Year 5 onwards (age 10) children are streamed into school types depending on ability. Both teachers and parents can nominate children for different types of schools. The resulting three school types represent different emphases on academic ability (Gymnasium), general work readiness (Realschule) and vocational/technical skills aimed at leading to apprenticeships (Hauptschule). There is also one type of school which is a mixture of these (Gesamtschule), more closely resembling the English state system of mixed ability schools (16).
Perhaps more importantly for the purposes of this article are differences in the legal status of alcohol and cannabis. For Britain’s ‘favourite drug’ (17) alcohol, the minimum age at which purchase is permitted is 18. Further, it is illegal ‘for anyone to buy alcohol for someone under 18 to consume in a pub [bar] or a public place’ (18). Drinking in the home is subject to parental discretion or personal choice, but it is illegal for under-fives to be given alcohol to drink. Possession of alcohol in the street depends on local laws – many cities and towns operate ‘alcohol-free’ zones (though these are rarely alcohol-free in reality), where alcohol can be confiscated by the police. Beer, wine or cider can be drunk by 16 and 17 year olds if they are dining and if they are accompanied by an adult (anyone over 18), and that adult purchases the alcohol for them.
For cannabis, the UK picture is mixed and somewhat confusing but it remains illegal to possess cannabis. The British government maintains a sliding scale of drug classification, Classes A, B and C, which groups together drugs under the Misuse of Drugs Act 1971. Class A is the highest category, with punishments for possession and distribution the most severe – examples of Class A drugs are heroin, cocaine and ecstasy (19). Cannabis was a Class B drug until 2004, when it was downgraded, following advice from the Advisory Council on the Misuse of Drugs (20), to a Class C (after penalties for distribution of Class C drugs had been increased). Political wrangling in the next few years resulted in cannabis being upgraded to a Class B drug in January 2009, against the advice of the ACMD. Possession of cannabis now carries a maximum of five years’ imprisonment, but police are able to give first time adult offenders a discretionary warning, or issue a fine. For young people, officers have the discretion to pursue a reprimand, caution or Final Warning (21). The effect of
changes to cannabis classification has been one of some confusion, and there is a question over whether the ‘right’ message was received by the population, in particular by young people (22). In Germany, children of 14 and older can drink undistilled beverages if accompanied by their parent(s). The minimum legal age for unaccompanied purchase and possession of alcohol depends on the type of alcohol. For most undistilled beverages (e.g. beer, cider and wine) the age is 16, whereas for spirits it is 18. Public drinking is allowed in the majority of cities, though drinking on public transport has been technically illegal in Berlin since 1999 (23). Recently one University City, Freiburg, had its public drinking ban overturned by the local Administrative Court, but some cities maintain such bans (24). Following a general trend in many EU countries (25) – drug policies over the last 20 years in Germany have shifted toward decriminalization for possession of small amount of drugs (especially cannabis), for personal use (26). So, although cannabis possession in Germany is still formally illegal, since 1994 limits for prosecuting possession of a “small amount” of cannabis depend on Federal State laws. As a result, individuals are rarely prosecuted for possessing fewer than 5g of cannabis, though there is some regional variability with the implementation of this law (27).

<table>
<thead>
<tr>
<th>England</th>
<th>Germany</th>
</tr>
</thead>
<tbody>
<tr>
<td>State/independent school systems with mixed ability classes (streaming within subjects).</td>
<td>Tiered school system.</td>
</tr>
<tr>
<td>Legal age for drinking: 18</td>
<td>Legal age for drinking: 16/18</td>
</tr>
<tr>
<td>Public drinking is subject to local By-laws with most city/town centres ‘alcohol-free zones’.</td>
<td>Public drinking (mostly) legal across Germany, with some cities introducing bans in recent years.</td>
</tr>
<tr>
<td>Cannabis is illegal to possess across the UK.</td>
<td>Cannabis use is illegal; prosecution for possession depends on Federal State laws which have public interest clauses.</td>
</tr>
</tbody>
</table>

