A note on the calculation of the Garskof-Houston relatedness coefficient

Una nota sobre el cálculo del coeficiente de relaciones de Garskof-Houston

LIBERATO CARDELLINI

Università Politecnica delle Marche, Via Brecce Bianche, Ancona, Italy, libero@univpm.it.

Abstract

The Word Association Test is an educational tool that enables the teacher to 'look inside the students' heads and see how the concepts are lodged in their cognitive structure and also to see the map of the connections between the concepts. In order to map the structures, the relatedness coefficient must be calculated. Because a different way of calculating was found in the literature, in this paper the calculations are performed according to the formula of Garskof & Houston.

Key words: *Word Association Test, cognitive structure, relatedness coefficient, rank order.*

Resumen

La prueba de asociación de palabras es una herramienta que facilita al docente 'ver dentro de la cabeza de los estudiantes, y comprender cómo los conceptos se encuentran ubicados en la estructura cognitiva y asimismo ver el mapa de las relaciones entre los conceptos. Para elaborar el mapa de las estructuras, el coeficiente de relaciones debe ser calculado. Teniendo en cuenta que, hay varias formas de calcular, en este trabajo los cálculos se realizaron con la fórmula de Garskof y Houston.

Palabras clave: *prueba de asociación de palabras, estructura cognitiva, coeficiente de relaciones, orden de rango.*

INTRODUCTION

In their book, *Probing Understanding*, WHITE and GUNSTONE (1992) present a variety of interesting methods for improving the understanding of science in our students and to probe the effectiveness of teaching. Between them, Word Association Testing is presented. This test is based upon the idea that, if a person is given a primer (starter word) and asked, under pressure to say the first word which comes to mind, the idea most strongly associated with the starter word will be revealed in the response. If the same starter is given again and a response is requested, it will represent a more distant association. The process can be repeated several times to give a string of associations, of decreasing strength, with the primer. A new primer can now be chosen and the process repeated and so on. If the primers are the key concepts in a particular area of study, an association map of how that study is laid down in the student's mind can be revealed. If two primers are associated with a very similar string, it is likely that these two starter words are strongly associated with each other. If, on the other hand, two primers share few associations, it is likely that these two starters are only weakly associated with each other.

WINDOWS INTO THE MIND

The test affords a map of the cognitive structure of a particular area of study, for example general chemistry. By a "structure" we mean an assemblage of identifiable elements and the relationships between those elements (SHAVELSON, 1972). "Cognitive structure" is an hypothetical construct referring to the organization (relationships) of concepts in memory (SHAVELSON, 1974). To learn about the cognitive structure is important because we can understand to what extent does the structure in the student's memory overlap with the subject structure, and in what measure it helps the student in problem solving or in the development of higher order cognitive skills.

In a study conducted on 85 freshman students enrolled for chemistry at the Faculty of Engineering in the now Polytechnic University of the Marche (CARDELLINI & BAHAR, 2000), ten key words from the main concepts of the General Chemistry curriculum where chosen to act as stimuli for the test. These words were *Reaction (1), Equilibrium (2), Chemical bond (3), Rate of reaction (4), pH (5), Oxidation-reduction (6), Molecule (7), Solution (8), Physical state of Matter(9), Atom (10)*. They were written at the top of the page and ten times down the left side of the page so that subjects were encouraged to return to the stimulus word after each association, in order to minimize the chain effect, in which each response, instead of the key word, becomes the stimulus for the next response (Bahar, Johnstone & Sutcliffe, 1999).

For each stimulus word, students were required to list up to ten words which they considered to be associates with the stimulus word, within thirty seconds and the time was controlled by the experimenter. This constraint is because research evidence suggests that a free association cannot reveal weak direct associations (McNamara, 1992). For example, a student response to the primer 'Chemical bond' was the following sequence of associations: 'covalent', 'ionic', 'polar', 'hydrogen', 'strength', 'simple', 'double', 'triple', 'coordinate'.

We can count the number of meaningful words as students' responses to each key word and this can be method for analysing the word association data (Shavelson, 1974). Using the procedure reported below, we can calculate a relative index of the overlaps between the key words, the Relatedness Coefficient, and we can obtain the cognitive structure of a single student and the cognitive structure of the class. To show some indications of the relative strength of the associations, some cut-off points were chosen: figure 1 reports the cognitive structure of the students at a cut-off point of 0.2.

