

The Classification of Welsh Medium Education

Secondary Schools

(1) INTRODUCTION

This paper provides a discussion of the categorization / classification of Secondary Schools regarding Welsh Language education. It concerns statistical analyses to see if WAG Secondary Schools data (from PLASC) naturally fits Schools into specific groups or types making a classification of Schools potentially empirically informed. The statistical tool to attempt this is called Cluster Analysis.

(2) USING PLASC DATA

The available PLASC Census (Pupil Level Annual School Census) derives from schools' returns in January 2108. (The 2018/2019 Census data will be available in July 2019). Fifteen PLASC Welsh language variables were used to compare Schools. **Appendix 1** provides fuller details. These variables are:

Variable 1: **WelshMediumSubjects**

Number of Subjects in a School taught through Welsh / Bilingually

NB: The range is from 0 to 10 subjects (Secondary Schools only)

Variable 2: **WelshGovernors**

% Welsh Governors in a School who are Welsh speaking

Variable 3: **CLASSKS3MAIN**

% KS3 Classes in a School where Welsh is the sole or **main** medium of instruction.

Variable 4: **CLASSKS3PART**

% KS3 Classes in a School where Welsh is used as a teaching medium for **part** of the curriculum (i.e. less than half)

Variable 5: **CLASSKS4MAIN**

% KS4 Classes in a School where Welsh is the sole or **main** medium of instruction.

Variable 6: **CLASSKS4PART**

% KS4 Classes in a School where Welsh is used as a teaching medium for **part** of the curriculum (i.e. less than half)

Variable 7: **TEACHER_FT_TCTO**

% Full Time Teachers in a School who are:

TC Qualified teachers teaching Welsh as a first language *plus*

TO Qualified teachers teaching other subjects through the medium of Welsh

Variable 8: TEACHER_PT_TCTO

% Part Time Teachers in a School who are:

TC Qualified teachers teaching Welsh as a first language *plus*

TO Qualified teachers teaching other subjects through the medium of Welsh

Variable 9: TEACHER_FT_NW

% Full Time Teachers in a School who are:

NW Qualified teachers able to teach Welsh or through the medium of Welsh, **but not doing so.**

NB: The Part-Time Teacher equivalent of this variable had only 23 teachers, so has not been reported although the initial analyses included this.

Variable 10: PUPIL_FLUENT

% Pupils in a School who are Fluent in Welsh

NB: The percentage calculation excludes 'information refused'.

Variable 11: PUPIL_NOTFLUENT

% Pupils in a School who can speak Welsh but not fluently

NB: The percentage calculation excludes 'information refused'.

Variable 12: WELSHATHOME

% Pupils in a School who can speak Welsh at home.

NB: The percentage calculation excludes 'not applicable'.

Variable 13: STUDYOFWELSHL1

% Pupils in a School who are taught Welsh as a first language

The percentage calculation includes 'disapplied from the National Curriculum' on the basis that ALL pupils create a language character of the school.

Variable 14: STUDYOFWELSHL2

% Pupils in a School who are taught Welsh as a second language

The percentage calculation includes 'disapplied from the National Curriculum' on the basis that ALL pupils create the language ethos of the school.

Variable 15: WELSHMEDIUM

% Pupils in a School who study any subject other than Welsh (1st or 2nd language) through the medium of Welsh.

A PLASC variable that was not entered into the Cluster analyses was a School's self-classification as a Welsh medium, English medium or Bilingual medium School using categorization system in the October 2007 WAG paper: '*Defining schools according to Welsh medium provision*'.

(<https://beta.gov.wales/sites/default/files/publications/2018-02/defining-schools-according-to-welsh-medium-provision.pdf>).

This variable is considered later in this paper for comparisons with the final Clusters generated in this Project. How does the current (2007) PLASC classification compare with the clusters found in these analyses? The coding for this variable is:

WM – Welsh Medium

EM – English Medium

EW – English (with significant Welsh)

BILINGUAL (with coding AB, BB, CB, CH).

AB – A Bilingual:

At least 80% of subjects apart from English and Welsh are taught only through the medium of Welsh to all pupils. One or two subjects are taught to some pupils in English or in both languages.

BB – B Bilingual

At least 80% of subjects (excluding Welsh and English) are taught through the medium of Welsh but are also taught through the medium of English.

CB – C Bilingual

50 - 79% of subjects (excluding Welsh and English) are taught through the medium of Welsh but are also taught through the medium of English.

CH – Ch Bilingual

All subjects, except Welsh and English taught to all pupils using both languages.

(3) CHOICE of DIMENSIONS

All 15 PLASC Welsh language dimensions were entered into the cluster analyses. This makes maximal use of the data to profile each school. However, there are risks. Just as in cooking, the inclusion or exclusion of individual ingredients affects the outcome, a question for this Project is what variables should be entered or excluded?

Are all variables included in these analyses as important as each other? For example, '*the percentage of Governors who are Welsh speaking*' may be an 'indicator measure' of Welsh community involvement but is it important to include? Later on in this paper there is a list of the variables that are relatively more important in creating the different clusters (Discriminant Analysis). This possible allows WAG to allocate Schools to clusters based on these variables alone.

(4) LIMITATIONS of DIMENSIONS

If the inclusion or exclusion of individual ingredients affects the outcome, then it is also important to list what variables might be included in future analyses if the relevant Welsh language data becomes available:

- (1) KS3 Teacher Assessment data, GCSE results, Welsh language progress of pupils across Key Stages.
- (2) *Derived variables* if pupil-level data can be appropriately aggregated (e.g. Pupils NOT Speaking Welsh at Home but who are studying Welsh as a L1 / through the *medium* of Welsh). Since these variables are important in reaching WAG's million Welsh speakers target, their inclusion in future analyses could be valuable.
- (3) Languages of feeder school **communities**. See a previous paper on how this could be achieved. As it stands, apart from 'Welsh speaking Governor' there is no variable that locates if a School is placed in a more Welsh language community / catchment area or is placed in a more Anglicized area. This would add in the locational difference between some Welsh medium / bilingual schools.

Ideally, PLASC would add a question for each School such as: How would you describe your school catchment area(s) in terms of its resident Welsh -speaking population?

70% and over Welsh speakers

50% to 69% Welsh speakers

30% to 49% Welsh speakers

10% to 29% Welsh speakers

Under 10% Welsh speakers

- (4) Provision in Welsh in Primary and Secondary schools for **ALNs**.
- (5) Languages of learning / teaching resources, curriculum materials, computer usage etc..
- (6) School language ethos e.g. announcements, displays on walls in classrooms and corridors, playground language(s), languages of Teacher Assistants and Ancillary staff.
- (7) Language of Further Study.
- (8) Leaver Employment.

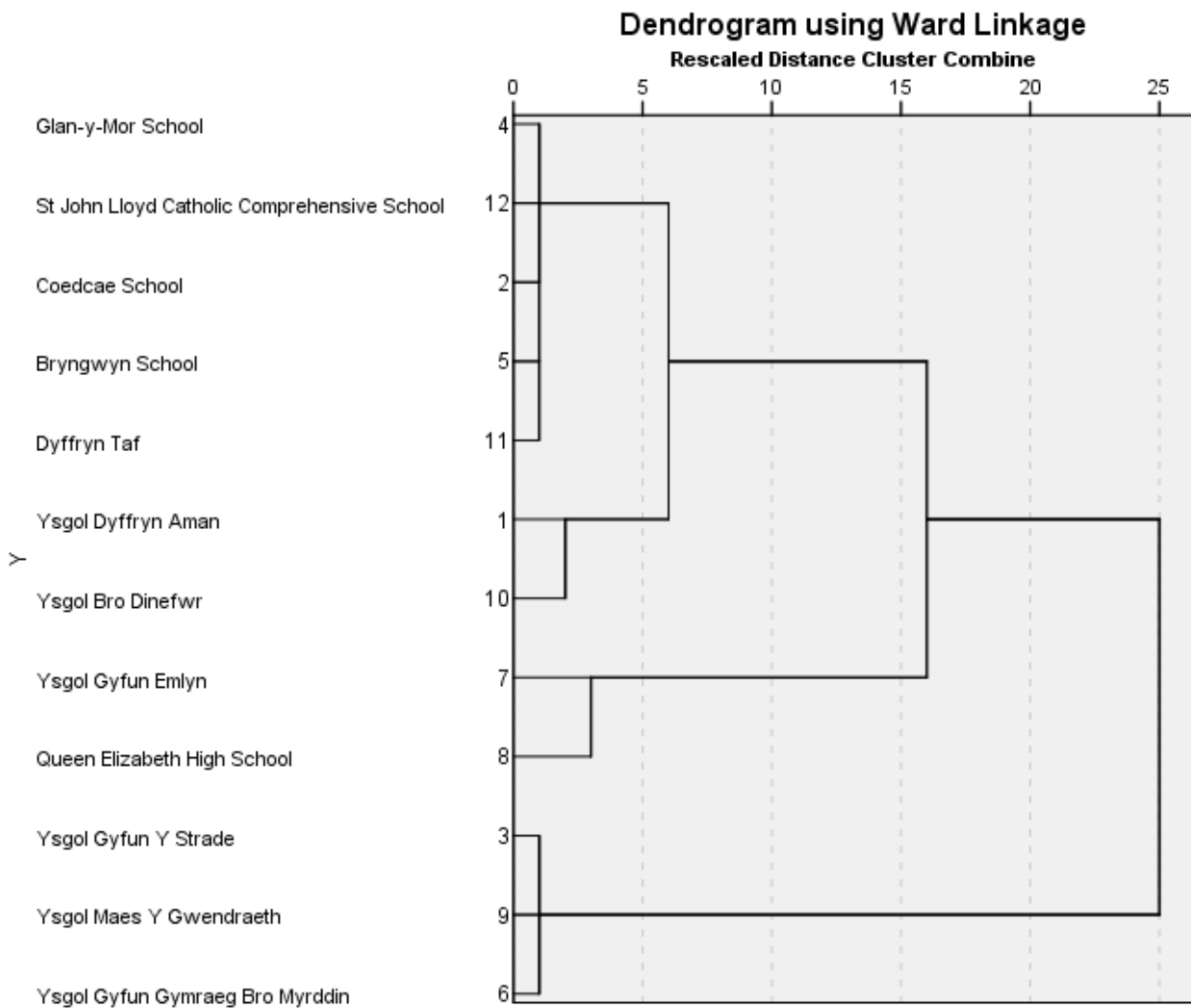
(5) CHOICE OF STATISTICAL ANALYSES

The statistical procedure used for classification of Schools is called Cluster Analysis. In **Appendix 2** of this paper, there is a more detailed and technical discussion of this technique (including its use in this Project), and its limitations.

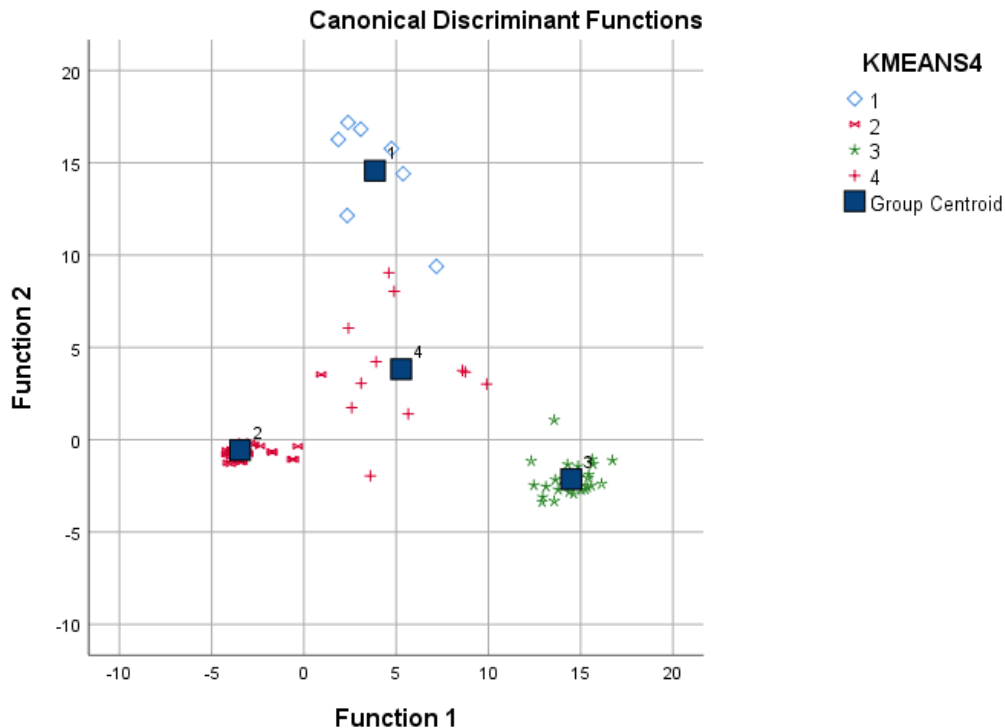
For the purposes of this project, Cluster Analysis takes each school and finds out which group they best fit. The final groups are called Clusters. Just as stars are not randomly distributed in the sky but form clusters, so schools are more or less like each other with clusters of similar schools being found.