Table 2: comparison of study contexts

6. Study comparison questions.

• Study representativeness

Slightly out of turn, we return to our comparison questions by first asking ourselves whether the studies are representative. As noted above, owing to random selection the data from PADS+ are representative of all young people in Peterborough and the East of England with very little attrition. However, the different data collection method in CRiMoC and the difficulties arising from anonymity meant that not all subjects completed data for all years of the study. There were also a number of occasions where individual data could not be matched to previous years. To derive panel data, only those individuals who were captured for the first five waves of CRiMoC were included in this analysis (n=1,552). The derived panel data differ from the cross-sectional sample in a number of ways. First, the panel data under-represents children from lower-class backgrounds (fewer children from Hauptschule are included) (28). Second, there are significant differences between the prevalence and frequency of alcohol and
cannabis use between the cross-sectional and panel data from CRiMoC. In short, there is a lower frequency of use reported by those included in the panel sample (data available on request). The difference between cross-sectional and panel data are larger for cannabis use than for alcohol. The frequency of alcohol consumption tends to be significantly underestimated in the panel data only at the first three measurement occasions. Similar results were obtained for CRiMoC when comparing the frequencies of crime between the original panel and a Full Information Maximum Likelihood (FIML)-estimated one, which addressed the problem of unit nonresponse and included all subjects who attended the study at least twice. The newly estimated sample reported higher level of crime although similar trajectories (29). This means that the CRiMoC panel data likely under-represents the level of use for these substances in the German sample, although should retains a very similar developmental trend over time.

- **Attrition**

Deriving the panel data from the cross-sectional sample means that there is no sample attrition for CRiMoC, but there are some differences between the panel and cross-sectional data (as noted above). For PADS+, sample attrition amounts to only 2% of the overall study sample across the first four waves of the study.

- **What is being compared?**

With these caveats in mind, we next ask ourselves what is being compared. Both studies collect one-year retrospective self-report data from young people on their drug and alcohol use (along with more detailed information on criminality). Matching the two dependent variables was relatively straightforward as count data were available for either cannabis or alcohol use in one of the studies, meaning this could be matched to categorical data used in the other. More problematic was the matching of the two cohorts. Aside from the difference in starting age for the studies (which is accounted for in later analyses), it seems the English school system does not tolerate skipping or repeating school years, meaning that school year groups are relatively homogenous in terms of age. This is not so in the German system – table three shows data from a cross-section of the first wave of CRiMoC data, corresponding to age 12/13 (school Year 8 in England and Wales; Year 7 in Germany). We can see that there are many individuals much older than might be expected which are part of the same school cohort. As such, making cross-national comparisons using school year alone would be misleading. At the very least, we might expect those aged 14 and over to be developmentally different from those at the ‘normal’ ages for this year group.

<table>
<thead>
<tr>
<th>Age, t&lt;sub&gt;1&lt;/sub&gt;</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>416</td>
<td>26.8</td>
</tr>
<tr>
<td>13</td>
<td>907</td>
<td>58.44</td>
</tr>
<tr>
<td>14</td>
<td>193</td>
<td>12.44</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>1.1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>0.13</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>Missing</td>
<td>15</td>
<td>0.97</td>
</tr>
<tr>
<td>Total</td>
<td>1552</td>
<td>100</td>
</tr>
</tbody>
</table>

In order to overcome this we ‘age-trimmed’ the data from CRiMoC so that all those aged 14 and over in the first year of the study, or where age
data were missing (the shaded area), were excluded from subsequent analyses. This resulted in a loss of 229 individuals from the CRiMoC data, giving a final n-size of 1,323 young people for the Duisburg sample. Those excluded at this stage from the sample did not differ significantly in their level of substance use.

We now return to our questions for comparison studies. Figure two (below) summarises the answers to our earlier questions.

- Who (or what) are we comparing?
  - Age-cohorts of young people in two cities with self-reported data on the prevalence and frequency of drug use.
- Study representativeness
  - Are the studies representative? Yes, but some caveats relating to CRiMoC.
  - Are those populations comparable? Yes, with ‘age-trimming’.
- Attrition
  - Does this affect one study more than another? ‘No’, with caveats.
  - Does this prevent comparison? No, but some caution is required with the results – likely to be underestimates for CRiMoC.

