

---

1 **What Value ‘Value Added’?**

2 **ANDREW RICHARDS**

3 **ABSTRACT** Two quantitative measures of school performance are currently used, the  
4 average points score (APS) at Key Stage 2 and value-added (VA), which measures the  
5 rate of academic improvement between Key Stage 1 and 2. These figures are used by  
6 parents and the Office for Standards in Education to make judgements and comparisons.  
7 However, simple statistical analysis suggests that the measures are correlated and,  
8 therefore, schools with high APS values have high VA. This calls into question whether  
9 the measures are objective and valuable as a means of assessing a school’s efficacy.

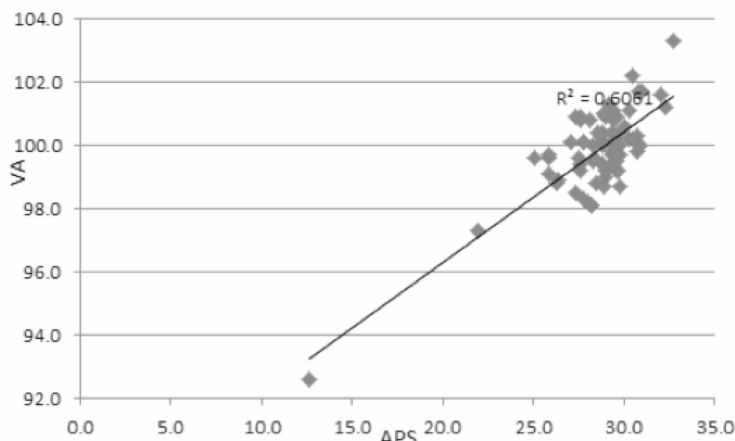
10 Two principal measures are commonly used to quantify a primary school’s  
11 ability to improve children’s academic performance. Key Stage 2 APS gives the  
12 average points score per pupil in the English and Mathematics Standard  
13 Assessment Tasks (SATs). The sum of the points that pupils achieved in the two  
14 subjects is then divided by the number of pupils eligible to take the tests. The  
15 concept of ‘value added’ (VA) attempts to measure the gain (or loss) from being  
16 in a school with respect to an average school; the average performance of  
17 schools taken from a data set. It provides a data-driven measure of school  
18 efficiency (Raudenbush, 2004; Timmermans et al, 2011). Contextual value  
19 added (CVA) measures the individual characteristics of a cohort against other  
20 measures such as gender, whether children are in receipt of government  
21 financial support, ethnicity or children with special education needs and/or  
22 disabilities. Value added is calculated for each student and then aggregated to  
23 give a score for the school. In this way an individual student’s progress is  
24 compared with the progress made by other students with the same or similar  
25 prior attainment.

26 The argument for using both values is that they report different things;  
27 the APS measure only reflects children’s attainment, and not progress. This  
28 means that it can be influenced by a range of external factors that do not reflect  
29 the capacity of a school in encouraging children’s academic development. VA  
30 recognises that similar types of children should show similar growth in their

31 skills and knowledge. Those children that show more growth than average  
32 must have had better teachers, while those with less growth than the average  
33 must have poorer teachers (McCaffrey et al, 2003; Rockoff & Speroni, 2010;  
34 Hanushek & Rivkin, 2010; Manzi et al, 2014).

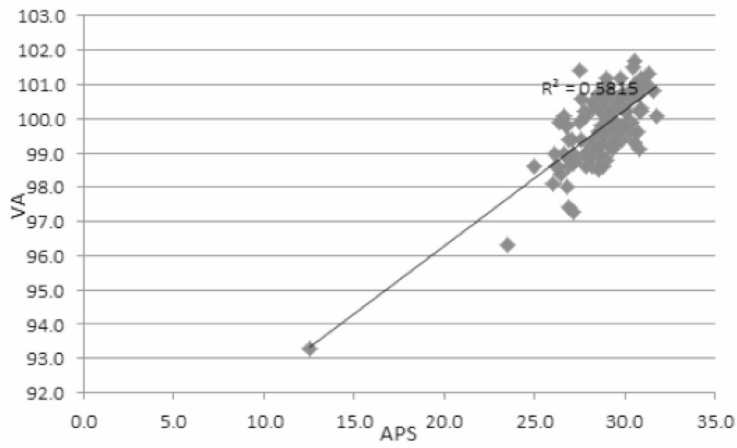
35 Parents and carers are encouraged to refer to these measures when they  
36 are deciding on where their children should go to school by both government  
37 and non-governmental organisations (i.e. The Good School Guide, BBC). The  
38 Office for Standards in Education (Ofsted) often uses the measures to  
39 pre-determine the results of school inspections and league tables are published  
40 in most broadsheets as soon as the data becomes available, readily comparing  
41 the efficacy of different schools. In the USA, value-added measures are  
42 frequently used to measure a teacher's impact – a trend that has caused a great  
43 deal of unrest (David, 2010; Berliner, 2013; *The Economist*, 2013).