One way to think of cluster analysis is as follows. The computer program begins by locating the two Schools that are most similar using the 15 PLASC dimensions. Then it finds a school that is either the closest to that pair or to another fourth school.

That linking procedure is continuous until every school is included and just two clusters remain. The decision is then (with the aid of a dendrogram – see below for an example based only on the Carmarthen Secondary schools) to find the number of clusters that best fit the data (e.g. three clusters in the Dendrogram below).



Below is a graph of the four-cluster solution showing how these clusters are distanced in space. Most of the 'dots / circles' representing schools are placed underneath the squares (the squares represent the centre of the cluster).



(6) ANALYSES

(i) **Data Checks.** The data provided by WAG was initially checked for outliers and possible issues. The PLASC data already had an 'extensive set of validation checks' negotiating 'errors and queries'. The data was found to be clean except for (i) some Welsh medium schools having a zero for the number of 'Welsh Medium Subjects' (ii) two schools having an unrealistic number (64 and 39) of teachers who could teach though Welsh but were not doing so, (iii) and no data from Ysgol Gyfun Gymraeg Glantaf for its pupils being fluent (or not) in Welsh.

Ysgol Gyfun Gwent Is Coed (680/4060) has 150 KS3 pupils (so far). Hence, this growing school was not entered into analyses for lack of comparative data. However, its categorization could be straightforward as it is a Welsh medium school (see later).

(ii) **Cluster Analyses** were conducted using the 15 variables in this project. All variables were entered into analyses because they all relate to the Welsh language / bilingual nature of the school. While not all variables are eventually needed to separate out schools into clusters (this is discussed at the end of the paper), conceptually it was decided to use the full set of 15 PLASC variables relevant to this project. That is more defensible than entering a subset of variables which could be debated and indeed dubious.

A typical approach is to first perform a hierarchical method to define the number of clusters followed by use the K-Means procedure to allocate Schools to Clusters. A K-Means analysis requires the number of clusters to be given at the start of the program, hence the need for an initial hierarchical classification with a dendrogram to suggest that number.

The K-Means procedure tends to result in a superior solution to a hierarchical analysis. This is partly as it allows individual schools to be constantly relocated in a cluster for best fit, whereas a hierarchical solution does not allow change from one cluster to another during its process. Technical details of the computer analyses for this dataset are found in **Appendix 2**.

The hierarchical analysis Dendrogram for this dataset of 15 variables and 194 Secondary Schools (see **Appendix 3**) suggested a cluster solution of a minimum of three clusters and a maximum of six clusters, with five clusters being possibly the 'best fit'. A Scree graph from the hierarchical analysis of the coefficients by the combining of clusters (see **Appendix 4**) suggests four. However, later this paper includes the interpretation of six clusters.

Experience suggests that human interpretation of cluster solutions is important. That is, neither a dendrogram nor a scree graph are rarely unambiguous as to the optimal number of clusters, and so examining several solutions is sensible.

In this case, three to six clusters were examined across the 15 variables. The results of these analyses are now reported.

(7) RESULTS FOR ALL 194 SCHOOLS

The first table shows to which cluster each School was allocated on four separate analyses: 3 clusters, 4 clusters, 5 clusters, 6 clusters.

For example, **Ysgol Syr Thomas Jones** was allocated to Cluster 1 on all of the four different analyses. **Ysgol Gyfun Llangefni** is allocated to Cluster 3 on all analyses. Some Schools move a little, especially as more clusters (e.g. moving from three clusters to four clusters) are added. For example, Ysgol Uwchradd Caergybi is placed in Cluster 2 (when there are three, four and five clusters), but moves to cluster 4 when there are six clusters.

To help understand the table, colour coding has been added.

Yellow coding denotes Welsh medium Schools.

Blue coding denotes English medium Schools.

White coding denotes Schools which change clusters (e.g. Ysgol Uwchradd Caergybi, Ysgol David Hughes and Ysgol Uwchradd Bodedern) and 'Bilingual' Schools that are not Welsh medium or English medium and change in their allocation. This becomes a key focus later.

Red coding denotes a Welsh medium school that moves when there are 6 clusters as with Ysgol Dyffryn Ogwen Bethesda. An important discussion later (when these Schools are profiled across the 15 variables) is whether this 6th Cluster is important to keep (or to merge with other Welsh medium Schools).

The numbering system in the table below is not meaningful, but simply reflects the cluster number allocated by the computer for different analyses. Since the three, four, five and six cluster solutions are from separate (independent) analyses, the numbering system the computer allocates can vary. Hence, the colour coding is important in understanding the table.

The final column provides a School's self-classification as a Welsh medium, English medium or Bilingual medium School using the October 2007 WAG paper: 'Defining schools according to Welsh medium provision' (see above for the codes). The degree of matching between the 'old' 2007 classification and the findings of this paper is considered later (see Appendix 6).

Enw Ysgol Uwchradd	3 CLUSTERS	4 CLUSTERS	5 CLUSTERS	6 CLUSTERS	OLD CLASS
Ysgol Syr Thomas Jones	1	1	1	1	BB
Ysgol Uwchradd Caergybi	2	2	2	4	EW
Ysgol Gyfun Llangefni	3	3	3	3	BB
Ysgol David Hughes	1	4	1	1	BB
Ysgol Uwchradd Bodedern	3	4	4	3	BB
Ysgol Dyffryn Ogwen Bethesda	3	3	3	6	AB
Ysgol Botwnnog	3	3	3	6	AB
Ysgol Brynrefail	3	3	3	3	AB
Ysgol Dyffryn Nantlle	3	3	3	3	AB
YSGOL EIFIONYDD	3	3	3	6	AB
Ysgol Y Moelwyn	3	3	3	6	AB
YSGOL UWCHRADD TYWYN	1	4	4	1	AB
Ysgol Y Berwyn	3	4	4	6	AB
Ysgol Ardudwy	3	3	3	3	AB
Ysgol Friars	1	2	2	4	EW
Ysgol Tryfan	3	3	3	3	AB
Ysgol Syr Hugh Owen	3	3	3	6	AB
Ysgol Glan y Mor	3	4	4	6	AB
Ysgol John Bright	2	2	2	4	EM
Ysgol Aberconwy	2	2	2	4	EM
YSGOL DYFFRYN CONWY	1	4	1	1	BB
YSGOL Y CREUDDYN	3	3	3	6	WM
Ysgol Emrys ap Iwan	2	2	2	2	EM
Eirias High School	2	2	2	4	EM
Ysgol Bryn Elian	2	2	2	2	EM
Rhyl High School	2	2	2	2	EM
Prestatyn High School	2	2	2	2	EM
Ysgol Uwchradd Glan Clwyd	3	3	3	6	WM
Denbigh High School	2	2	2	2	EM
Ysgol Dinas Bran	1	1	5	5	BB
Ysgol Brynhyfryd	3	4	4	3	BB
Blessed Edward Jones High Sch	2	2	2	2	EM
Hawarden High School	2	2	2	2	EM
Alun School	2	2	2	2	EM
Elfed High School	2	2	2	2	EM
Ysgol Treffynnon	2	2	2	2	EM
St. David's High School	2	2	2	2	EM
Castell Alun High School	2	2	2	2	EM

Maes Garmon	3	3	3	3	WM
Flint High School	2	2	2	2	EM
Connah's Quay High School,	2	2	2	2	EM
Argoed School	2	2	2	2	EM
St. Richard Gwyn Catholic High School	2	2	2	2	EM
Grango	2	2	2	2	EM
YSGOL MORGAN LLWYD	3	3	3	3	WM
Ysgol Bryn Alyn	2	2	2	2	EM
Darland High School	2	2	2	2	EM
Ysgol Rhiwabon	2	2	2	2	EM
Rhosnesni High School	2	2	2	2	EM
Ysgol Clywedog	2	2	2	2	EM
St Joseph's Catholic and Anglican High	2	2	2	2	EM
THE MAELOR SCHOOL	2	2	2	2	EM
Ysgol Uwchradd Caereinion High School	3	4	4	3	BB
Llanfyllin High School	2	2	2	2	CB
LLANIDLOES HIGH SCHOOL	1	1	5	5	EW
Newtown High School	2	2	2	2	EM
WELSHPOOL HIGH SCHOOL	2	2	2	2	EM
LLANDRINDOD HIGH SCHOOL	2	2	2	2	EM
BUILTH WELLS HIGH SCHOOL	2	2	2	2	CB
Ysgol Maesydderwen	2	2	2	4	EM
BRECON HIGH SCHOOL	2	2	2	2	EW
Gwernyfed High School	2	2	2	2	EM
Crickhowell High School	2	2	2	4	EM
GYFUN ABERAERON COMPREHENSIVE	1	4	4	1	CB
Ysgol Uwchradd Aberteifi	1	1	1	1	CB
Penglais School	2	2	2	4	EM
YSGOL GYFUN PENWEDDIG	3	3	3	6	AB
Ysgol Bro Gwaun	1	1	5	5	EW
Ysgol Dewi Sant	2	2	2	4	EM
Ysgol Greenhill School	2	2	2	4	EM
Pembroke School	2	2	2	2	EM
SIR THOMAS PICTON	2	2	2	2	EM
Milford Haven School	2	2	2	2	EM
Ysgol y Preseli	3	3	3	3	WM
Tasker-Milward V.C. School	2	2	2	2	EM
Ysgol Dyffryn Aman	1	4	4	1	BB
Coedcae School	2	2	2	2	EM
Ysgol Gyfun Y Strade	3	3	3	3	AB
Glan-y-Mor School	2	2	2	4	EM
Bryngwyn School	2	2	2	2	EM
Ysgol Gyfun Gymraeg Bro Myrddin	3	3	3	3	WM
Ysgol Gyfun Emlyn	1	1	5	5	EW
Queen Elizabeth High School	1	1	5	5	EW
Ysgol Maes Y Gwendraeth	3	3	3	6	AB

Ysgol Bro Dinefwr	1	4	4	3	BB
Dyffryn Taf	2	2	2	2	EM
St John Lloyd Catholic Comprehensive School	2	2	2	4	EM
Cefn Hengoed	2	2	2	2	EM
OLCHFA SCHOOL	2	2	2	2	EM
MORRISTON COMPREHENSIVE	2	2	2	2	EM
PENTREHAFOD SCHOOL	2	2	2	2	EM
BISHOP GORE SCHOOL	2	2	2	2	EM
Penyrheol Comprehensive School	2	2	2	2	EM
Gowerton Comprehensive School	2	2	2	2	EM
BISHOPSTON COMPREHENSIVE	2	2	2	2	EM
Pontarddulais Comprehensive School	2	2	2	2	EM
Ysgol Gyfun Gwyr	3	3	3	3	WM
BIRCHGROVE	2	2	2	2	EM
Dylan Thomas Community School	2	2	2	2	EM
Ysgol Gyfun Gymraeg Bryn Tawe	3	3	3	3	WM
Bishop Vaughan School	2	2	2	2	EM
Cymer Afan Comprehensive School	2	2	2	2	EM
Dyffryn School	2	2	2	2	EM
Cefn Saeson Comprehensive School	2	2	2	4	EM
Cwmtawe Community School	2	2	2	2	EM
Llangatwg Community School	2	2	2	2	EM
Dwr Y Felin Comprehensive School	2	2	2	2	EM
St Joseph's RC School and 6th Form Centre	2	2	2	2	EM
Cynffig Comprehensive	2	2	2	2	EM
Bryntirion Comprehensive.	2	2	2	2	EM
Maesteg Comprehensive School	2	2	2	2	EM
PENCOED COMPREHENSIVE	2	2	2	2	EM
Brynteg School	2	2	2	2	EM
PORTHCAWL COMPREHENSIVE SCHOOL	2	2	2	2	EM
Ysgol Gyfun Gymraeg Llangynwyd	3	3	3	3	WM
Coleg Cymunedol Y Dderwen	2	2	2	2	EM
Archbishop McGrath Catholic School	2	2	2	2	EM
LLANTWIT MAJOR SCHOOL	2	2	2	2	EM
BARRY COMPREHENSIVE SCHOOL	2	2	2	2	EM
BRYN HAFREN COMPREHENSIVE SCHOOL	2	2	2	2	EM
COWBRIDGE COMPREHENSIVE SCHOOL	2	2	2	2	EM
St Cyres Comprehensive School	2	2	2	2	EM
St Richard Gwyn Catholic High School	2	2	2	2	EM
Stanwell School	2	2	2	2	EM
Bryncelynnog Comprehensive School	2	2	2	2	EM
The Pontypridd High School	2	2	2	2	EM
Hawthorn High School	2	2	2	2	EM
Mountain Ash Comprehensive School	2	2	2	4	EM
Ysgol Gyfun Garth Olwg	3	3	3	3	WM
TONYREFAIL SCHOOL	2	2	2	2	EM