**Figure 2: questions for data comparison**

### 7. Method.

- Descriptive statistics
  
  As a first stage of analysis, we look simply at the prevalence and frequency of use for both studies. Where used, the name of statistical tests and p-values are given.

- Multiple-group multiple-cohort models
  
  In a second stage, we compare the development of substance use in adolescence; for this we carry out a longitudinal analysis in order to best use all the information at our disposal. Latent growth curve models (LGM) are an ad hoc technique for the analysis of growth processes (30). The behaviour of interest is modelled as a function of time, and the developmental process is described by means of latent variables. The latter are used to define the sample mean growth trajectory, which in the simplest case is described by an intercept (mean onset level) and a linear slope (mean growth rate). This model can be further expanded to include polynomial (e.g. quadratic) terms, which can be used when a curvilinear development is expected (31). Further, individual deviation from the sample mean trajectory is captured by the variance measured around the growth parameters.

  Another important feature of latent growth models is the possibility of carrying out multiple group comparisons, as well as cohort-sequential LGM for so-called ‘accelerated’ designs (32). In this particular study, however, we have to face the problem that the two groups were one year apart chronologically, and this age difference should be accounted for in the model. For this special case, Muthén and Muthén (33) propose an extension of multiple group analysis, the so-called multiple group multiple cohort LGM, which is nothing more than a LGM multiple group analysis which takes into account the existence of different aged cohorts.

  Since the data for the studies are in one dataset, another advantage of this modelling strategy is the possibility to estimate a single model, where the trajectories for each group are calculated and compared against each other. In this way it is
possible to test various model specifications with different equality constraints on the two trajectories, and thus find out how much the two groups differ or are similar in their development. This is, in the end, the aim of this study.

8. Expectations.
Judging from the ESPAD data, we can surmise a number of expected differences between the two cohorts, which are presented below.

H3. Therefore, different average individual trajectories of use will be reported in the samples for both (a) alcohol and (b) cannabis.

9. Results.

• Frequency of use
The data presented below in tables four (alcohol) and five (cannabis) are those school years where the two cohorts are the same age. Comparisons at the aggregate level reveal that there are indeed differences between the two groups (t-test p<.001 for both substances in all three directly comparable years). For alcohol, specifically the number of times ‘drinking until drunk’ in the previous year, we can see that the Peterborough group are drinking earlier and more frequently (table four). By 13/14 years old, a quarter of the Peterborough sample report getting drunk up to once a month, compared with only 7% of the Duisburg sample. Similarly, by the time both cohorts are 14/15, 6% of the Peterborough sample report getting drunk ‘more than once a week’, compared to only 2.3% of the Duisburg group. Although startling, these findings are in line with the results reported earlier from ESPAD (34).

<table>
<thead>
<tr>
<th>Age</th>
<th>12/13 PADS+</th>
<th>12/13 CriMoC</th>
<th>13/14 PADS+</th>
<th>13/14 CriMoC</th>
<th>14/15 PADS+</th>
<th>14/15 CriMoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not in previous year</td>
<td>49.9</td>
<td>53.9</td>
<td>34.8</td>
<td>62.3</td>
<td>25.8</td>
<td>48.0</td>
</tr>
<tr>
<td>Once or twice</td>
<td>20.4</td>
<td>9.2</td>
<td>19.1</td>
<td>22.7</td>
<td>10.9</td>
<td>28.3</td>
</tr>
<tr>
<td>Up to once a month</td>
<td>20.4</td>
<td>2.5</td>
<td>25.8</td>
<td>7.0</td>
<td>32.0</td>
<td>11.1</td>
</tr>
<tr>
<td>More than once a month</td>
<td>7.1</td>
<td>0.8</td>
<td>15.2</td>
<td>4.2</td>
<td>23.3</td>
<td>8.2</td>
</tr>
<tr>
<td>More than once a week</td>
<td>1.0</td>
<td>0.3</td>
<td>3.5</td>
<td>1.3</td>
<td>6.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Missing data</td>
<td>1.3</td>
<td>33.3</td>
<td>1.5</td>
<td>2.6</td>
<td>1.8</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Table 4: frequency of drunkenness in previous year

