Methodological Appendix

EU Kids Online Country Classification Report

Index

General methodology EU Kids Online:
Adopting an approach that is child-centred, comparative, critical and contextual, EU Kids Online II designed and conducted a major quantitative survey of 9-16 year olds experiences of online use, risk and safety in 25 European countries. In each household one parent and one child were interviewed. The sample was a random household sample stratified by region in each country. The child interview was conducted face-to-face, with a self-completion component for the sensitive questions on online risks.

For more information, and to receive project updates, visit www.eukidsonline.net.
More information on the questionnaires and general methodology can be found here as well.

Note on Country Classification Report Methodology:

All sections in the report use the following statistical techniques: first scales were created in relation to the topic of the particular section. A description is given in each section of these scales and how they were created. Then cluster analyses were used to group children according to the scores on these scales. Subsequently, the proportion of children in each group was calculated for the different countries. These proportions were used in a cluster analysis in each section that led to a classification of countries for each section. The final step (section 4) used a K-Means cluster analysis on the 25 countries based on the proportion of children in each group per country. In the report Ns, averages and percentages are for the total sample, weighted by overall weight. Multi-variate analyses were performed without weights.

This appendix describes the measures, the clustering procedures for the grouping of children and the clustering procedures for country classifications used in the Country Classification report.

Methodology used in the Opportunities Section (1)

This section of the appendix describes the measures, clustering procedures for the grouping of children and the clustering procedure for country classifications used in the section of the report related to the opportunities taken up by young Europeans.

Measures

In this section of the report a few key indicators related to the take up of online opportunities by children were used:

- *Duration of use* (About how long do you spend using the internet on a normal school/weekend day?)*:* these provide plausible information on the quantitative presence of the internet in young people's everyday lives.

- *Breadth of use*: Breadth of use was calculated as the number out of 17 activities the young person undertook in the previous month (Which of these things have you done in the past month on the internet?). Range of activities is substantially correlated with duration of use (r=.46).

- *Nature of activities*: Factor analyses were performed to explore the dimensions of these activities: on the basis of *whether* (see breadth of use) they have done the activities within the last month, and among 11–16 year olds, with what *frequency* (Please tell me how often you have done it in the past month?). To the 17 indicators for online activities, we added two other aspects of online behaviour that seemed particularly important:

- *Having an own profile on a social networking site* (or even more than one profile) is linked to a whole range of possible opportunities. We included a variable 0 - 'No profile', 1- 'One profile' to 2 – 'More than one profile';
- We also included a separate question on social, web2.0 related functionalities. These activities were assessed in the survey through the options 0- 'Never' to 4 - 'Almost every day'. We defined the frequency of these activities as an additional variable ranging from 0 – none of the activities are ever performed, to 1 all five activities performed almost every day).

The respective factor solutions are not statistically distinct; several activities have double loadings on different factors. The double loadings would seem to reflect the hybrid character of many online activities. For example, playing online games with others has both a strong communicative component and involves dealing with interactive content. Visiting a social networking profile might be for communicative or self-presentation reasons.

- *Factor 1 ('Communication'):* Visiting social networking profiles is the marker variable; thus, the factor includes several activities which are mainly communicative. The fact that watching video clips and downloading music or films have the highest loadings may be because these activities are closely related to peer-to-peer communication;
- *Factor 2 ('Creativity'):* Although the loadings are rather moderate all activities involve some degree of creativity or productivity;
- *Factor 3 ('Gaming'):* This factor clearly represents gaming and activities linked to it. Two items have double loadings with 'Creativity';
- *Factor 4 ('Learning'):* The main variable is using the internet for schoolwork, but includes

also reading or watching news on the internet.

Note that 'social online activities' cannot be attributed to any one of these factors. Social online activities are moderately linked to communicative and creative activities, reflecting the link between these activities and use of social media.