Treorchy Comprehensive School	2	2	2	4	EW
Ferndale Community School	2	2	2	2	EM
Porth County Community	2	2	2	2	EM
Tonypandy Community College	2	2	2	2	EM
Y Pant Comprehensive	2	2	2	2	EM
Ysgol Gyfun Cwm Rhondda	3	3	3	3	WM
Ysgol Gyfun Rhydywaun	3	3	3	3	WM
Aberdare Community School	2	2	2	2	EM
Cardinal Newman R.C.	2	2	2	2	EM
St.John Baptist High School	2	2	2	2	EM
Afon Taf High School	2	2	2	2	EM
PEN-Y-DRE HIGH SCHOOL	2	2	2	2	EM
Cyfarthfa High School	2	2	2	2	EM
Bishop Hedley High School	2	2	2	2	EM
Newbridge School	2	2	2	2	EM
Blackwood Comprehensive School	2	2	2	2	EM
St Cenydd School	2	2	2	2	EM
Risca Community Comprehensive	2	2	2	2	EM
St Martin's School	2	2	2	2	EM
HEOLDDU COMPREHENSIVE SCHOOL	2	2	2	2	EM
Lewis School Pengam	2	2	2	2	EM
Lewis Girls' Comprehensive School	2	2	2	2	EM
Bedwas High School	2	2	2	4	EM
Ysgol Gyfun Cwm Rhymni	3	3	3	3	WM
Islwyn High School	2	2	2	2	EM
Cwmcarn High School	2	2	2	2	EM
Tredegar Comprehensive School	2	2	2	2	EM
Brynmawr Foundation School	2	2	2	2	EM
Croesyceiliog School	2	2	2	2	EM
Abersychan Comprehensive	2	2	2	2	EM
West Monmouth School	2	2	2	2	EM
YSGOL GYFUN GWYNLLYW	3	3	3	3	WM
Cwmbran High School	2	2	2	2	EM
St. Albans R.C. High School	2	2	2	2	EM
Monmouth Comprehensive School	2	2	2	2	EM
King Henry VIII Comprehensive	2	2	2	2	EM
Chepstow Comprehensive School	2	2	2	2	EM
Caldicot School	2	2	2	2	EM
St Julian's School	2	2	2	2	EM
The John Frost School	2	2	2	2	EM
Llanwern High School	2	2	2	2	EM
Newport High School	2	2	2	2	EM
Lliswerry High School	2	2	2	2	EM
Bassaleg School	2	2	2	2	EM
Caerleon Comprehensive School	2	2	2	2	EM
St Joseph's RC High School	2	2	2	2	EM

CARDIFF HIGH SCHOOL	2	2	2	2	EM
Willows High School	2	2	2	2	EM
Fitzalan High School	2	2	2	2	EM
Cantonian High School	2	2	2	2	EM
Llanishen High School	2	2	2	2	EM
Cathays High School	2	2	2	2	EM
Radyr Comprehensive School	2	2	2	2	EM
Ysgol Gyfun Gymraeg Glantaf	3	3	3	6	WM
Ysgol Gyfun Gymraeg Plasmawr	3	3	3	3	WM
Ysgol Gyfun Gymraeg Bro Edern	3	3	3	3	WM
Cardiff West Community High School	2	2	2	4	EM
Eastern High	2	2	2	2	EM
St. Illtyd's Catholic High School	2	2	2	2	EM
Mary Immaculate High School	2	2	2	2	EM
Bishop of Llandaff Church in Wales High School	2	2	2	2	EM
St Teilo's C-in-W High School	2	2	2	2	EM
CORPUS CHRISTI CATHOLIC HIGH SCHOOL	2	2	2	2	EM
Whitchurch High School	2	2	2	2	EM

To better understand the difference between the clusters in each solution, the scores for each School are given in **APPENDIX 5**.

There follows an outline of the three, four, five and six cluster solutions in turn. The label of the clusters are temporary and for discussion.

(8)THE CHOICES

A decision for the team appears to be whether Secondary Schools in Wales fit into three, four, five or six groups (clusters). Below is an **overview** of the four choices, followed by tables with the detail.

Choice 1. Three Clusters of Schools without sub division: Welsh Medium, English Medium, Bilingual.

Choice 2. Four Clusters of Schools with a sub division of Bilingual Schools:

1. Welsh Medium
2. English Medium
3. Bilingual subdivided into:
 - 3A 'Bilingual with relatively **more** Welsh medium teaching and learning'.
 - 3B 'Bilingual with relatively **less** Welsh medium teaching and learning'.

Choice 3. Five Clusters of Schools with a threefold sub division of Bilingual Schools:

1. Welsh Medium

2. English Medium
3. Bilingual subdivided into:
 - 3A 'Bilingual with relatively **more** Welsh medium teaching and learning'.
 - 3B 'Bilingual with relatively **less** Welsh medium teaching and learning'.
 - 3C 'Schools with more Welsh medium teaching and learning at **KS3 than KS4**'.

Choice 4. Six Clusters of Schools with a threefold sub division of Bilingual Schools and a twofold division of Welsh Medium Schools:

1. English Medium
2. Bilingual subdivided into:
 - 3A 'Bilingual with relatively **more** Welsh medium teaching and learning.
 - 3B 'Bilingual with relatively **less** Welsh medium teaching and learning.
 - 3C 'Schools with more Welsh medium teaching and learning **KS3 than KS4**at.
3. Welsh Medium subdivided into:
 - 4A Welsh Medium 'A' Schools
 - 4B Welsh Medium 'B' Schools.
 (see later for the differences between 'A' and 'B').

(9) CLUSTER ANALYSIS RESULTS

3 CLUSTER SOLUTION	Bilingual Schools	English Medium Schools	Welsh Medium Schools
Average Number of Welsh Medium Subjects	5	0	8
% Welsh Governors	56	7	82
%CLASSKS3_SOLE/MAIN	10	0	88
%CLASSKS3_PART	69	0	3
%CLASSKS4_SOLE/MAIN	11	0	88
%CLASSKS4_PART	60	0	5
%TEACHER_FT_TCTO	33	1	72
%TEACHER_PT_TCTO	9	0	18
%TEACHER_FT_NW	7	2	3
%PUPIL_FLUENT	47	6	83
%PUPIL_PART FLUENT	53	78	12
%WELSH_AT_HOME	31	7	52
%STUDY OF WELSH L1	55	1	95
%STUDY OF WELSH L2	42	96	5

%WELSH MEDIUM	27	1	67
NUMBER OF SCHOOLS IN CLUSTER	14	145	35

4 CLUSTER SOLUTION	'ENGLISH BILINGUAL'	English Medium	Welsh Medium	'WELSH BILINGUAL'
Average Number of Welsh Medium Subjects	6	0	9	4
% Welsh Governors	37	7	81	86
%CLASSKS3_SOLE/MAIN	8	0	92	38
%CLASSKS3_PART	90	0	2	31
%CLASSKS4_SOLE/MAIN	6	1	96	26
%CLASSKS4_PART	85	0	0	35
%TEACHER_FT_TCTO	29	1	74	49
%TEACHER_PT_TCTO	6	0	18	18
%TEACHER_FT_NW	9	2	3	4
%PUPIL_FLUENT	41	6	85	63
%PUPIL_PART FLUENT	59	78	8	37
%WELSH_AT_HOME	27	7	51	48
%STUDY OF WELSH L1	42	1	99	75
%STUDY OF WELSH L2	55	96	1	25
%WELSH MEDIUM	10	1	68	56
NUMBER OF SCHOOLS IN CLUSTER	7	146	30	11

5 CLUSTER SOLUTION	Bilingual KS3>KS4	English Medium	WELSH BILINGUAL	Welsh Medium	ENGLISH BILINGUAL
Average Number of Welsh Medium Subjects	5	0	4	9	6
% Welsh Governors	83	7	85	81	23
%CLASSKS3_SOLE/MAIN	0	0	46	92	11
%CLASSKS3_PART	97	0	17	2	86
%CLASSKS4_SOLE/MAIN	12	1	30	96	3
%CLASSKS4_PART	59	0	31	0	93
%TEACHER_FT_TCTO	59	1	46	74	19
%TEACHER_PT_TCTO	18	0	17	18	3
%TEACHER_FT_NW	2	2	3	3	12
%PUPIL_FLUENT	66	6	62	85	32
%PUPIL_PART FLUENT	34	78	38	8	68
%WELSH_AT_HOME	49	7	46	51	21
%STUDY OF WELSH L1	84	1	73	99	25
%STUDY OF WELSH L2	16	96	26	1	71
%WELSH MEDIUM	42	1	51	68	12
NUMBER OF SCHOOLS IN CLUSTER	4	146	9	30	5

6 CLUSTER SOLUTION	English Medium	English Bilingual	Welsh Medium B	Welsh Bilingual	Mid Bilingual	Welsh Medium
N Welsh Medium Subjects	0	1	9	6	3	8
% Welsh Governors	5	20	74	23	83	97
%CLASSKS3_SOLE/MAIN	0	1	86	11	8	88
%CLASSKS3_PART	0	4	1	86	70	8
%CLASSKS4_SOLE/MAIN	0	3	83	3	16	90
%CLASSKS4_PART	0	2	4	93	46	8
%TEACHER_FT_TCTO	1	2	74	19	45	64
%TEACHER_PT_TCTO	0	1	13	3	15	28
%TEACHER_FT_NW	1	12	4	12	3	2
%PUPIL_FLUENT	5	10	84	32	59	79
%PUPIL_PART FLUENT	78	77	12	68	41	12
%WELSH_AT_HOME	7	12	41	21	41	72
%STUDY OF WELSH L1	1	4	93	25	80	98
%STUDY OF WELSH L2	96	91	7	71	19	2
%WELSH MEDIUM	1	2	58	12	41	82
NUMBER OF SCHOOLS IN CLUSTER	129	17	24	5	7	12

(10) Understanding Results from the three, four, five and six cluster solutions

The tables above suggest that there are two relatively solid clusters: English medium schools who number over 128, and Welsh medium schools who number up to 35. The variations between the three cluster to six cluster solutions are mostly the Bilingual Schools that do not fall consistently into these two groups.

- (1) The **three cluster solution** shows a separation between Welsh medium schools and Bilingual schools (see the three cluster solution table above), plus English medium schools as the third group. Therefore, one possibility is to choose a three cluster solution (Welsh medium; Bilingual; English medium).

This parallels the **Basque** three-fold categorization: Spanish medium *Model A Schools* where Spanish as the language of teaching and Basque is taught as a second language. *Model B Schools* have both Basque and Spanish as languages of teaching and are therefore Bilingual schools, and *Model D Schools* with Basque as the language of instruction.

However, the four and five cluster solutions (see above) separate out 'bilingual schools' into (admittedly small-sized) two and three subdivisions (the four and five cluster solution respectively). In terms of '**progression**', separating these bilingual schools may be helpful, giving some schools a target of reaching a 'higher' grouping.