For cannabis, similar but perhaps more striking differences are noted in table five. At age 12/13 only 2.2% of the Duisburg sample report using cannabis at all, compared with nearly one-sixth (approx. 15%) of the Peterborough study sample. By the time they reach 14/15 years old, these differences have ossified – 30.2% of the Peterborough sample report using cannabis in that year, with one-third of the users having done so eleven or more times in the previous twelve months. By comparison, only 12.2% of the Duisburg sample report any use in the same year,
and only 4.9% report using ‘11+ times in the previous year’.

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency of cannabis use</th>
<th>PADS+</th>
<th>CRiMoC</th>
<th>PADS+</th>
<th>CRiMoC</th>
<th>PADS+</th>
<th>CRiMoC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not in previous year</td>
<td>84.5</td>
<td>95.5</td>
<td>77.7</td>
<td>88.8</td>
<td>68.0</td>
<td>83.3</td>
</tr>
<tr>
<td></td>
<td>Once or twice</td>
<td>5.6</td>
<td>0.8</td>
<td>6.2</td>
<td>3.0</td>
<td>8.9</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>3-5 times</td>
<td>3.5</td>
<td>0.6</td>
<td>4.1</td>
<td>1.5</td>
<td>5.6</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>6-10 times</td>
<td>1.3</td>
<td>0.4</td>
<td>2.9</td>
<td>1.4</td>
<td>4.8</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>11+ times</td>
<td>3.9</td>
<td>0.4</td>
<td>7.7</td>
<td>2.0</td>
<td>10.9</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Missing data</td>
<td>1.3</td>
<td>2.3</td>
<td>1.5</td>
<td>3.3</td>
<td>1.8</td>
<td>4.4</td>
</tr>
</tbody>
</table>

**Table 5: frequency of cannabis use in previous year**

Bearing in mind that the Peterborough group contains both those at school and those usually missing from studies of this type, we might expect these differences to be large, but merely an artefact of including those who are known to use substances more frequently (35). However, when excluding those listed as ‘not registered at school’, ‘attending special educational needs school’, ‘not attending school’ or ‘attending school outside Peterborough’ (according to administrative data), the results for the Peterborough sample do not change dramatically. There are some differences recorded when examining the proportions of users and non-users for cannabis in waves three and four (chi² p<.05, p<.01 respectively). Similarly, there are significant differences between the frequency of cannabis use reported for these two groups in waves three and four (two tailed t-test, p<.001 in both cases). However, no differences were found for the proportion of alcohol users and non-users when excluding those ‘not at school’ (results available from first author, upon request).

When examining the frequency of alcohol use between the ‘school’ and ‘non-school’ groups in the PADS+ sample, there are no significant differences (results from first author upon request). The findings for alcohol use are a little surprising, as they suggest that the proportion of young people using alcohol, and the frequency of that use, is not a function of being in school or not. It should be noted that the total number of those categorised as ‘not at school’ in PADS+ increases over time (w1: 26, w2: 46, w3: 58, w4: 105), and the non-significant results for cannabis in the first wave might be a result of this small n-size. Despite this, the differences found later in PADS+ suggest that ‘school only’ samples in England (and possibly elsewhere) may underrepresent the proportion (and frequency) of cannabis use by young people. Perhaps of more interest is the finding that ‘school only’ samples may accurately approximate the proportion of users and the frequency of alcohol use for all young people in a particular age cohort.

Although there are some differences noted for cannabis, it seems safe to suggest that the aggregate differences between the Duisburg and Peterborough groups are robust. This seems more likely when examining the data from CRiMoC alongside the PADS+ data with non-school...
attendees removed. For the three directly comparable years of data, there are significant differences in the proportions of users/non-users and the level of use between the Peterborough and Duisburg samples (results available from first author).