Opportunities: Groups of children

Duration of use, breadth of activities and the four factors representing different kinds of online activities were subjected to cluster centre analyses. As a first approach we ran a two-step cluster analysis that provided a 3-cluster solution. Given the high dimensionality of the constructs that have been entered in this analysis, this solution was rejected; the resulting clusters included very heterogeneous patterns of use.

Instead, we ran a cluster centre analysis; based on the criteria of interpretability, stability of cluster membership, and the respective F-values of the variables included, a solution with six groups was selected. The quality indicator of a six cluster solution is not substantially lower than the three cluster solution as proposed by the two-step procedure; both are classified as "medium quality".

Table 1 Average scores on opportunities per opportunities group of children

	Restricted learners	Young networkers	Moderate users	All-round explorers	Intensive gamers	Experienced networkers	EU average
Duration of use	.23	.61	.63	.82	.93	.84	.60
Standardised duration of use	-.10	.04	.05	-.02	.09	.09	.01
Breadth of use	2.95	5.81	7.69	13.19	9.75	9.57	7.20
Standardised breadth of use	-1.22	-.44	.08	1.58	.64	.59	-.06
Communication activities (out of 6)	.91	4.14	3.80	5.27	5.08	5.33	3.53
Communication factor score	-1.38	.55	.02	.11	.77	.55	-.15
Creative activities (out of 6)	.32	1.47	1.41	4.28	2.71	3.31	1.83
Creative factor score	-.21	-.45	-.61	1.68	-.25	.81	.01
Gaming activities (out of 4)	.37	.51	.70	2.31	1.24	.26	.76
Gaming factor score	-.39	-.28	.00	1.16	.57	-1.10	-.11
Learning activities (out of 3)	1.08	.89	2.24	2.36	2.04	2.30	1.75
Learning factor score	-.21	-1.02	.92	.34	.27	.47	.17

Table 1 illustrates the average scores on the variables that stood at the basis of the clustering procedure. The restricted learners, young networkers and moderates have the shortest duration and narrowest breadth of use. Communication activities are taken up most by young networkers, all round explorers, intensive gamers and experienced networkers. Creative activities are taken up more than average by the all-round explorers, the intensive gamers and the experienced networkers. Gaming activities by the all-round explorers and the intensive gamers, the latter has the longest duration of use. Learning activities are taken up more by the moderates, all-round explorers, intensive gamers and the experienced networkers. For a more detailed description of the groups, see the report.

Table 2 Distribution of age, gender and highest level of parental education within groups of children (percentages)

	Restricted learners	Young networkers	Moderates	All-round explorers	Intensive gamers	Experienced networkers	All children
Boys	50%	45%	52%	62%	63%	33%	50%
Girls	50%	55%	48%	38%	37%	67%	50%
Average age	11.1	12.7	12.7	13.5	13.6	14.1	12.6
Primary education	20%	14%	15%	12%	11%	9%	14%
Secondary education	60%	64%	61%	65%	67%	63%	63%
Tertiary education	20%	22%	24%	23%	22%	29%	23%

Table 2 illustrates the distribution within gender and age groups. Girls and boys differ with regard to the likelihood of belonging to the all-round explorers (more likely boys), intensive gamers (more likely boys) and experienced networkers (more likely girls) groups. The youngest users are in the restricted learners group and the oldest in the all-round explorers, intensive gamers and experienced networkers groups. The educational level of the parents is highest in the experienced networkers group and lowest in the restricted learners group.

Opportunities: Classification of Countries

The distribution of the six clusters of children (described in the previous section) was used to create a country classification based on risk and harm. The following procedure was followed: for each country the proportion of children in each group was calculated, this data was used to run a hierarchical cluster analysis of countries as cases, squared Euclidian distance and furthest neighbour methods. A five cluster solution was the outcome based on the distance matrix and the dendrogram. The characteristics of the clusters are described in detail in the report.