44 When plotted against each other, Key Stage 2 APS and overall VA  
45 measures produce some interesting results. Figures 1a, b and c show the 2014  
46 data for primary schools in Herefordshire, Shropshire and Birmingham  
47 respectively. Figure 2 shows the 2014 data for every primary school in England  
48 (excluding those where less than 10 children sat the Key Stage 2 tests). R-  
49 squared is a statistical measure of how close the data are to the fitted regression  
50 line, or the percentage of the response variable variation that is explained by a  
51 linear model. 0 indicates that the model explains none of the variability of the  
52 response data around its mean while 1 indicates that the model explains all the  
53 variability of the response data around its mean. Therefore, in general, the  
54 higher the R-squared, the better the model fits the data. As a further measure,  
55 the Pearson product-moment correlation coefficient was calculated for the two  
56 sets of values as a separate measure using the covariance of the standard  
57 deviation from the mean.



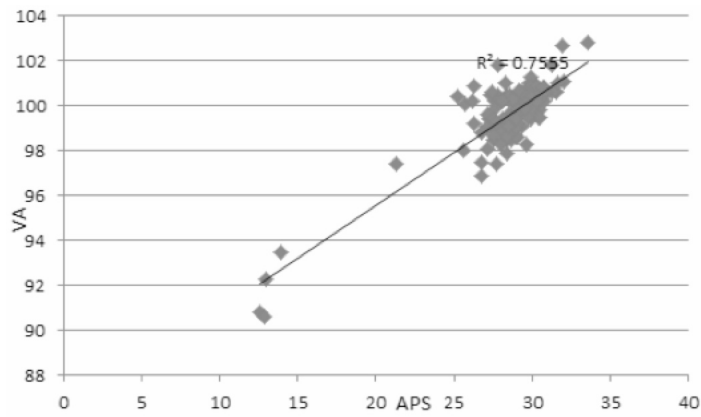
58  
59 Figure 1(a). Scatterplot VA vs. KS2 APS: Herefordshire (correlation coefficient: 0.778).  
60

61  
62  
63



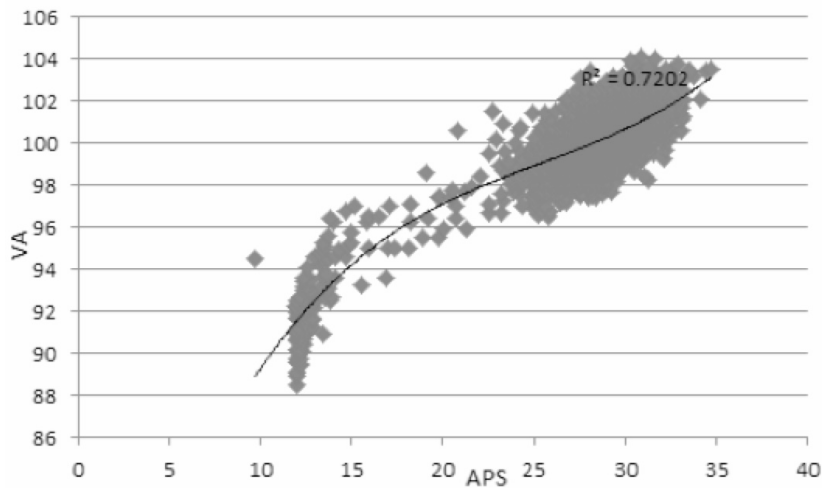
64  
65  
66  
67  
68

Figure 1(b). Scatterplot VA vs KS2 APS: Shropshire (correlation coefficient: 0.763).