- (2) The **four cluster solution** has solid Welsh medium and English medium clusters, plus two Bilingual School clusters which differ in their 'bilingualism'. A provisional label for these two clusters is 'Welsh Bilingual' and 'English Bilingual'. While there are just seven (7) schools in the 'English Bilingual' cluster and 11 schools in the 'Welsh Bilingual' cluster, the four cluster table (see above and below) shows a definite difference between these two clusters to make the difference conceptually important despite the small numbers in each cluster.

For example, the 'Welsh Bilingual' cluster has 38% of KS3 classes where Welsh is the **main** medium compared with 8% in the 'English Bilingual' cluster. The 'Welsh Bilingual' cluster has 31% of KS3 classes where Welsh is the **part** medium compared with 90% in the 'English Bilingual' cluster. The same pattern is found at KS4.

The four cluster solution shows that the 'Welsh Bilingual' cluster has different percentages compared with the 'English Bilingual' cluster on the following 12 variables and therefore seems an enhancement on the three cluster solution:

4 CLUSTER SOLUTION	'ENGLISH BILINGUAL'	'WELSH BILINGUAL'
% Welsh Governors	37	86
%CLASSKS3_SOLE/MAIN	8	38
%CLASSKS3_PART	90	31
%CLASSKS4_SOLE/MAIN	6	26
%CLASSKS4_PART	85	35
%TEACHER_FT_TCTO	29	49
%PUPIL_FLUENT	41	63
%PUPIL_PART FLUENT	59	37
%WELSH_AT_HOME	27	48
%STUDY OF WELSH L1	42	75
%STUDY OF WELSH L2	55	25
%WELSH MEDIUM	10	56
NUMBER OF SCHOOLS IN CLUSTER	7	11

Therefore, one possibility is to decide on a **four cluster solution** either (i) leaving the four clusters as they are, or (ii) nesting the 'Bilingual School' cluster with a subdivision of 'Welsh Bilingual' and 'English Bilingual'.

- (3) **The five cluster solution** has Welsh medium and English medium clusters which do not change from the four cluster solution. However, there are now three Bilingual School clusters which differ in their 'bilingualism', numbering 18 schools in all.

The table below shows how the three 'bilingual' cluster differ.

The 'Welsh Bilingual' cluster has changed from the four cluster solution to having a higher percentage of KS3 classes with Welsh as the **main** medium of study, and less KS3 classes with Welsh as the **part** medium of study. Also, the 'Welsh Bilingual' cluster has a higher percentage of KS3 and KS4 classes with Welsh as the main medium of study compared with the other two 'bilingual' clusters.

The 'English Bilingual' cluster has changed from the four cluster solution to having **less** pupils learning Welsh as a L1 and **more** as a L2.

The provisional label for the other cluster is 'BilingualKS3>KS4'. It has many similarities with the Welsh medium cluster. However, its percentage of teachers who can teach through Welsh is high (59% compared with 46% in the 'Welsh medium' cluster). Also, its percentage of Pupils studying Welsh as a first language is relatively high.

The 'BilingualKS3>KS4' cluster has zero schools where KS3 Classes Welsh as a Sole/Main Medium, with 97% KS3 Classes Welsh as a Part Medium. KS4 Classes are similar except 12% of classes are

Welsh as the main medium, 59% with Welsh as a part medium, (and the remaining 29% are English medium classes). Further comparisons are provided in the table below.

To summarize: the **five cluster solution** contains (apart from Welsh medium and English medium schools) three Bilingual School clusters. These three clusters separate the Welsh Bilingual Schools from the English Bilingual Schools, but also adds on a new cluster of four schools. These four schools appear to have the teacher and pupil potential for increased Welsh medium teaching but currently appear to move to a more English medium approach in KS4. This is portrayed in the table below.

This provides a choice of either (i) leaving the five clusters as they are or (ii) a Bilingual School cluster with a threefold subdivision of 'Welsh Bilingual', 'English Bilingual' and 'Bilingual KS3>KS4 Cluster'.

Variable	'Bilingual KS3>KS4' Cluster	'Welsh Bilingual' Cluster	'English Bilingual' Cluster
Number of Schools in Cluster	4	9	5
Percentage of Welsh speaking Governors	83	85	23
Number of Subjects taught through Welsh / Bilingually	5	4	6
Percentage KS3 Classes Welsh as Sole/Main Medium	0	46	11
Percentage KS3 Classes Welsh as Part Medium	97	17	86
Percentage KS4 Classes Welsh as Sole/Main Medium	12	30	3
Percentage KS4 Classes Welsh as Part Medium	59	31	93
% Full Time Teachers of Welsh as L1 / Welsh medium	59	46	19
Percentage of Pupils fluent in Welsh	66	62	32
Percentage of Pupils partly fluent in Welsh	34	38	68
Percentage of Pupils speaking Welsh at home	49	46	21
Percentage of Pupils studying Welsh as a L1	84	73	25
Percentage of Pupils studying Welsh as a L2	16	26	71
Percentage of Pupils studying any subject through the medium of Welsh	42	51	12

(4) The **six cluster solution** has some interesting changes. 'Welsh Medium' Schools are split into two groups as is highlighted below.

Variable	Welsh Medium 'B'	New Welsh Medium 'A'
Number of Schools in Cluster	24	12
Number of Subjects taught through Welsh / Bilingually	9	8
Percentage of Welsh speaking Governors	74	97
Percentage KS3 Classes Welsh as Sole/Main Medium	86	88
Percentage KS3 Classes Welsh as Part Medium	1	8
Percentage KS4 Classes Welsh as Sole/Main Medium	83	90

Percentage KS4 Classes Welsh as Part Medium	4	8
% Full Time Teachers of Welsh as L1 / Welsh medium	74	64
% Part Time Teachers of Welsh as L1 / Welsh medium	13	28
Percentage of Pupils fluent in Welsh	84	79
Percentage of Pupils partly fluent in Welsh	12	12
Percentage of Pupils speaking Welsh at home	41	72
Percentage of Pupils studying Welsh as a L1	93	98
Percentage of Pupils studying Welsh as a L2	7	2
Percentage of Pupils studying any subject through the medium of Welsh	58	82

This sixth cluster contains 12 Schools in the ‘Welsh Medium ‘A’ Cluster. They are:

1. Ysgol Dyffryn Ogwen Bethesda
2. Ysgol Botwnnog
3. Ysgol Eifionydd
4. Ysgol Y Moelwyn
5. Ysgol Syr Hugh Owen
6. Ysgol Y Berwyn
7. Ysgol Glan Y Mor
8. Ysgol Y Creuddyn
9. Ysgol Uwchradd Glan Clwyd
10. Ysgol Gyfun Penweddig
11. Ysgol Maes Y Gwendraeth
12. Ysgol Gyfun Gymraeg Glantaf

One difference from Welsh medium ‘B’ schools is that a significantly *higher* number of ‘A’ Schools have: (i) pupils who come from Welsh speaking homes and (ii) pupils who study any subject through the medium of Welsh. Also, the percentage of Welsh speaking Governors is higher in this Welsh medium ‘A’ cluster compared with the Welsh medium ‘B’ cluster (97% compared with 74%).

Should there be just one Welsh medium grouping as in Cluster Solution five, or should Welsh medium be subdivided as in Cluster Solution six?

(11) Categorizing the 34 ‘Bilingual Schools’ that change across Solutions

Most of the 194 schools do not change categories across the three to six clusters. The allocation of English medium schools is particularly stable across clusters.

However, there are some 34 schools who change classification across the solutions, mostly due to varying in their ‘bilingualism’, but also a few that are ‘more Welsh’ than some of the other ‘Welsh medium schools’. The following table lists those 34 schools. The tables are colour coded for ease of reading. Bilingual Schools are shaded WHITE. English Medium Schools are shaded BLUE. Welsh Medium Schools are shaded YELLOW. RED Shading on the six cluster solution denotes those 12

Schools in the 'Welsh Medium 'A' Cluster. Welsh Medium 'B' Schools have not changed from previous Welsh medium Clusters (and are not listed in the table below).

School.Name	THREE CLUSTERS	FOUR CLUSTERS	FIVE CLUSTERS	SIX CLUSTERS
Ysgol Uwchradd Caergybi				
Ysgol Uwchradd Bodedern				
Ysgol Dyffryn Ogwen Bethesda				
Ysgol Botwnnog				
YSGOL EIFIONYDD				
Ysgol Y Moelwyn				
Ysgol Y Berwyn				
Ysgol Friars				
Ysgol Syr Hugh Owen				
Ysgol Glan y Mor				
Ysgol John Bright				
Ysgol Aberconwy				
YSGOL Y CREUDDYN				
Ysgol Emrys ap Iwan				
Eirias High School				
Ysgol Uwchradd Glan Clwyd				
Ysgol Brynhyfryd				
Ysgol Uwchradd Caereinion High School				
Ysgol Maesydderwen				
Crickhowell High School				
Penglais School				
YSGOL GYFUN PENWEDDIG				
Ysgol Dewi Sant				
Ysgol Greenhill School				
Glan-y-Mor School				
Ysgol Maes Y Gwendraeth				
Ysgol Bro Dinefwr				
St John Lloyd Catholic Comprehensive School				
Cefn Saeson Comprehensive School				
Mountain Ash Comprehensive School				
Treorchy Comprehensive School				
Bedwas High School				
Ysgol Gyfun Gymraeg Glantaf				
Cardiff West Community High School				

The choice of how many clusters the team wish to choose is mostly about these 34 schools.

Which one of the 3 or 4 or 5 or 6 cluster solutions best allocates these 34 schools to how many clusters? How should those chosen clusters be labelled?

(12) Comparing the 2007 WAG Classification system with the three to six Cluster Solutions

The current PLASC data from schools includes a school's self-classification as a Welsh medium, English medium or Bilingual medium School using the October 2007 WAG paper: *'Defining schools according to Welsh medium provision'*.

How does the current PLASC classification compare with the clusters found in these analyses?

Appendix 6 provides the tables of comparison. However, we are not comparing like with like.

There is a proviso that the WAG 2007 Secondary School classification system concerns different 'text defined' dimensions of Welsh and English schools. In comparison, this analysis uses 15 PLASC dimensions. Therefore, we cannot expect 100% similarity.

The most apt comparison is probably the 2007 Classification used in PLASC 2017/18 (aggregated to four groups – see below) and for this report, compared with the Four Cluster Solution in the analyses presented in this paper. This basically compares the following four groups in each system: Welsh Medium; English Medium; 'Welsh Bilingual' and 'English Bilingual'. The result is **80.9% similarity**, or one on every five schools is differently allocated across the two systems.

The results (see **Appendix 6**) suggest that (i) The classification of English medium schools is similar between the two systems (ii) Welsh Medium and Bilingual Schools show a more varied allocation. For example (see Appendix 6 for the table), on the four cluster '2007' solution, only 44.8% of Welsh Medium schools were in that same Welsh medium group using the 15 PLASC variables, with the remaining Welsh medium schools being distributed across the English Bilingual (10.3%), English medium (6.9%) and Welsh Bilingual (37.9%) groups.

The current self-classification using the '2007' system tends to show Welsh medium schools and Bilingual schools as variable in distribution.

(13) What Variables are the Key Ones?

From the PLASC database, 15 variables were analysed. Which of these variables were more or less important in allocating Schools to Clusters? The statistical tool to examine this is called Discriminant Analysis. **Appendix 7** provides details of the analyses.

The results suggest that three variables are dissimilar from the other variables and have a different influence compared with the other variables: (i) the percentage of KS3 Classes where Welsh is a part medium for teaching (ii) the percentage of KS4 Classes where Welsh is a part medium for teaching and (iii) the percentage of teachers in a school who are qualified teach Welsh or through the medium of Welsh but are not doing so. However, conceptually, classes where Welsh is a part medium for teaching is important and helps distinguish some cluster from others.

(14) How might WAG use PLASC data to Group Schools?

If the team recommends a classification to WAG (e.g. the four or five cluster solution), it may need to keep in mind how WAG would operationalize this. Any classification system needs to be deliverable.

To commence the discussion, three early possibilities are briefly outlined below.

Possibility 1: Replicate the Cluster Analyses in this paper. Since much aggregation is involved and the analyses can be time-consuming, this may not be the preferred option. Also, groupings may not be the same every year (e.g. seven clusters may appear!).