Table 3 Percentage of children in different groups in each opportunities country cluster

Groups of children → Country clusters↓	Restricted learners	Young networkers	Moderate users	All round explorers	Intensive gamers	Experienced networkers
Young networkers	18%	**33%**	17%	11%	8%	12%
Diversity	16%	16%	**26%**	12%	**16%**	**15%**
Moderates	**25%**	13%	**32%**	10%	13%	8%
Restricted learners	**34%**	10%	22%	12%	8%	14%
Advanced	12%	**22%**	9%	**15%**	14%	**27%**
All countries	22%	17%	24%	12%	12%	13%

Note: **Bold** indicates which groups of children are relatively highly presented in each country cluster.
Base: All countries (N=25)

Methodology used in the Risk and Harm Section (2)

This section of the appendix describes the measures, the clustering procedure for the grouping of children and the clustering procedure for country classification used in the section of the report related to the risks and harm encountered online up by young Europeans.

Measures

Four key risky online activities and the level of harm experienced by these if these caused upset, were measures in the EU Kids Online II survey: seeing sexual images, sexting, meeting strangers and bullying.

Since the data for risk and harm were nested (i.e. only children who experienced a risk could experience the related harm) it was necessary to create 3 scales to estimate clusters of children in relation to risk and harm. One scale each was created for sexual images, meetings strangers and bullying. Other questions about risks (Including sexting) were asked only to 11+ year olds so were left out of the analysis.

The scales ran from 0 – 'no experience of the risk' to 6 – 'the child experienced risk online and was very upset'. The rest of the scale was generally divided as follows: 1 – risk experienced offline but not online, 2 - risk experienced online but not bothered, 3 - risk experienced, child bothered but not upset, 4 – risk experienced online and child was a bit upset, 5 – risk experienced online and child was fairly upset. For the grouping of children described below only children who had experienced the risk online (scores 2 through 6) were entered into the analysis.

For bullying '2' marked children who said they had been bullied online but did not answer the question on how upset they were (N=61) and '3' indicated children who said they were bullied online but had not been upset by this (1%, N=189), only 3% of children in Europe were bullied online and fairly or very upset (N=707).

For meeting strangers '1' indicated that they had made friends online but did not meet them offline, '2' that they had met them offline but that they were not bothered (6%, N=1587), and the rest of the scale was as above. Less than 0.5% of children in Europe met people offline they had first met online and were subsequently fairly or very upset (N=67) by this experience.

13% of European children (N=3228) had seen sexual images online, but only 2% (N=430) indicated being fairly or very upset by this experience.

Also included was a contact risk scale, a scale that reflects the propensity to give and share (personal) information, this scale ranged from 0 – none of the contact risks taken to 5 – all contact risks taken, 4% of children (N=837) had taken all 5 contact risks.

Table 4 Correlations between risk/harm scales

	Bullying scale	Sexual images	Meeting offline
Sexual images scale	-0.06**		
Meeting strangers scale	-0.01	-0.09**	
Contact risks scale	-0.01	-0.02	0.42**

Base: Children who have run at least one of the risks online (N=5,722).

Table 3 shows that effect sizes for correlations between the different risk and harm scales are relatively small with the exception of the correlation between meeting strangers offline and contact risks which makes sense in light of the necessity to give away personal information and interact on social media to be able to meet these persons in the first place.

Risk and harm: Groups of children

The clustering was done only on children who had experienced at least one of the risks online of the scales described in the previous section (i.e. scores 2 through 6 on the sexual images, bullying and meeting strangers offline scales) . This was done because otherwise there would just be two groups of children; one group that had not experienced any risk and the other who had experienced some form of risk. While this is a reflection of the reality, i.e. children can be divided into those who have not experienced any risks and those who have, this is not very informative if trying to cluster countries based on the *different types of risks and harms* run by children.

This approach created a default group of children who had not run any risk (N=19,420) and a clustering typology of children who had experienced at least one of the risks (offline or online) from seeing sexual images, meeting strangers or bullying (N=5,722). Across all countries only 2% of children (N=519) have run all these risks.

A two-step cluster approach was taken; this technique is designed to handle large datasets and is capable of handling both continuous and categorical variables. If the desired number of clusters is unknown, the SPSS TwoStep Cluster procedure indicates the most appropriate number of clusters automatically. This led to the following distribution of children across 3 different groups.