69  
70  
71  
72  
73  
74  
75  
76

Figure 1(c). Scatterplot VA vs KS2 APS: Birmingham (correlation coefficient: 0.869).



77

78 Figure 2. Scatterplot VA vs KS2 APS: England (correlation coefficient: 0.841).

79

80 When the national indices are compared (Figure 2), it seems that a ‘normal’  
81 linear fit does not suffice, while a 3-power polynomial or curvilinear model is  
82 more appropriate – as at a local level. Not only do the Key Stage 2 APS and VA  
83 measures appear to be correlated, the measures involve some statistical  
84 relationships that are not applicable for all schools – and in this case, suggesting  
85 that a cohort with low APS at Key Stage 1 is predetermined to make progress at  
86 a different rate to those with APS values that are more comparable with the  
87 national trend. When the data set that includes the 4500+ primary schools in  
88 England is analysed, different rates of progress are applicable. While this is  
89 relevant regarding the way that the VA assessment measures these schools, it  
90 also questions the validity of using both Key Stage 2 APS average and the VA  
91 measure as a collective, and independently relevant, measure of a school’s  
92 performance.

93

94 *Questions:*

- 95 • Do the indices offer anything different in terms of measuring that
- 96 effectiveness of a school?
- 97 • Is the VA measure independent of children’s inherent abilities, the influence
- 98 of localised socio-economic factors and/or the power of a school’s ability to
- 99 influence these indices through pedagogic practice?
- 100 • If we use similar models for assessing teacher efficacy as in the USA, will
- 101 there be sufficient awareness of problems associated with whole-population
- 102 heterogeneity and in-school endogeneity that cannot accurately be measured
- 103 using current models?

104        **References**

- 105        Berliner, D.C. (2013) Problems with Value-Added Evaluations of Teachers? Let Me  
106        Count the Ways!, *The Teacher Educator*, 48, 235-243.  
107        <http://dx.doi.org/10.1080/08878730.2013.827496>
- 108        David, J.L. (2010) What Research Says about Using Value-added Measures to Evaluate  
109        Teachers, *Educational Leadership*, 67(8), 81-82.
- 110        *Economist* (2013) Value-Added Remodelling, March 16.
- 111        Hanushek, E.A. & Rivkin, S.G. (2010) Generalisations about Using Value-added  
112        Measures of Teacher Quality, *American Economic Review*, 100, 267-271.  
113        <http://dx.doi.org/10.1257/aer.100.2.267>
- 114        Manzi, J., Marton, E.S & Bellegem, S.V. (2014) School System Evaluation by Value-  
115        added Analysis under Endogeneity, *Psychometrika*, 79(1), 130-153.  
116        <http://dx.doi.org/10.1007/s11336-013-9338-0>
- 117        McCaffrey, D.F., Lockwood, J.R., Koretz, D.M. & Hamilton, L.S. (2003) *Evaluating*  
118        *Value-Added Models for Teacher Accountability*. **LOCATION OF PUBLISHER?**  
119        Rand Education.
- 120        Raudenbush, S.W. (2004) *Schooling, Statistics, and Poverty: can we measure school*  
121        *improvement?* William H. Angoff Memorial Lecture Series. **TOWN/CITY?**  
122        Australia: Educational Testing Service.
- 123        Rockoff, J.E. & Speroni, C. (2010) Subjective and Objective Evaluations of Teacher  
124        Effectiveness, *American Economic Review*, 100, 261-266.  
125        <http://dx.doi.org/10.1257/aer.100.2.261>
- 126        Timmermans, A., Doolaard, S. & de Wolf, I. (2011) Conceptual and Empirical  
127        Differences among Various Value-Added Models for Accountability, *School*  
128        *Effectiveness and School Improvement*, 22, 393-413.  
129        <http://dx.doi.org/10.1080/09243453.2011.590704>

130

131

132        **ANDREW RICHARDS**, Marlbrook Teaching School Alliance, Hereford.  
133        *Correspondence:* [arichards@littledewchurch.hereford.sch.uk](mailto:arichards@littledewchurch.hereford.sch.uk)