- **Multiple-group multiple-cohort models (MGMC)**
  What is lost with comparisons of aggregate data, such as those presented above, is that a person could report drinking or using cannabis heavily in one year, then report no use the year after – the continuity of use is lost. As noted above, a well-known method to psychological and biological sciences for charting change over time is the estimation of latent growth models (36), and more precisely an extension of this technique called multiple-group multiple-cohort latent growth models (MGMC-LGM) (37). The latter allows one to take into account the difference in age between the two cohorts, and to model that difference within a single statistical model. The result is that a multiple group comparison between the two observed groups (the Duisburg and the Peterborough samples) is carried out in which the observed outcome is a function of age and not of the measurement points (as it would be otherwise) (38).

  Hereafter we present the result for alcohol and cannabis use respectively. In all analyses, the ordinal variables detailed above are treated as continuous.

- **Alcohol use**
  The frequencies of alcohol use – as outlined above – were used to investigate the individual development of alcohol consumption over time. The best model was selected on the basis of model fit indices ($\chi^2 = 5.980$ with p>0.05; RMSEA = 0.022; CFI = 0.999) (39), and resulted in two completely different trajectories for the two cohorts. The results are shown in figure three below.

![Figure 3: Average individual alcohol trajectories for Peterborough and Duisburg cohorts.](image)

In both cases the developmental trajectories show an increasing pattern across the observed time span. On the one hand, the Peterborough cohort is characterized by a slightly curvilinear growth which increases more rapidly between the age of
14 and 15. The Duisburg group, on the other hand, report a constant and linear growth (40). The most striking difference concerns the mean level of alcohol consumption. At the age of 12, the Peterborough cohort already reports a higher level of use compared to their 13 year old Duisburg counterparts. This trend can be seen all across the observed time periods: although there is only one school year separating the two cohorts, the Peterborough sample drink at the equivalent level of pupils two years older than them in the Duisburg sample. Thus, although similar in shape, the two trajectories differ visibly in the number of times the pupils reported being drunk in the last 12 months.

- **Cannabis use**

The frequency of cannabis use was also employed for the estimation of MGMC-LGM. The best model, selected on the basis of model fit indices ($\chi^2 = 11.785$ with $p>0.05$; RMSEA = 0.031; CFI = 0.998), resulted again in two completely different trajectories for the two cohorts. The results are shown in figure four below.

![Figure 4: Average individual cannabis trajectories for Peterborough and Duisburg cohorts.](image)

In this case the differences between the two cohorts are more evident than for alcohol. Although both trajectories show a clear growth, the Peterborough cohort is best described by a steep linear development; whereas the Duisburg group is characterized by a negative curvilinear trajectory. This suggests that the frequency of cannabis use might stabilize in late adolescence for the Duisburg sample (further data from CRiMoC might confirm this). Similarly to alcohol consumption, we can argue that although younger, the Peterborough children tended to use cannabis more frequently (41) and the frequency of use increases more steadily across adolescence than for their Duisburg counterparts.

**10. Summary of results.**

The results from this paper suggest that there are stark differences between Peterborough and Duisburg young people in terms of the
proportions of users (H1), their frequency of substance use (H2), and the trajectories of this use within these two groups (H3). The different analytical approaches, aggregated t-tests and intra-individual trajectory analysis, reflect one another which is reassuring. Further, within the considered age-span, all cohorts report a significant growth in substance use. Finally, although one year younger, the UK cohort shows higher trajectories of substance use.

11. Limitations.

We attempted to compare two cohorts of young people from cities in England and Germany. A critical issue is the actual comparability of the two samples. As noted above, the CRiMoC researchers encountered difficulties with tracking individuals over time for the study. Primarily, this was the result of not being allowed to use the names of individuals in the study. In order to participate, young people had to recall a number of unique identifiers (e.g. the first letter of their eye colour). As recall is likely to be a function of IQ, those who can successfully remember these identifiers in each year are more likely to have a higher intelligence than those who cannot. The implication is that the five wave panel data used here consists of more intelligent individuals who are (statistically) less likely to use drugs or alcohol (perhaps owing to higher education aspirations), and who are also more likely to participate in longitudinal research (42). Problems of recall consistency and reliability are further compounded with cannabis use, as recent use can affect memory (44). If there is a dose-response effect of cannabis use on memory, then those using cannabis the most might self-select out of the CRiMoC study in the long run. Equally, those who smoked cannabis immediately prior to the research might have been unable to recall the identifiers required.