Table 5 Distribution of different risk/harm categories in groups of children (only children who have run at least one risk online).

		Sexual risks	Higher risk/harm	Contact risks	EU average
Bullying	not bullied	74%	0%	77%	59%
	offline but not online	24%	0%	21%	18%
	not or only a bit upset	2%	44%	2%	11%
	fairly/very upset	0%	56%	0%	12%
Sexual images	not seen sexual images	13%	52%	47%	35%
	sexual images offline but not online	7%	10%	10%	9%
	not or only a bit upset	68%	29%	41%	49%
	fairly/very upset	12%	9%	2%	8%
Meeting strangers	Not met anyone offline or online	60%	45%	13%	39%
	Made friends online but did not meet offline	32%	30%	28%	30%
	not or only a bit upset	8%	22%	57%	30%
	fairly/very upset	0%	3%	2%	1%
Contact risks	Average number out of 5	.85	1.95	3.16	1.97
	N=	2299	1250	2173	5,722

Base: Children who have run at least one of the risks online (N=5,722).

Table 4 shows that the highest level of upset (i.e. harm) for bullying can be found in the children in the higher risk and harm group, higher levels of upset in relation to meeting strangers can also be found here. The greatest number of contact risks are found in the contact risk group but this group also has relatively high risks in meeting strangers (although low upset). The sexual risk group clearly has the highest levels of experience with seeing online sexual images and the levels of upset are also relatively high, this group has the lowest levels of harm (i.e. upset) in the other risk categories however.

Table 6 Gender, age and parental level of education distribution over risk/harm groups of children.

	Sexual risks	Higher risk and harm	Contact risks	No risk
Boys	58%	41%	55%	50%
Girls	42%	59%	45%	50%
Average age	13.88	13.48	14.29	12.26
Primary education	14%	11%	10%	15%
Secondary education	58%	64%	64%	63%
Tertiary education	28%	25%	26%	22%

Base: All children (N=25,142).

Table 5 shows that boys are more likely to be found in the sexual and contact risk groups than girls but that girls are more likely to be found in the higher risk and harm group. The youngest group of children is the no risk group and the oldest the contact risk group. The level of education of the parents is highest for children who fall in the sexual risks group and lowest for those who are in the no risk group.

Risk/Harm: Classification of countries

The distribution of the three clusters of children was used to create a country classification based on risk and harm. The following procedure was applied: for each country the proportion of children in each cluster was calculated, this data was used to run a hierarchical cluster analysis of countries as cases, squared Euclidian distance and furthest neighbour methods. A four cluster solution was the outcome based on the distance matrix and the dendrogram. Details on the risk/harm country classification can be found in the main report.

Table 7 Percentage of children in different groups in each risk/harm country cluster

Groups of children→ Country cluster ↓	Sexual risks	Higher risk and harm	Contact risks	No risk
Higher risk and harm	12%	**7%**	**15%**	66%
Lower risk and harm	8%	4%	7%	**81%**
Sexual risks	**18%**	6%	10%	66%
All countries	11%	5%	10%	74%

Note: **Bold** indicates which groups of children are relatively highly represented in each country cluster.
Base: All countries (N=25)

Methodology used in the Parental Mediation Section (3)

This section of the appendix describes the measures, the clustering procedure for the grouping of children and the clustering procedure for country classification used in the section of the report related to parental mediation of young Europeans' online activities.

Measures

The EU Kids Online survey asked a number of questions about several types of parental mediation. Furthermore, matched questions were asked of the child and the parent most involved in the child's internet use. To reduce the number of original variables while at the same time retaining and generalising the maximum amount of information, we created three compound scales of parental mediation, based on nine original mediation scales asked of all children:

- *Active mediation* (ranging from 0 to 22): Active mediation of internet use (Children's questionnaire, CQ) + Active mediation of internet use (Parents' questionnaire, PQ) + Active mediation of internet safety (CQ) + Active mediation of internet safety (PQ)
- *Restrictive mediation* (ranging from 0 to 12): Restrictive mediation (CQ) + Restrictive mediation (PQ)
- *Monitoring and technical restrictions* (ranging from 0 to 12): Monitoring (CQ) + Monitoring (PQ) + Technical restrictions (PQ).

