The 2009 VARK SCORING TRIAL

After an informative meeting with Dr Walter Leite, psychometrician and Assistant Professor at Florida University in 2009 we decided to trial a new scoring algorithm. The existing algorithm was based on arithmetic differences between each respondent's VARK scores noting the differences between his/her highest scores and the next-ranked scores relative to the total score for all four.

Dr Leite suggested that if VARK was to be used more reliably in research projects it needed a statistically-based scoring system. He suggested checking each score for the VARK questionnaire with the means and standard deviations from the extensive VARK database. If the score, for say Visual, was more than three standard deviations above the mean for all Visual scores from the database it could be allocated a *very strong Visual* category. A score between two and three standard deviations could be allocated a *strong Visual* category. A score between one and two standard deviations would be allocated a *mild Visual* category. Similarly for a respondent's Aural, Read/write and Kinesthetic scores.

It would be possible that a respondent could have a *mild Visual* preference, a *strong Aural* preference and a *very strong Kinesthetic* preference. The means and standard deviations were to be selected on the basis of the respondent's total score for all four modes so that all those in the database with a total score of say, 24, would form the source for the four means and four standard deviations. This is now called the Research algorithm and has its own spreadsheet for analysis. It can be purchased online. Means and standard deviations for it will be regularly updated.

The trial began in March 2009 and ended on 1 May 2009. During that time those who responded to the VARK questionnaire online, received a profile description based on the Research algorithm. The distribution of VARK categories changed quite dramatically as shown in Table One. The data from September 2008 were also used to study the changes and are shown here for comparison.

TABLE ONE : Changes between the Standard and the Research Scoring Systems

	Research	**Standard**	**% Change. Research minus Standard**
Visual single preferences. (Mild, Strong and Very Strong)	12.7%	3.0%	9.7%
Aural single preferences. (Mild, Strong and Very Strong)	12.9%	7.6%	5.3%
Read/write single preferences. (Mild, Strong and Very Strong)	14.7%	14.1%	0.6%
Kinesthetic single preferences. (Mild, Strong and Very Strong)	13.4%	12.5%	1.0%
Bi-modes (VA, VR, AR, AK, RK and VK).	5.3%	15.5%	-10.2%
Tri-modes (VAR, ARK, VRK and VAK).	0.0%	12.5%	-12.5%
VARK (All four modes).	40.9%	34.9%	6.0%
	100.0%	100.0%	0.0%

In Table One we noticed the large shift into the VARK category and the single preferences of Visual and Aural and the smaller additions to the Read/write and Kinesthetic categories. In a zero-sum table, these increases had to be matched by reductions elsewhere and that can be seen in the loss of almost all the trimodal categories and a 10% reduction in the bi-modes. (There are a small number of tri-modes but less than one-tenth of a percent). The Research algorithm allocated very few respondents as having two and three preferences.

A single preference became more common as it was relatively easy for a respondent to have one standout score in a single mode - the respondent's total would be relatively low and therefore the means and standard deviations at that level were *"attainable."* To have two or three modes meant

that the total scores were comparatively higher and therefore the algorithm was feeding off a total line that was substantially higher and therefore the appropriate standard deviations and means were less likely to be reached.

The ten respondents who had the tri-mode (VAK preferences) under the new Research scoring had the scores as shown in Table Two below. Note that in six instances their fourth mode had a comparatively low score. All were students. Six were aged under 18 and four under the age of 25. Under the Standard scoring system that existed in September when this data was collected, eight said that the VAK categorisation *"matched"* their perception of how they learned and the other two chose *"do not know"* how I learn). There were no other tri-modes (ARK, VAR or VRK) in the Research scoring data, signalling that it is difficult to reach a preference for the Read/write (R) mode.

TABLE TWO: Tri-modes

V	A	R	K	Total
11	12	0	12	35
12	12	1	13	38
12	13	2	13	40
13	13	1	14	41
16	13	0	13	42
14	13	2	14	43
15	15	6	15	51
15	15	7	15	52
15	15	7	15	52
16	16	9	16	57

In Table One we noted the changes that took place between the two scoring systems. Table Three below shows where some of those changes came from. Notice that a large number didn't change their profile. Columns three and four show the numbers and percentages for all the single preferences and the VARK category. Read/write and Kinesthetic profiles under the Research scoring were very similar to those under the Standard scoring with 69% and 76% staying with the same categorisation. VARK and Aural had the same designation for over 50% of their respondents but only 21.6% of those with a Visual profile under the Research algorithm had a Visual profile under the Standard scoring system.

TABLE THREE: Sources of Change between the Standard and the Research Scoring Systems
(Shifts to and from Bi- and Tri-modes are not shown). Data Used: September 2008 n=62603

Mode	Research Scoring	Profile stayed the same as for Standard scoring	% Same	Profile changed Net Increase	% Increase
V Single preference	7973	1724	21.62%	6249	78.38%
A Single preference	8077	4200	52.00%	3877	48.00%
R Single preference	9185	7019	76.42%	2166	23.58%
K Single preference	8415	5844	69.45%	2571	30.55%
VARK multimodal preference	25606	13996	54.66%	11610	45.34%

The Visual single preference had a dramatic increase in size with an additional 6249 (78% increase) being categorised with that preference. The Aural single preference also increased markedly (48%) as did the VARK preference (45%). To track where these increases came from, we assembled Table Four.