However, there is little that can be done about these issues now – they are something to be acknowledged and worked around. We believe that the strategy employed here ameliorates some of the obstacles to making comparisons between the two studies. First, by trying to match the two groups as closely as possible in terms of chronological age, rather than school year. This eliminated those from the Duisburg sample who may have been developmentally different from the main cohort. Second, by making multiple comparisons where those classed as ‘not attending school’ were removed from the Peterborough cohort. This tries to match the (likely) selection effects resulting from the data-collection problems encountered during the CRiMoC study. Results from this indicate that even when excluding those ‘not at school’, there were still appreciable differences between the two cohorts. Finally, comparative studies may have problems with their dependent variable(s). Here however this was straightforward; question phrasing was the same across the studies, so it was simply a question of matching count data to ordinal categories across the two studies, meaning that the outcomes are being measured in the same way. From our interpretation of the data, it seems unlikely that even with sample bias fully accounted for, the Duisburg cohort would ‘catch up’ the
Peterborough one with respect to levels of both alcohol and cannabis consumption.

In the process of making “good” comparisons using longitudinal studies, we acknowledge the importance of a step by step approach to important issues like representativeness, sample attrition, context differences, and last but not least matching the object(s) of research. The latter, although obvious, is a paramount problem in comparative research: the incompatibility of constructs, in particular the dependent variable.

All in all, these problems have prevented many researchers from undertaking comparative studies, especially among the ongoing longitudinal studies in criminology. The few publications available on the topic bear testament to this situation; Farrington and Wikström (1993), Wikström and Svensson (2008), and Pauwels and Svensson (2009) use longitudinal data to compare crime rates in different countries; Link (2008) focuses specifically on drug use but uses only cross-sectional data (45).

12. Discussion.

Research which focuses solely on one context is sometimes limited in what it can tell us, particularly in situations where one wonders what the relationship between $x$ and $y$ might be, and specifically if one is concerned whether $x$ causes $y$. Comparative research offers the possibility of straightforward counterfactual examples where any number of single studies from within a particular context cannot achieve this. Perhaps the best illustration of this is Zimring’s work (46) on the ‘great crime decline’ in the US. Much time and money has been spent attempting to prove that, for instance, zero tolerance policing or increasing use of custody or indeed criminal justice policies, affected the crime rate in America. However, simply by looking north to Canada, Zimring was able to convincingly demonstrate that none of these can have been the case on their own.

We attempt something similar here – by highlighting the differences between the two contexts it becomes clear that some factors may not be effective ways of managing or changing behaviour. Despite strong evidence and a wide consensus on the harms caused by alcohol and other drugs, if not on how to rank those harms, there is little agreement on how to approach the issue of behavioural change. In Scotland for instance, serious consideration is being given to minimum pricing per unit of alcohol (47), as suggested by the Chief Medical Officer, Sir Liam Donaldson. This suggestion, for reasons unknown or that are unclear, has met with opposition in England and Wales, notably from the (then) British Prime Minister, Gordon Brown. Some (48) have advocated increasing the minimum legal age for purchasing alcohol to 21, on the basis that doing so in America cut the number of alcohol related driving deaths amongst young adults by 1,000 each year (49). At least on the evidence presented here, the minimum age of purchase seems to have the reverse relationship that might be hypothesised given the evidence from the US – Germany has the more ‘relaxed’ legal system with a lower age limit, but lower levels of use by adolescents.