Possibility 2: Use only a small number of variables that are powerful in creating separate groups (see Appendix 7 of the technical report for details of these). Then either (i) a cluster analysis could be used or (ii) a simple calculation be performed across these variables.

This may be worth considering, as then classification can avoid the problems of a self-classification by Schools themselves. Currently, analyses in this project suggest that schools are variable in self-categorization to Welsh Medium and Bilingual Schools. This is despite the October 2007 WAG paper: *'Defining schools according to Welsh medium provision'* being quite detailed about self-classification.

Possibility 3: Provide sufficient details (with or without an algorithm) for a school to classify itself. This may have the advantage of Schools becoming aware of 'progression' routes.

Appendix 1

The Specification of a Secondary School Data file from the 2017/2018 Pupil Level Annual School Census (PLASC)

Green highlighting = Welsh language dimensions (variables) for a Cluster Analysis. At the end of the paper, three further possible Derivable Variables are added.

Yellow Highlighting = Background data needed for post cluster analyses (e.g. names of Schools in each Cluster). but not be used for categorization.

KEY:

NS	Nursery School
PS	Primary School
SS	Secondary School
SP	Special School
MS	Middle School

School Identifiers

Field Name	Field Length	Field Type	Sample Data	Schools to complete				
LEA number	3	Alphanumeric	660	NS	PS	SS	SP	MS
Establishment	4	Alphanumeric	4099	NS	PS	SS	SP	MS
School name	100	Alphanumeric	Anglesey Comprehensive School	NS	PS	SS	SP	MS
School phase			MS	NS	PS	SS	SP	MS

School Characteristics

Field Name	Field Length	Field Type	Sample Data	Schools to complete				
School type	2	Alphanumeric	52	NS	PS	SS	SP	MS
Full time attendance census day	4	Alphanumeric	123	NS	PS	SS	SP	MS
Welsh medium school/subjects	2	Alphanumeric	5			SS		MS
Welsh medium school type	2	Alphanumeric	EW		PS	SS		MS

Governance

Field Name	Field Length	Field Type	Sample Data	Schools to complete			
Male serving governors	2	Alphanumeric	3	PS	SS	SP	MS
Female serving governors	2	Alphanumeric	3	PS	SS	SP	MS
Welsh speaking governors	2	Alphanumeric	3	PS	SS	SP	MS

A variable (at School level) was created of % of Governors who are Welsh speaking: i.e. (WELSH SPEAKING GOVERNORS ÷ TOTAL N GOVERNORS) X100.

CLASSES data

Field Name	Field Length	Field Type	Sample Data	Schools to complete				
Class name	30	Alphanumeric	OAK	NS	PS	SS	SP	MS
Class category	1	Alphanumeric	O		PS			MS
Class year group	2	Alphanumeric	5		PS	SS	SP	MS
Class key stage	1	Alphanumeric	2		PS	SS	SP	MS
Level of Welsh teaching of the class	1	Alphanumeric	1		PS	SS	SP	MS

In **Level of Welsh teaching of the Class** the codes are: 1=Welsh sole/main medium; 2=Welsh as medium for less than half curriculum; 3=second language only; 4=no Welsh used or taught. From this, Classes were aggregated to School level separating KS3 and KS4.

% KS3 Classes in a School with code=1 (i.e. Welsh as sole or main medium).

% KS3 Classes in a School with code=2 (i.e. Welsh as part medium).

% KS4 Classes in a School with code=1 (i.e. Welsh as sole or main medium).

% KS4 Classes in a School with code=2 (i.e. Welsh as part medium).

TEACHERS

Field Name	Field Length	Field Type	Sample Data	Schools to complete				
Headcount	2	Alphanumeric	50	NS	PS	SS	SP	MS

Teaching of Welsh

Field Name	Field Length	Field Type	Sample Data	Schools to complete			
Teaching Welsh category	2	Alphanumeric	TC		PS	SS	MS
Tenure	1	Alphanumeric	F		PS	SS	MS
Headcount	2	Alphanumeric	50		PS	SS	MS

For the Teaching of Welsh variable, the codes are:

TC Qualified teachers teaching Welsh as a first language

TW Qualified teachers teaching Welsh as a second language only

TO Qualified teachers teaching other subjects through the medium of Welsh

NW Qualified teachers able to teach Welsh or through the medium of Welsh, but not doing so

NT Not qualified to teach Welsh or through the medium of Welsh.

From this raw data, three variables were constructed at School level.

TEACHER_FT_TCTO

% Full Time Teachers in a School who are:

TC Qualified teachers teaching Welsh as a first language **plus**

TO Qualified teachers teaching other subjects through the medium of Welsh

TEACHER_PT_TCTO

% Part Time Teachers in a School who are:

TC Qualified teachers teaching Welsh as a first language **plus**

TO Qualified teachers teaching other subjects through the medium of Welsh

TEACHER_FT_NW

% Full Time Teachers in a School who are:

NW Qualified teachers able to teach Welsh or through the medium of Welsh, **but not doing so.**

NB: The Part-Time Teacher equivalent of this variable had very small numbers indeed, so has **not** been included although the initial analyses encompassed this.

PUPIL LEVEL DATA

Pupil-level data has a high level of confidentiality (as individual pupils are identifiable), and so could not be transferred. School level variables derived from pupil data were aggregated to School level on transfer, with a breakdown into KS3 and KS4 not available.

Welsh Language

Field Name	Field Length	Field Type	Sample Data	Schools to complete			
Fluency in Welsh	1	Alphanumeric	1	PS	SS		MS
Speaking Welsh in the home	1	Alphanumeric	1	PS	SS		MS

Study of Welsh	1	Alphanumeric	1		PS	SS		MS
Welsh medium education	5	True/False	0			SS		MS

For the **Fluency in Welsh** variable, the codes are: 1=fluent in Welsh; 2=speaks Welsh but not fluent; 3=cannot speak Welsh; 4= information refused. For percentage calculations, pupils with 'information refused' were excluded. The two variables that were created are:

% Pupils in a School 'fluent in Welsh'

% Pupils in a School 'speaks Welsh but not fluent'

For the **Speaking Welsh in the Home** variable: Codes of 0 = does not speak Welsh at home; 1= speaks Welsh at home; 2=not applicable. For percentage calculations, pupils with 'not applicable' were excluded. The variable created is:

% Pupils who speak Welsh at home

For the **Study of Welsh variable**: Codes of 1= taught Welsh as a L1; 2= taught Welsh as a L2; 5=disapplied. Hence L2 Welsh was included as a separate variable. The percentage calculation *includes* 'disapplied from the National Curriculum' on the basis that all pupils create the language character of the school. The variables created are:

% Pupils in a School taught Welsh as a L1

% Pupils in a School taught Welsh as a L2

For the **Welsh Medium Education variable** is 'Pupils who take any subject (other than L1/L2 Welsh) through the *medium* of Welsh'. The codes are: 1=true ;0=false. A breakdown into Year 7,8,9,10,11 was requested but was not available. The variable created is:

% Pupils in a School who take any subject (other than L1/L2 Welsh) through the medium of Welsh.

APPENDIX 2: CLUSTER ANALYSIS

1. Choosing Dimensions

Cluster analysis begins with a specification of the variables to be entered into the cluster analysis. Since this specification has an effect on the results, agreeing a specification is important. This is a matter for the team. It is not a statistical decision but a conceptual one regarding how to profile Welsh medium education. This is an analogous to deciding what ingredients to enter into a cake, with the final flavour being affected by that decision.

However, it is wise to enter slightly different combinations of ingredients into successive cluster analyses. That is, by varying the dimensions entered into different analyses, it is possible to see how those dimensions perform and how robust are the final preferred results.

On this Project, it was found early on that the Teacher variable 'Part-time Qualified Teachers able to teach Welsh or through the medium of Welsh had such small numbers (n=23) that it was important to exclude. Other runs were made excluding 'Welsh speaking Governors', but on some analyses this variable was influential and so was included in all final analyses. Appendix 7 contains a further analysis of the strength of importance in the variables included. A conclusion was that 14 of the 15 PLASC variables were important to include on all analyses.

2. Cluster Analysis as a Statistical procedure

Cluster analysis is not one simple procedure. It is an umbrella term under which there are a variety of methods. These methods do not just differ in their formula and calculations (algorithms) but also in their assumptions and styles. Some methods try to locate clusters with very small spatial distances between their members and large distances between the clusters themselves. Other methods trying to work hierarchically, combining members with the shortest distances between them and then working iteratively until there is just one cluster. Sometimes, statistical distributions are employed as are graph-based models.

So long as the statistician / researcher avoids fishing for a preferred solution, then a sensible strategy tends to be to use two or three different approaches to see if they suggest the same or a very similar solution. This is the approach taken with this project whereby hierarchical clustering (connectivity modelling) was used as well as a centroid model (the K-Means algorithm) to test the robustness and reproducibility of solutions.

Experience suggests that the K-Means algorithm tends to produce interpretable and sensible results when the database is particularly large (as with our Primary school data) and concerns 'human' data. Hierarchical clustering tends to produce similar outcomes with a dendrogram helping a decision about the final number of clusters. Hierarchical clustering is possibly more intuitively understandable.

Reservations about the above approach include: (i) it assumes hard clustering whereby each School will belong to a cluster or not. An alternative conceptualisation is soft clustering (or fuzzy clustering) whereby each School would have a probability of belonging to each cluster. While statisticians sometimes regard this as more acceptable, in terms of this project, the classification of schools needs to be exact (e.g. one School cannot be located in two different clusters). (ii) Neither hierarchical clustering nor the K-means algorithm tend to work with outliers. For example, a school which is completely unlike any other school would be considered an outlier. In some individual pupil datasets, for example, this can happen. However, given that our classification of schools needs to be into groups, and taking the assumption that there are no real outliers among our schools in Wales, this is not regarded as a problem.

(iii) Another limitation is that the number of clusters in the data is not always agreed. Even the dendrogram does not produce a definitive expression of the 'right' number of clusters in the data. While various statistical significance tests were run with this data, experience suggests that *interpretability* of the final clusters is important. For example, it seems valuable to carefully inspect several different final solutions and see which one is the most interpretable among the experts involved in the team. My experience is to try one solution beyond the dendrogram and scree plot (in this case the six cluster solution). Occasionally, this is the most interpretable. Often not. So with this data, for example, solutions for three, four, five and six clusters have been reproduced for the team for their final decision as to the most sensible final result.