TABLE FOUR: SOURCES FOR INCREASES

	SOURCE OF INCREASE (numbers) These five columns sum to the Net Increase column above.				
Mode	From Bi-modes containing their mode.	From Tri-modes containing their mode.	From VARK	From other single preferences	Another Source
V single preference	1915	1743	2481	110~	0
A single preference	1287	654	35	0	1901@
R single preference	452	293	1421	0	0
K single preference	590	439	1542	0	0
VARK multimodal	3858*	4253#		3499^	0

~ From R and K mild.
@ From VK.
* For VARK this should read *"From all Bi-modes."*
# For VARK this should read *"From all Tri-modes."*
^ For VARK this should read *"From Single Preferences."*

The Visual single preference, with its large increase in numbers, gained most from re-categorising the VARK modes. Presumably there were 2481 respondents who under Standard scoring had no particular preference (VARK) but who had a Visual single preference under the Research scoring algorithm. The criteria for a Visual preference under the Standard scoring is set much higher than under the Research scoring. The Visual preference also gained 3658 from a reduction in those with bi- and tri-mode profiles. It was the only single preference to *"steal"* some from other single preferences – all 110 came from *Read/write* and *Kinesthetic mild*. No explanation is offered for this.

The Aural preference gained most (1901) by re-categorising the bimodal VK preference as an Aural preference. Its other changes came from bi-modes that had some Aural in them (AR, AK, VA) (1287) or from tri-modes with some Aural in them (ARK, VAR, VAK) (654). Only 35 were redesignated from VARK.

The Read/write preference grew by almost a quarter at the expense of re-designations from VARK, bi-modes and tri-modes in that order. The Kinesthetic profiles increased by 30% from the same sources as for Read/write.

The VARK profiles that previously made up 35% of the total Standard database moved to 41% under the Research scoring. Almost one half stayed the same under both systems. As shown above, VARK was a heavy loser to single preferences but, as a profile, it gained from redesignating 3499 back from single preferences and 8000 from bi- and tri-modes.

TABLE FIVE: CHANGES IN BI-MODES AND TRI-MODES

	New	Old	**Difference O-N**
Bi modes	3323	9702	-6379
Tri Modes	10	7816	-7806

Table Five shows that the bi-modes and tri-modes lost large numbers when categorised by the Research scoring system with only ten left in the tri-modes from a total database of some 62000.

Of the Bi-modes the numbers were split between them as shown below in Table Six. They make up only 5% of the total database.

TABLE SIX: Bi-modes

Mode	Number
VA	322
VR	468
VK	1185
AK	816
AR	341
RK	202

CONCLUSION: Pros and Cons

We were not able to find out from respondents which scoring system they would have preferred for their own categorisation. Would those who had a VARK profile under the standard scoring have seen that as a better (or worse) match to their perceptions of how they learned? That would be an interesting research exercise for the future. There is no definitive answer to what is the correct categorisation of a population and it should always be noted that the VARK online questionnaire is already biased towards those who use computers and who are students and teachers. Maybe it is also biased towards those who have a Read/write preference!

While the Research scoring system has a sound statistical rationale, it reduced the multimodal variety significantly and enhanced the single preferences especially of Visual and Aural. It increased the four-mode profile (VARK) to 40% of the total database. Its strength is that it is more amenable to statistical manipulation and would allow the use of more sophisticated statistical tools. It compares a respondent's score for a particular mode with other respondents who have the same total for their four scores. It takes little notice of the relationships between the scores of an individual. Those undertaking a major research project need to beware that the two algorithms produce quite different categorisations and norms and are conceptually different.

For the online scoring system we decided to continue with the Standard algorithm that has been used since 2001. This system has an advantage. It compares a respondent's scores only with himself/herself. There is no external reference. While that may seem strange, it means that a high number in a set of VARK scores is high for that respondent regardless of whether it is high or significant for the whole database population of not. What matters is the size of that score as compared with the total for that respondent, not the size of that score compared with the other individuals in the database. The standard scoring system is therefore more personal but less objective.

The standard system is based on a *"row"* of scores where a respondent's four scores and total can be computed into a VARK category. The research system is based on a *"column"* of scores where a respondent's four scores are compared with other respondents' four scores and computed into a VARK category. As an example here are a respondent's scores showing both systems.

Algorithm used	**Visual**	**Aural**	**Read Write**	**Kinesthetic**	**Total**	**Categorisation**
Research	4	4	8	4	20	VARK

Standard	4	4	8	4	20	Strong R

The research algorithm investigates whether the score of 8 is *"unusual"* when compared with other respondents' scores for Read/write. It interrogates the database to make that decision and decides that 8 is not in the range of scores that it should categorise as *mild,* or *strong,* or *very strong Read/write*. The standard algorithm investigates whether the score of 8 for Read/write is *"unusual"* among the other scores for that individual and compared with the end row total (20). It interrogates the other three scores and decides that 8 is different and decides to categorise that individual as having a s*trong Read/write* preference.

Using the standard system in a workshop setting, participants can quickly note whether they have a high score (or high scores) relative to their total score. They have more difficulty when their highest score is not significant when compared with others who are not in the workshop but online or elsewhere. In that sense VARK, as Marilla Svinicki notes, provides an *instructional* tool supporting the workshop aims which are to encourage students to use their own preferences (not somebody else's) and for teachers and trainers to provide more variety in their presentations.

There is no right or wrong in this process as the *"correct"* categorisation of a population's modal preferences are not known. We were somewhat swayed in our decision to stay with the standard system for our online questionnaire because of the reasons above and because the *"match"* statistic decreased when the research system was used. Having over 60% of respondent identify their categorisation as a *"match"* to their perception of how they learn is a hallmark of VARK and we were reluctant to move to a system where that declined by 5-7 percentage points.

Neil D Fleming
Designer of the VARK questionnaire
May 2009