However, the results in this paper are only one example and might not represent the wider picture. But evidence from across Europe suggests that alcohol use (specifically drunkenness) by young people is patterned by legal drinking age,
but not in the way anticipated by Nutt. Figure five shows that even where countries share the same minimum legal age for drinking there are large disparities in the reported prevalence of adolescent drunkenness. This includes countries renowned for binge drinking (such as Poland, Russia and the UK), where reported prevalence varies a great deal. We do observe that the maximum prevalence of drunkenness seems to be related to age – it is lowest in the 16 age band, followed by the 18 group, and highest in the group with no minimum age. We can see that within each age band there are variations which cannot solely be attributable to the minimum legal drinking age; ‘something else’ must be going on. It might be that the legal age and substance use behaviour are unrelated to one another, as was the case with smoking – it was in decline across Britain well before the UK government increased the minimum age of purchase to 18.

![Graph showing the prevalence of drunkenness in students aged 15/16 in 35 European Countries (ESPAD 2003) sorted by minimum legal drinking age.](image)

**Figure 5:** Minimum legal drinking age and drunkenness in last 30 days reported by schoolchildren across Europe

As noted above, cannabis classification has been used as a political football in the UK in recent years. There has been an increased emphasis on the pros and cons of decriminalization or legalization, with comparisons being made to countries with relaxed legislative contexts for possession such as the Netherlands and Portugal. Such comparisons tend to be incomplete, in that they only include these extreme cases and ignore other countries with similar legal contexts but different levels of use.

If we re-examine the relationship between legal sanctions and cannabis using data from ESPAD (figure six), we see that as with alcohol use, there are large variations in prevalence of use even in countries with (arguably) similar penalties for possession (50). In addition to the well-trodden examples of the Netherlands and Portugal, it is also noticeable that countries with far more
restrictive regimes and harsher penalties for possession, such as Cyprus, also have lower prevalence rates than the UK. Tentative though it may be, this suggests that legal sanctions against cannabis possession do not act as a deterrent to young people, and equally more relaxed approaches may not encourage greater prevalence of use (51).

Figure 6: drug use by European students ESPAD data with UK and Germany indicated

13. Conclusion.
To quote some now infamous research from the UK, ‘the challenges of dealing with the harms of alcohol [are] probably the biggest challenge[s] that we have in relation to drug harms today’ (52). Given the long-term effects of alcohol use, particularly heavy alcohol use, it is some surprise and perhaps of some concern that there is such a disparity between two European countries well known for their drinking. But reaching for ‘culture’ as an explanation is unappealing to us. Whilst there seems little else that can adequately explain the different patterns reported here, there are a range of candidate factors which have not been included which might otherwise be important, notably parental behaviour. Equally, there seems to be a fundamental ‘something’ differing between Germany and England in terms of the desirability of cannabis use which is not accounted for by legal context.

We have identified a number of possible ways to extend this paper. Obviously, a key question is trying to explain the trajectories noted. Research by both authors examining explanations for change over time (53), suggest a range of possible factors. However, many empirical assessments of
proposed theories often fall short of explanation, and end up reading like a list of all possible correlates (54). If we are to avoid this trap in the future, we believe that much more attention has to be paid to the discrete social mechanisms operating at the individual level and which are able to link the putative cause to individual action (55). This automatically removes many variables which have otherwise been the focus of research in this area, such as gender and ethnicity, but opens up the inquiry to a wide range of plausible candidate factors.

A clear ‘next step’ from the first proposal would be to extend the intra-individual trajectory explanations to attempt to explain the differences between the cohorts over time. Even more than with the present paper, this would require that theoretical factors are measured in the same way and have the same meaning in both contexts. For more sociological concepts this might not be possible, but if theories are truly general in nature, then cross-cultural comparisons will be able to demonstrate relationships more easily than innumerable within-country studies. Finally, there is the risk that because the LGM results represent the ‘average’ individual, they actually represent no-one at all. Using the technique set out here (growth mixture models) allows for a subtler analysis of substance using sub-groups. One avenue for further exploration might be examining whether there are distinct sub-groups within each study which conform to the kind of individual trajectories found in other longitudinal research on substance use by young people (56).