Handling of missing values: those individuals who could not use monitoring or technical restrictions (due to the child not using the internet at home) were assigned the value of zero on the scale of Monitoring and technical restrictions. The number of cases where missing values were replaced with the value 0 varied from 2045 to 2774, depending on the particular scale out of 3.

For those who had answered 'Don't know', the average values (calculated with country weights switched on) were used to replace missing data. The number of cases where the average was used to replace 'Don't know' ranged from 110 to 493.

Table 8: Correlations between mediation types

	Active mediation	Restrictive mediation
Restrictive mediation	0.21**	
Monitoring and technical restrictions	0.11**	0.07**

Base: All Children (N=25,142).

Table 6 shows that all mediation types are related. The smallest effect size was found in the correlation between monitoring and technical restrictions and restrictive mediation. This means that parents who apply one mediation technique are also likely to apply another mediation technique.

Parental mediation: groups of children

To cluster individual children, the three compound mediation scales (previously standardised to Z-scores) were used as input variables. Firstly, TwoStep Clustering, suitable for very large

datasets, was applied, but it resulted in six very poorly interpretable clusters. Secondly, we used K-Means Clustering, trying out solutions with different numbers of clusters. The 4-cluster solution was a reasonable fit and interpretable in light of the literature on parental mediation.

Table 9: Average standardised values of compound mediation scales in mediation clusters

	All-rounders	Active mediation preferred	Restrictive mediation preferred	Passive
Active mediation	0.89	0.33	0.09	-1.25
Restrictive mediation	0.43	-0.65	1.22	-0.69
Monitoring and technical restrictions	1.46	-0.11	-0.39	-0.79
N=	5583	7320	6350	5889

Table 7 shows that the parents of the all-rounders group of children are most likely to use all the different mediation strategies, while the parents of the passive mediation group are the least likely to use any of the strategies. The parents of other groups of children are very clearly clustered around active mediation for the active mediation group and restrictive mediation for the restrictive mediation group.

Table 10: Gender, age and parental level of education in mediation clusters

	All-rounders	Active mediation preferred	Restrictive mediation preferred	Passive
Girls	51%	51%	50%	46%
Boys	49%	49%	50%	54%
Average age	12.2	13.6	11.4	13.8
Primary education	8%	10%	17%	22%
Secondary education	66%	62%	61%	62%
Tertiary education	26%	29%	21%	16%

Table 8 shows that there are more boys in the passive mediation group and that the average age of children is also highest in this group. The youngest children fall in the restrictive mediation group. The children whose parents use mainly active mediation are also relatively older and in addition their parents have the highest levels of education. The lowest levels of education can be found for the restrictive and passive mediation groups.

Parental mediation: Classification of countries

The distribution of the four clusters of children (described above) was used to create a country classification based on parental mediation. The following procedure was followed: for each country the percentage of children in each cluster was calculated, this data was used to run a hierarchical cluster analysis of countries as cases, with squared Euclidian distance and furthest neighbour methods. More details on the characteristics of these clusters can be found in the main report.

Table 11 Percentage of children in different groups in each parental mediation country cluster

Groups of children → / Country clusters ↓	All-rounders	Active mediation preferred	Restrictive mediation preferred	Passive
Restrictive mediation	**24%**	21%	**32%**	23%
Passive	14%	30%	20%	**36%**
All-rounders	**25%**	**37%**	13%	**25%**
Active mediation	21%	**47%**	17%	15%
All countries	**22%**	**31%**	**24%**	**24%**

Note: **Bold** indicates which groups of children are relatively highly represented in each country cluster.
Base: All countries (N=25)

er