3. Cluster Analysis Decisions

Decisions are made when conducting Cluster Analysis both for hierarchical clustering and for K-Means analysis. For this project, these were as follows:

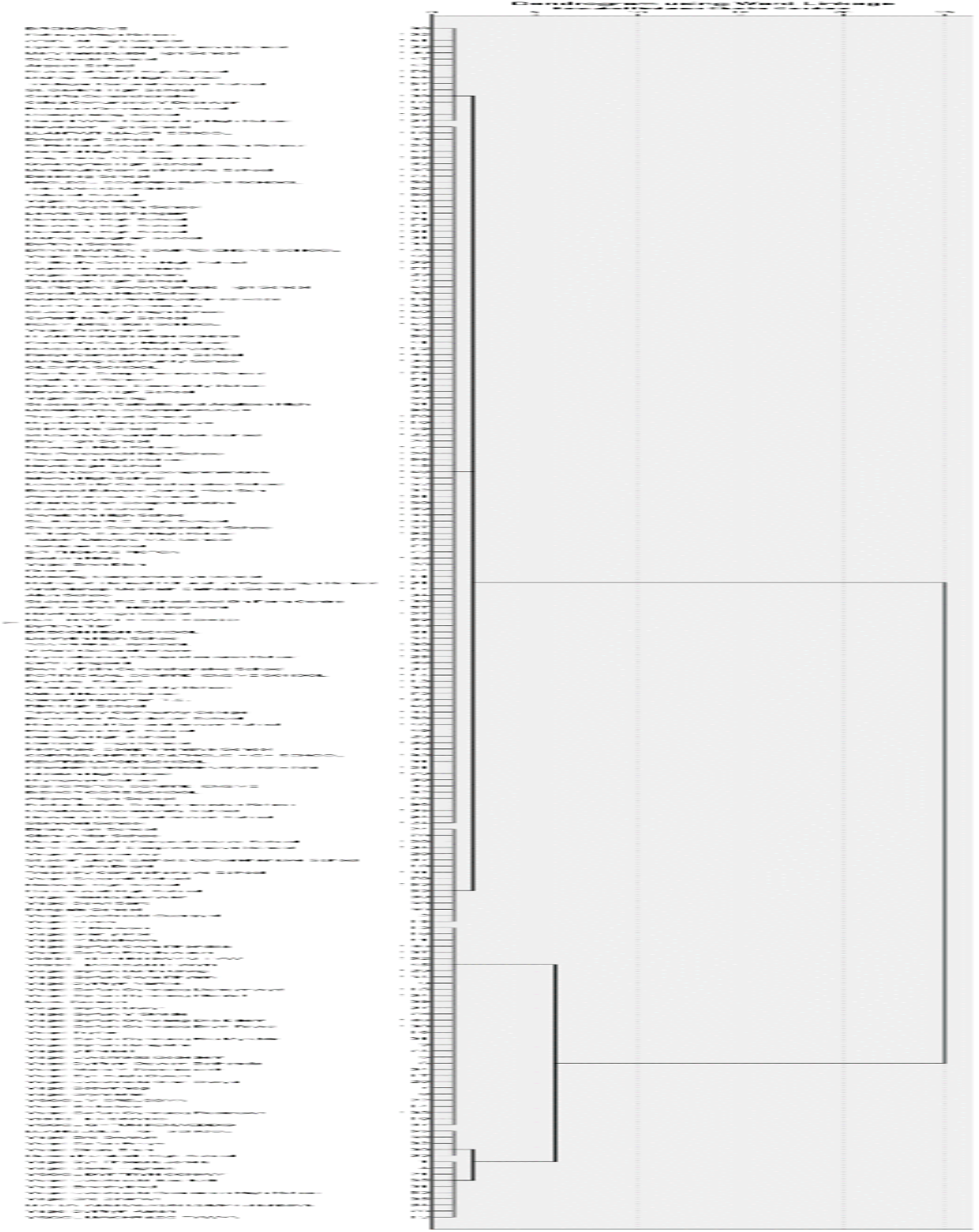
- (i) Proximity (distance / similarity) Coefficient = Squared Euclidean Distance.
- (ii) Distance between Clusters used Wards' Method.
- (iii) Outlier analysis performed as a check. This is important before a K-Means analysis.
- (iv) Standardization of variables (as those variables with a larger scale tend to add disproportionate importance to an analysis) was via Z scores. Variables with large values tend to contribute more to the distance measure than variables with small values, making variable standardization important. All 15 variables were converted to z scores. There was no weighting of variables, although this is possible in SPSS.
- (v) Use of a Dendrogram to inform the non-hierarchical K-means analyses.
- (vi) K-Means analyses were varied from three to six clusters

4. Cluster Analysis program

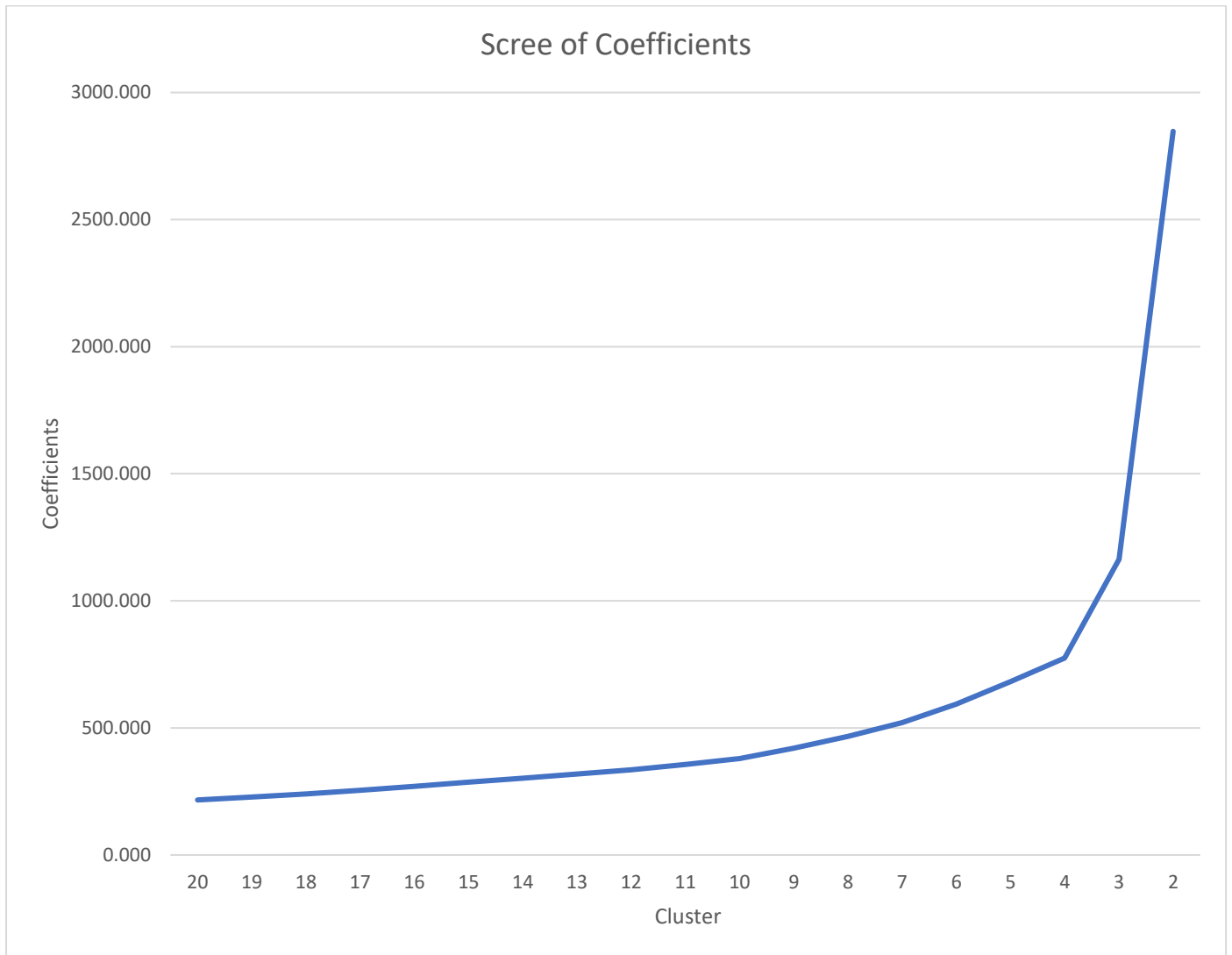
- (i) The cluster analyses were conducted in **SPSS**. SPSS has sub programs for K-Means analyses and Hierarchical cluster analysis. **SPSS** (also called SPSS IBM) reached Version 25 in 2017 have commenced in 1968. SPSS is a widely used program for a considerable variety of statistical analysis. It is used by for research in education, social science, market research, health research, surveys, government etc..

Reference: IBM Corporation, 2017, IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corporation.

APPENDIX 3 DENDROGRAM



APPENDIX 4 SCREE OF COEFFICIENTS



Appendix 5

Scores on the 15 PLASC Variables for each School

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PUP PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Ysgol Syr Thomas Jones	0	100	0	100	8	92	82	15	0	77	23	56	96	4	2
Ysgol Uwchradd Caergybi	1	43	12	0	20	0	0	0	26	12	88	16	15	84	2
Ysgol Gyfun Llangefni	10	73	100	0	100	0	93	5	0	82	18	72	96	4	83
Ysgol David Hughes	0	95	0	88	18	53	65	18	0	69	31	46	78	21	69
Ysgol Uwchradd Bodedern	10	100	81	6	40	50	83	13	3	72	28	68	85	15	83
Ysgol Dyffryn Ogwen Bethesda	10	100	82	0	100	0	62	28	0	90	10	78	99	1	56
Ysgol Botwnnog	10	100	92	8	100	0	68	29	0	98	2	78	99	1	99
Ysgol Brynrefail	10	100	88	0	100	0	84	14	0	93	7	89	99	1	36
Ysgol Dyffryn Nantlle	10	100	100	0	100	0	78	7	15	93	7	89	99	1	24
YSGOL EIFIONYDD	10	100	100	0	100	0	62	31	8	80	20	70	97	3	96
Ysgol Y Moelwyn	0	100	100	0	100	0	50	41	5	83	17	86	96	4	94
YSGOL UWCHRADD TYWYN	0	73	0	50	43	14	24	14	10	35	65	18	95	5	92
Ysgol Y Berwyn	0	100	83	17	63	38	57	35	0	85	15	75	92	8	66
Ysgol Ardudwy	10	89	67	0	100	0	68	16	8	89	11	34	89	11	86
Ysgol Friars	6	29	0	40	24	24	17	6	15	18	80	6	20	63	5
Ysgol Tryfan	10	100	70	0	100	0	82	6	6	81	19	67	100	0	91
Ysgol Syr Hugh Owen	10	100	88	0	77	0	70	20	0	93	6	87	99	0	88
Ysgol Glan y Mor	0	100	86	14	50	50	61	36	0	83	17	80	96	4	91
Ysgol John Bright	0	25	0	0	0	0	0	0	10	4	95	25	0	100	0
Ysgol Aberconwy	0	21	0	0	0	0	2	5	7	8	88	15	6	93	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
YSGOL DYFFRYN CONWY	10	90	0	100	0	58	60	28	9	67	33	64	86	13	84
YSGOL Y CREUDD YN	10	100	100	0	100	0	66	32	0	62	38	70	100	0	34
Ysgol Emrys ap Iwan	0	6	0	0	0	0	0	0	0	3	97	6	0	100	0
Eirias High School	0	0	0	0	0	0	3	0	10	11	71	16	6	94	0
Ysgol Bryn Elian	0	17	0	0	0	0	0	0	0	4	96	9	0	100	0
Rhyl High School	0	0	0	0	0	0	0	0	0	2	96	2	0	100	0
Prestatyn High School	0	0	0	0	0	0	0	0	0	3	93	10	0	100	0
Ysgol Uwchradd Glan Clwyd	10	100	100	0	100	0	81	19	0	97	3	62	98	2	95
Denbigh High School	0	0	0	0	0	0	0	0	0	5	78	14	0	100	0
Ysgol Dinas Bran	0	6	9	91	13	67	19	10	6	27	73	14	11	75	13
Ysgol Brynhyfryd	10	63	36	0	38	0	52	10	6	43	56	40	37	63	24
Blessed Edward Jones High Sch	0	0	0	0	0	0	0	0	0	1	99	2	0	100	0
Hawarden High School	1	0	0	0	0	0	0	0	0	1	99	3	0	99	0
Alun School	0	20	0	0	0	0	0	0	0	2	98	10	0	77	0
Elfed High School	0	5	0	0	0	0	0	0	6	1	99	7	0	98	0
Ysgol Treffynnon	0	0	0	0	0	0	0	0	0	3	86	6	0	98	0
St. David's High School	0	6	0	0	0	0	0	0	0	0	0	23	0	100	0
Castell Alun High School	0	0	0	0	0	0	0	0	0	4	93	9	0	90	0
Maes Garmon	10	56	100	0	100	0	67	31	0	98	2	27	100	0	77
Flint High School	0	6	0	0	0	0	0	0	0	3	66	2	0	85	0
Connah's Quay High School,	0	0	0	0	0	0	0	0	0	1	85	1	0	100	0
Argoed School	0	13	0	0	0	0	0	0	0	0	0	3	0	94	0
St.Richard Gwyn HS	0	0	0	0	0	0	0	0	0	1	96	9	0	99	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Grango	1	31	0	0	0	0	0	0	0	3	93	2	0	99	0
YSGOL MORGAN LLWYD	10	60	100	0	57	0	69	16	9	89	11	28	97	3	32
Ysgol Bryn Alyn	0	0	0	0	0	0	0	0	0	5	94	4	0	99	0
Darland High School	0	5	0	0	0	0	0	0	4	3	89	2	0	93	0
Ysgol Rhiwabon	0	6	0	0	0	0	0	0	0	2	92	0	0	92	0
Rhosnesni High School	0	13	0	0	0	0	0	0	2	1	50	1	0	100	0
Ysgol Clywedog	1	5	0	0	0	0	0	0	0	1	99	2	0	99	0
St Joseph's Catholic and Anglican High	1	0	0	0	0	0	0	0	0	1	89	2	0	99	0
THE MAELOR SCHOOL	0	11	0	0	0	0	0	0	0	0	100	1	0	86	0
Ysgol Uwchradd Caereinion High School	9	100	42	0	9	27	40	18	3	58	42	29	58	42	54
Llanfyllin High School	9	27	20	7	14	7	30	9	0	23	76	21	18	64	21
LLANIDLO ES HIGH SCHOOL	10	25	0	100	0	100	28	5	10	16	84	15	16	84	1
Newtown High School	0	0	0	0	0	0	1	0	6	2	92	3	2	98	0
WELSH POOL HIGH SCHOOL	1	16	0	0	0	0	4	0	2	4	95	5	2	98	0
LLANDRI NDOD HIGH SCHOOL	0	0	0	0	0	0	0	0	0	1	90	4	0	100	0
BUILTH WELLS HIGH SCHOOL	6	20	14	0	0	0	16	2	0	40	60	18	31	68	31
Ysgol Maesydde rwen	0	56	8	0	0	0	3	0	11	14	86	15	1	98	0
BRECON HIGH SCHOOL	5	20	9	0	11	0	13	0	3	12	88	5	4	83	1
Gwernyfed High School	1	0	0	0	0	0	0	0	3	4	91	10	0	100	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Crickhowe II High School	0	5	0	0	0	0	0	0	14	3	69	7	0	89	0
GYFUN ABERAER ON COMPRE	0	83	31	23	0	44	24	17	0	61	39	36	79	18	17
Ysgol Uwchradd Aberteifi	9	47	0	100	22	33	30	12	0	52	48	30	73	26	13
Penglais School	3	44	0	4	7	7	7	4	14	24	73	17	18	62	6
YSGOL GYFUN PENWED DIG	10	81	50	50	89	11	56	21	5	84	16	67	100	0	99
Ysgol Bro Gwaun	6	32	15	85	0	100	16	0	8	31	69	22	24	73	20
Ysgol Dewi Sant	0	35	0	0	0	0	0	0	14	8	89	10	0	100	0
Ysgol Greenhill School	0	11	0	0	0	0	0	0	10	2	96	2	0	100	0
Pembroke School	0	6	0	0	0	0	0	0	1	2	97	4	0	90	0
SIR THOMAS PICTON	0	16	0	0	0	0	0	0	0	7	93	4	0	99	0
Milford Haven School	0	0	0	0	0	0	0	0	0	0	68	1	0	100	0
Ysgol y Preseli	10	100	100	0	100	0	64	0	0	95	5	40	100	0	100
Tasker- Milward V.C. School	0	17	0	0	0	0	0	0	0	2	98	3	0	100	0
Ysgol Dyffryn Aman	0	95	24	27	24	24	32	3	0	54	46	35	53	46	12
Coedcae School	0	18	0	0	0	0	0	0	0	0	100	0	0	100	0
Ysgol Gyfun Y Strade	10	75	92	0	83	0	83	16	0	99	1	41	100	0	81
Glan-y- Mor School	0	0	0	0	0	0	0	0	10	17	83	9	0	100	0
Bryngwyn School	0	0	0	0	0	0	0	0	0	20	79	14	0	100	0
Ysgol Gyfun Gymraeg Bro Myrddin	10	100	82	0	100	0	80	7	7	100	0	76	100	0	86
Ysgol Gyfun Emlyn	9	28	25	67	0	100	20	0	7	49	50	21	62	38	22