This paper reflects a first attempt to compare two important longitudinal studies of young people. We acknowledge the many limitations of it, but also recognize its advantages. The latter reflect, first, our intent to report, step by step, the work we deemed necessary to make such a comparison work. Secondly, we were able to describe the development of substance use in adolescence from a longitudinal perspective, applying a new statistical technique (MGMC-LGM) which perfectly suits the need of cross-national comparative analysis, and that – to our knowledge – has not yet been applied in criminology. Thirdly, we acknowledge the important role played by longitudinal research and the Developmental Life-Course Criminology in the contemporary drug use discussion for giving new perspectives and stimulating new research; however, there is still a lot to do in terms of cross-national comparative research in this field. Finally, although some questions have been answered, many new ones have been formulated which we hope to address in the future.

Endnotes.

(1) Note: we use the phrase “substance use” to refer to the use of alcohol, tobacco and all other drugs.
(6) This is very different than asking about ‘trying’


(14) Further details of the study can be found at www.uni-bielefeld.de/soz/krimstadt/.


(34) There is a large amount of missing data in the first
wave of CRiMoC for alcohol use (33.3%). While this would usually be of some concern, the following year shows that those with missing data in wave 1 included the majority of those omitted appear to have been ‘once or twice’ or ‘not in the previous year’. As such, we do not believe that the missing data in the first year of CRiMoC adversely affects the substantive difference noted. 


(38) The MCMG-LGM is a special case of a classical multiple group comparison within the framework of structural equation models (SEM). In a classical multiple group comparison, the basic equation for a general LGM remain the same: \( y_{tk} = \lambda_{21} \eta_{tk} + \lambda_{22} \eta_{tk} + \epsilon_{tk}. \) The suffix \( k \) specifies that for each group \( k = 1, 2, \ldots K \) of the observed group variable, a new growth equation is calculated which results in a different growth trajectory. In the special case of a multiple cohort comparison, the equation does not change its form, only the group specific factor loadings matrix \( \lambda \) change. These values are specified in a sequential fashion in order to reproduce the age difference of the two cohorts. For instance, for the younger PADS+ cohort followed over four time points, the factor loading for the random slope assumes the following values: \( \lambda_{21} = 0, \lambda_{22} = 0.1, \lambda_{23} = 0.2, \lambda_{24} = 0.3. \) For the one-year older CRiMoC cohort: \( \lambda_{21} = 0.1, \lambda_{22} = 0.2, \lambda_{23} = 0.3, \lambda_{24} = 0.4. \) In this way it is possible to match the time points where the subjects have the same age in both groups (see Muthen and Muthen, 1998-2007).

(39) The model fit indices used here represent widely accepted means for defining the goodness of a model in structural equation models (see Bollen, 1989; Preacher et al. 2008). The \( \chi^2 \) tests the baseline-model (the one with no restrictions on the parameters) against the estimated one; the null-hypothesis states that the former is the true model. In this case being the p-value larger than the significance level of 0.05 (\( p = 0.227 \)), we reject the null-hypothesis in favour of the estimated model. The RMSEA suggests a good model when its value is smaller than 0.05, with a boundary value at zero representing a perfect model. Similarly, a CFI value close to 1 are indicative of good model fit. All in all, all these indices support the goodness of our model.

(40) The results of the MGMC-LGM, in fact, showed no need for a curvilinear slope for the German cohort (which was fixed to zero), whereas the English cohort reported a small but significant positive curvilinear growth parameter (standardized curvilinear slope for the English cohort \( Q = 0.153 \), p-value = 0.015)

(41) With regard to the level of use reported in the first wave of both studies (the intercept term) for the two cohorts, it should be noticed that although similar in their values, the intercept for the Duisburg and Peterborough group were statistically different. Thus, we can argue that the Peterborough cohort, although younger, reported a higher frequency of use than the Duisburg one in the first year of the study. 


References.


• Mariotti L., Application of longitudinal analysis techniques to the study of drug use behaviour among adolescents, unpublished doctoral thesis.


• Spiegel, “Drinking in Germany: Soused on the Sidewalk”, in Spiegel Online, 31st March,