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Queen Elizabeth High School	5	25	4	89	0	100	11	0	27	36	64	34	13	86	1
Ysgol Maes Y Gwendraeth	10	94	72	11	100	0	76	20	0	96	4	64	100	0	65
Ysgol Bro Dinefwr	10	50	35	13	4	30	42	4	11	67	33	36	61	37	21
Dyffryn Taf	5	20	6	0	0	0	8	2	0	34	66	18	19	81	11
St John Lloyd Catholic Comprehensive School	0	27	0	0	0	0	0	0	6	15	84	12	0	100	0
Cefn Hengoed	5	10	0	0	0	0	0	0	0	0	0	43	0	99	0
OLCHFA SCHOOL	0	0	0	0	0	0	0	0	1	3	97	2	0	99	0
MORRISTON COMPREHENSIVE	0	5	0	0	0	0	0	0	0	1	99	1	0	100	0
PENTREHAFOD SCHOOL	0	0	0	0	0	0	0	0	0	7	93	19	0	97	0
BISHOP GORE SCHOOL	0	0	0	0	0	0	0	0	0	19	74	11	0	100	0
Penyrheol Comprehensive School	0	0	0	0	0	0	0	0	0	5	94	13	0	100	0
Gowerton Comprehensive School	1	10	0	0	0	0	0	0	0	17	75	25	0	99	0
BISHOPSTON COMPREHENSIVE	0	0	0	0	0	0	0	0	0	21	79	15	0	99	0
Pontarddulais Comprehensive School	0	30	0	0	0	0	0	0	0	15	85	21	0	99	0
Ysgol Gyfun Gwyr	10	80	100	0	100	0	73	24	0	93	7	30	100	0	65
BIRCHGROVE	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
Dylan Thomas Community School	0	7	0	0	0	0	0	0	2	1	99	0	0	100	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Ysgol Gyfun Gymraeg Bryn Tawe	10	95	100	0	100	0	75	22	0	86	14	11	100	0	100
Bishop Vaughan School	0	0	0	0	0	0	0	0	0	0	100	0	0	89	0
Cymer Afan Comprehe nsive School	0	6	0	0	0	0	0	0	0	0	0	0	0	100	0
Dyffryn School	0	0	0	0	0	0	0	0	0	3	95	4	0	100	0
Cefn Saeson Comprehe nsive School	0	0	0	0	0	0	0	0	8	3	53	16	0	100	0
Cwmtawe Communit y School	0	21	0	0	0	0	0	0	0	16	83	26	0	97	0
Llangatwg Communit y School	0	0	0	0	0	0	0	0	0	9	84	11	0	100	0
Dwr Y Felin Comprehe nsive School	0	0	0	0	0	0	0	0	0	7	50	14	0	100	0
St Joseph's RC School and 6th Form Centre	0	5	0	0	0	0	0	0	0	5	95	11	0	78	0
Cynffig Comprehe nsive	0	11	0	0	0	0	0	0	0	2	6	16	0	100	0
Bryntirion Comprehe nsive.	0	6	0	0	0	0	0	0	0	1	98	2	0	100	0
Maesteg Comprehe nsive School	0	10	0	0	0	0	0	0	0	6	92	4	0	85	0
PENCOE D COMPRE HENSIVE	0	5	0	0	0	0	0	0	0	5	85	7	0	100	0
Brynteg School	0	0	0	0	0	0	0	0	0	3	72	7	0	100	0
PORTHC AWL COMPRE HENSIVE SCHOOL	0	5	0	0	0	0	0	0	0	11	39	14	0	80	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Ysgol Gyfun Gymraeg Llangynwy d	10	60	100	0	100	0	81	11	0	0	0	26	100	0	99
Coleg Cymunedo IY Dderwen	1	13	0	0	0	0	0	0	4	1	26	7	0	100	0
Archbisho p McGrath Catholic School	1	10	0	0	0	0	0	0	0	5	93	6	0	82	0
LLANTW I MAJOR SCHOOL	0	0	0	0	0	0	0	0	5	2	98	3	0	100	0
BARRY COMPRE HENSIVE SCHOOL	0	5	0	0	0	0	0	0	0	8	91	0	0	100	0
BRYN HAFREN COMPRE HENSIVE SCHOOL	0	0	0	0	0	0	0	0	0	4	96	5	0	100	0
COWBRID GE COMPRE HENSIVE SCHOOL	1	5	0	0	0	0	0	0	1	6	92	20	0	100	0
St Cyres Comprehe nsive School	0	10	0	0	0	0	0	0	0	2	98	1	0	100	0
St Richard Gwyn Catholic High School	1	6	0	0	0	0	0	0	7	0	100	2	0	100	0
Stanwell School	0	0	0	0	0	0	0	0	0	22	78	49	0	100	0
Bryncelyn nog Comprehe nsive School	0	16	0	0	0	0	0	0	3	39	55	27	0	100	0
The Pontypridd High School	0	0	0	0	0	0	0	0	0	2	98	2	0	100	0
Hawthorn High School	0	18	0	0	0	0	0	0	2	5	88	8	0	83	0
Mountain Ash Comprehe nsive School	0	16	0	0	0	0	0	0	10	21	68	24	0	100	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Ysgol Gyfun Garth Olwg	10	75	100	0	100	0	69	18	11	90	10	36	100	0	8
TONYREF AIL SCHOOL	0	5	0	0	0	0	2	0	2	44	50	36	0	95	0
Treorchy Comprehensive School	4	10	0	18	0	0	2	5	8	15	85	14	2	77	13
Ferdale Community School	0	13	0	0	0	0	0	0	3	3	22	12	0	96	0
Porth County Community	0	7	0	0	0	0	0	0	0	6	91	0	0	100	0
Tonypany Community College	0	0	0	0	0	0	0	0	0	1	54	3	0	100	0
Y Pant Comprehensive	0	0	0	0	0	0	0	0	0	52	45	32	0	100	0
Ysgol Gyfun Cwm Rhondda	4	0	100	0	100	0	85	13	0	95	5	41	100	0	0
Ysgol Gyfun Rhydywau n	4	36	100	0	100	0	92	8	0	95	5	23	100	0	18
Aberdare Community School	0	5	0	0	0	0	0	0	0	1	78	3	0	100	0
Cardinal Newman R.C.	0	0	0	0	0	0	0	0	2	4	76	5	0	100	0
St.John Baptist High School	0	5	0	0	0	0	0	0	0	6	94	3	0	100	0
Afon Taf High School	1	6	0	0	0	0	0	0	0	0	0	0	0	100	0
PEN-Y-DRE HIGH SCHOOL	0	0	0	0	0	0	0	0	0	10	90	0	0	100	0
Cyfarthfa High School	0	0	0	0	0	0	0	0	0	6	93	2	0	100	0
Bishop Hedley High School	0	7	0	0	0	0	0	0	3	0	0	1	0	100	0
Newbridge School	0	0	0	0	0	0	0	0	0	0	100	0	0	100	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Blackwood Comprehe nsive School	0	0	0	0	0	0	0	0	0	15	55	0	0	100	0
St Cenydd School	0	10	0	0	0	0	0	0	0	0	0	0	0	98	0
Risca Communit y Comprehe nsive	0	0	0	0	0	0	0	0	0	0	100	0	0	100	0
St Martin's School	0	7	0	0	0	0	0	0	0	2	96	3	0	100	0
HEOLDD U COMPRE HENSIVE SCHOOL	0	0	0	0	0	0	0	0	0	2	96	1	0	82	0
Lewis School Pengam	0	0	0	0	0	0	0	0	0	0	91	0	0	91	0
Lewis Girls' Comprehe nsive School	0	0	0	0	0	0	0	0	0	1	99	0	0	100	0
Bedwas High School	0	9	0	0	0	0	0	0	13	0	99	1	0	83	0
Ysgol Gyfun Cwm Rhywni	10	44	97	0	89	0	79	16	2	100	0	14	100	0	15
Islwyn High School	0	0	0	0	0	0	0	0	0	0	100	0	0	99	0
Cwmcarn High School	0	0	0	0	0	0	0	0	0	3	97	0	0	100	0
Tredegar Comprehe nsive School	0	5	0	0	0	0	0	0	3	0	0	0	0	100	0
Brynmawr Foundatio n School	0	0	0	0	0	0	0	0	0	1	57	1	0	100	0
Croesyceil iog School	0	0	0	0	0	0	0	0	7	0	0	1	0	100	0
Abersycha n Comprehe nsive	0	0	0	0	0	0	0	0	0	0	98	1	0	100	0
West Monmouth School	0	0	0	0	0	0	0	0	0	0	100	2	0	100	0
YSGOL GYFUN GWYNLL YW	0	72	89	0	83	0	83	16	1	99	1	6	100	0	51

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Cwmbran High School	0	0	0	0	0	0	0	0	0	1	99	1	0	99	0
St. Albans R.C. High School	0	0	0	0	0	0	0	0	0	0	100	3	0	100	0
Monmouth Comprehensive School	0	0	0	0	0	0	0	0	0	0	100	2	0	79	0
King Henry VIII Comprehensive	0	5	0	0	0	0	0	0	4	0	99	1	0	100	0
Chepstow Comprehensive School	1	18	0	0	0	0	0	0	0	1	99	3	0	100	0
Caldicot School	0	5	0	0	0	0	0	0	0	1	99	1	0	87	0
St Julian's School	0	0	0	0	0	0	0	0	0	0	100	1	0	100	0
The John Frost School	0	6	0	0	0	0	0	0	0	1	99	0	0	100	0
Llanwern High School	0	0	0	0	0	0	0	0	0	1	91	1	0	93	0
Newport High School	0	0	0	0	0	0	0	0	0	2	95	2	0	100	0
Liswerry High School	0	0	0	0	0	0	0	0	0	1	98	2	0	94	0
Bassaleg School	0	0	0	0	0	0	0	0	0	0	99	1	0	79	0
Caerleon Comprehensive School	0	0	0	0	0	0	0	0	1	1	88	4	0	100	0
St Joseph's RC High School	0	0	0	0	0	0	8	0	0	0	0	0	0	84	0
CARDIFF HIGH SCHOOL	0	0	0	0	0	0	5	0	0	4	96	3	0	100	0
Willows High School	0	0	0	0	0	0	16	0	0	25	75	0	0	100	0
Fitzalan High School	0	0	0	0	0	0	0	0	2	9	91	14	0	100	0
Cantonian High School	0	0	0	0	0	0	0	0	0	9	84	18	0	100	0
Llanishen High School	0	0	0	0	0	0	0	0	0	1	99	2	0	97	0

SCHOOL	Wels hMe dium Subj ects	Wel sh Go ver nor s	CL AS SK S3 MA IN	CL AS SK S3 PA RT	CL AS SK S4 MA IN	CL AS SK S4 PA RT	TEA CHE R_F T_T CTO	TEA CHE R_P T_T CTO	TEA CH ER_ FT_ NW	PU PIL _F LU EN T	PUP IL_N OTF LUE NT	WE LS HA TH OM E	STU DYO FW ELS HL1	STU DYO FW ELS HL2	WE LS HM EDI UM
Cathays High School	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
Radyr Comprehensive School	0	0	0	0	0	0	0	0	0	5	80	5	0	99	0
Ysgol Gyfun Gymraeg Glantaf	10	83	100	0	100	0	65	22	7	0	0	45	100	0	97
Ysgol Gyfun Gymraeg Plasmawr	8	71	95	0	100	0	69	19	5	98	2	51	100	0	69
Ysgol Gyfun Gymraeg Bro Edern	10	79	100	0	100	0	83	10	0	97	3	20	100	0	94
Cardiff West Community High School	0	11	0	0	0	0	0	0	8	0	0	0	0	99	0
Eastern High	0	13	0	0	0	0	0	0	0	1	91	1	0	100	0
St. Illtyd's Catholic High School	0	0	0	0	0	0	0	0	0	2	98	5	0	100	0
Mary Immaculate High School	0	6	0	0	0	0	0	0	0	0	0	0	0	100	0
Bishop of Llandaff Church in Wales High School	0	15	0	0	0	0	0	0	0	5	95	6	0	84	0
St Teilo's C-in-W High School	1	18	0	0	0	0	0	0	0	4	96	3	0	100	0
CORPUS CHRISTI CATHOLIC HIGH SCHOOL	0	0	0	0	0	0	0	0	0	3	96	16	0	100	0
Whitchurch High School	0	10	0	0	0	0	0	0	0	2	95	3	0	91	0

APPENDIX 6

Comparing the 2007 WAG Classification system with the Three to Six Cluster Solutions

The coding for the 2007 WAG classification system is:

WM – Welsh Medium

EM – English Medium

EW – English (with significant Welsh)

BILINGUAL (with coding AB, BB, CB, CH).

AB – A Bilingual:

At least 80% of subjects apart from English and Welsh are taught only through the medium of Welsh to all pupils. One or two subjects are taught to some pupils in English or in both languages.

BB – B Bilingual

At least 80% of subjects (excluding Welsh and English) are taught through the medium of Welsh but are also taught through the medium of English.

CB – C Bilingual

50 - 79% of subjects (excluding Welsh and English) are taught through the medium of Welsh but are also taught through the medium of English.

CH – Ch Bilingual

All subjects, except Welsh and English taught to all pupils using both languages.

For ease of understanding, the 2007 classification was recoded into four categories:

Welsh Medium Schools labelled **WM**

English Medium labelled **EM**

English (with significant Welsh) labelled **EW**

Bilingual Schools (combination of AB, BB, CB, CH) labelled **BL**.

The 15 PLASC dimensions used for comparison are listed in the main report. Below are the SPSS tables providing the analyses, with the results summarized in the main text of this report. The codes in the tables are as below:

Welsh Medium Schools labelled **WM**

English Medium labelled **EM**

English Bilingual labelled **EB**

Welsh Bilingual Schools labelled **WB**

Bilingual Schools **BL**.

2007 CLASSIFICATION compared with a THREE CLUSTER SOLUTION

			THIS PROJECT CLUSTERS			
			BL	EM	WM	Total
2007 CLASSIFICATION	WM	Count	9	2	18	29
		% within OLD CLASSIFICATION	31.0%	6.9%	62.1%	100.0%
	EM	Count	0	140	0	140
		% within OLD CLASSIFICATION	0.0%	100.0%	0.0%	100.0%
	EW	Count	5	3	0	8
		% within OLD CLASSIFICATION	62.5%	37.5%	0.0%	100.0%
	BL	Count	0	0	17	17
		% within OLD CLASSIFICATION	0.0%	0.0%	100.0%	100.0%
	Total	Count	14	145	35	194
		% within OLD CLASSIFICATION	7.2%	74.7%	18.0%	100.0%

2007 CLASSIFICATION compared with a FOUR CLUSTER SOLUTION

			THIS PROJECT CLUSTERS				
			EB	EM	WM	WB	Total
2007 CLASSIFICATION	WM	Count	3	2	13	11	29
		% within OLD CLASSIFICATION	10.3%	6.9%	44.8%	37.9%	100.0%
	EM	Count	0	140	0	0	140
		% within OLD CLASSIFICATION	0.0%	100.0%	0.0%	0.0%	100.0%
	EW	Count	4	4	0	0	8
		% within OLD CLASSIFICATION	50.0%	50.0%	0.0%	0.0%	100.0%
	BL	Count	0	0	17	0	17
		% within OLD CLASSIFICATION	0.0%	0.0%	100.0%	0.0%	100.0%
	Total	Count	7	146	30	11	194
		% within OLD CLASSIFICATION	3.6%	75.3%	15.5%	5.7%	100.0%

2007 CLASSIFICATION compared with a FIVE CLUSTER SOLUTION

			THIS PROJECT CLUSTERS					
			BILING_ KS3>KS4	EM	WB	WM	EB	Total
2007 CLASSIFICATION	WM	Count	4	2	9	13	1	29
		% within OLD CLASSIFICATION	13.8%	6.9%	31.0%	44.8%	3.4%	100.0%
	EM	Count	0	140	0	0	0	140
		% within OLD CLASSIFICATION	0.0%	100.0%	0.0%	0.0%	0.0%	100.0%
	EW	Count	0	4	0	0	4	8
		% within OLD CLASSIFICATION	0.0%	50.0%	0.0%	0.0%	50.0%	100.0%
	BL	Count	0	0	0	17	0	17
		% within OLD CLASSIFICATION	0.0%	0.0%	0.0%	100.0%	0.0%	100.0%
Total	Count	4	146	9	30	5	194	
	% within OLD CLASSIFICATION	2.1%	75.3%	4.6%	15.5%	2.6%	100.0%	

2007 CLASSIFICATION compared with a SIX CLUSTER SOLUTION

			THIS PROJECT CLUSTERS						Total
			EM	EB	WM2	WB	BL	WM1	
2007 CLASSIFICATION	WM	Count	2	0	10	1	7	9	29
		% within OLD CLASSIFICATION	6.9%	0.0%	34.5%	3.4%	24.1%	31.0%	100.0%
	EM	Count	126	14	0	0	0	0	140
		% within OLD CLASSIFICATION	90.0%	10.0%	0.0%	0.0%	0.0%	0.0%	100.0%
	EW	Count	1	3	0	4	0	0	8
		% within OLD CLASSIFICATION	12.5%	37.5%	0.0%	50.0%	0.0%	0.0%	100.0%
	BL	Count	0	0	14	0	0	3	17
		% within OLD CLASSIFICATION	0.0%	0.0%	82.4%	0.0%	0.0%	17.6%	100.0%
Total	Count	129	17	24	5	7	12	194	
	% within OLD CLASSIFICATION	66.5%	8.8%	12.4%	2.6%	3.6%	6.2%	100.0%	

APPENDIX 7

What Variables are the Key ones?

From the PLASC database, 15 variables were analysed. Which of these was more or less important in allocating Schools to Clusters? The statistical tool to examine this is called Discriminant Analysis.

First, a Principal Factor Analysis with a Varimax rotation suggests that all the 15 PLASC variables except three variables are influential. The three variables that are different and might be excluded in future classifications are: (i) the percentage of KS3 Classes where Welsh is a part medium for teaching (ii) the percentage of KS4 Classes where Welsh is a part medium for teaching and (iii) the percentage of teachers in a school who are qualified teach Welsh or through the medium of Welsh but are not doing so. The results are below:

	Factor	
	1	2
(WelshMediumSubjects)	.800	.185
(WelshGovernors)	.894	.248
(CLASSKS3MAIN)	.963	-.086
(CLASSKS3PART)	.069	.938
(CLASSKS4MAIN)	.961	-.096
(CLASSKS4PART)	.077	.926
(TEACHER_FT_TCTO)	.958	.148
(TEACHER_PT_TCTO)	.834	.182
(TEACHER_FT_NW)	.080	.316
(PUPIL_FLUENT)	.924	.178
(PUPIL_NOTFLUENT)	-.685	-.061
(WELSHATHOME)	.804	.222
(STUDYOFWELSHL1)	.972	.187
(STUDYOFWELSHL2)	-.961	-.194
(WELSHMEDIUM)	.875	.108

Second, Discriminant Analyses were conducted across the 3 Cluster, 4 Cluster, 5 Cluster and 6 Cluster solutions to show the **order of importance** of the 15 variables in the final solution. The results are below. Standardized (z) scores were used in these analyses as they were in the Cluster Analyses.

Only the variables that were important in separating Schools into clusters are listed.

THREE CLUSTERS

Discriminant Variables in order of importance in creating three clusters:

(CLASSKS3MAIN)
(CLASSKS3PART)
(WelshGovernors)
(CLASSKS4MAIN)
(STUDYOFWELSHL1)
(CLASSKS4PART)
(TEACHER_FT_TCTO)
(TEACHER_PT_TCTO)
(WELSHMEDIUM)

FOUR CLUSTERS

Discriminant Variables in order of importance in creating four clusters:

(CLASSKS4MAIN)
(CLASSKS4PART)
(STUDYOFWELSHL1)
(CLASSKS3PART)
(WELSHMEDIUM)
(CLASSKS3MAIN)
(TEACHER_PT_TCTO)
(TEACHER_FT_TCTO)

FIVE CLUSTERS

Discriminant Variables in order of importance in creating five clusters:

(STUDYOFWELSHL1)
(CLASSKS4PART)
(CLASSKS4MAIN)
(CLASSKS3PART)
(TEACHER_FT_TCTO)
(WelshGovernors)
(WelshMediumSubjects)
(CLASSKS3MAIN)
(TEACHER_PT_TCTO)

SIX CLUSTERS

Discriminant Variables in order of importance in creating six clusters:

(STUDYOFWELSHL1)

(CLASSKS3PART)

(CLASSKS3MAIN)

(TEACHER_FT_NW)

(TEACHER_PT_TCTO)

(TEACHER_FT_TCTO)

(CLASSKS4PART)

(CLASSKS4MAIN)

(WELSHATHOME)

(WelshMediumSubjects)

(WelshGovernors)