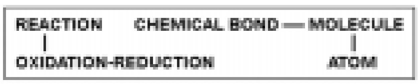

Figure 1. The cognitive structure of the students at a cut-off point of 0.2. (Cardellini & Bahar, 2000, reported with the permission from the editor)

This relatively strong relatedness coefficient reflects a strong interconnect ness between the keywords as revealed from the responses: the majority of students can associate the right words to the concepts 'Reaction' and 'Oxidation-reduction' and to the concepts 'Chemical bond', 'Molecule' and 'Atom'. Lowering the cut-off point at 0.1, a connection between the Equilibrium and Rate of reaction appears; still the words are separate "islands", as shown in figure 2:

Figure 2. The cognitive structure of the students at a cut-off point of 0.1.

The few isolated "islands" very slowly begin to come together as a reasonable interconnected network only at the weakest RC value of 0.05; at this level, all the 10 primers are connected, as shown in Figure 3. A RC value of 0.05 is far from 1; this means that the complete connections between these concepts in the mind of the students are quite weak, or just few students "see" this network in their mind.

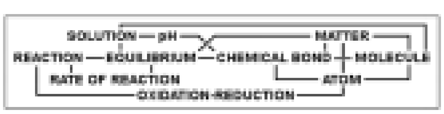

Figure 3. The cognitive structure of the students at a cut-off point of 0.05.

What emerges from those maps is that the students are not likely to see all the concepts represented by the ten key words, as linked together. The responses obtained to each primer can be analysed in several ways, usually by deriving for each student a matrix representing the similarities between the responses to each pair of primers. It is also possible to look for the relations between the students' cognitive structures using the Waern's technique (Waern, 1972). From the students' response to every primer, a frequency table can be obtained: the frequencies vary from 187 to 1 for the 285 different words. For example, the word 'ionic' was associated 2 times to the primer *1*, 24 times to *2*, 71 times to *3*, 54 times to *5*, 7 times to *6*, 4 times to *7*, 17 times to *8*, 1 time to *9* and 7 times to *10*. In a similar way, a count was made for every word. From the frequency table we consider three different groups of students using three arbitrary lines of separation: 68, 42 and 17, considering in this way 80%, 50% and 20% of students and so, the corresponding "maps" of concepts can be obtained. (Cardellini & Bahar, 2000)

Few words are frequently associated to the key words: 68 students associated the words 'electron' to 'atom', 'ionic' and 'covalent' to 'chemical bond', 'solid' to 'matter', 'acidic' and 'basic' to 'pH'. By lowering the fraction of students considered (42 students), other associations appear, some "islands" slowly begin to come together and all precedent "islands" become richer (see figure 4).

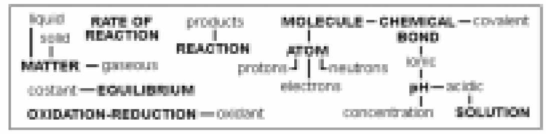

Figure 4. The interconnect ness between key and response words for 50% of students.

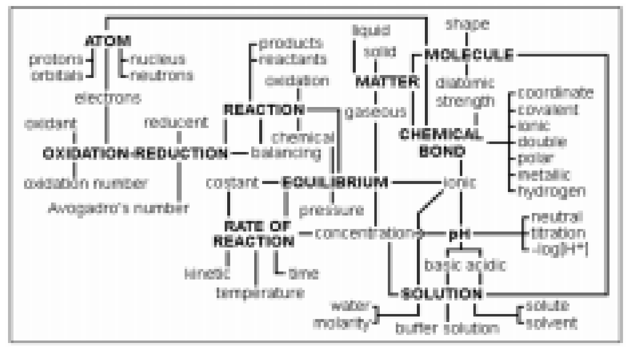

Figure 5. The interconnect ness between key and response words showed by 20% of students.

Eventually, the whole picture emerges, at the weakest level of cut-off point. It is necessary to mention that the whole network of the words at the 20% level does not mean that 17 students out of 85 have this network between the concepts. Some of the students may have some part of the connections, but not necessarily all. From Figure 5, it also emerges that the key word 'Chemical bonding' has the greatest number of associated words; this can be an indication that students know this topic better than the other topics.

Because we can know the interconnect ness between the concepts for every student, this method is useful for monitoring the acquired knowledge of our students and affords to discover the semantic network in the student's long-term memory. This test can be a probe for the teacher: choosing the appropriate key words he can monitor the understanding of the subject by the students, for every topic and the evolution during the course. This test can be a tool for teachers to make their teaching more effective.

This test is valid for externalizing the structure of declarative knowledge and has been used by several researchers in science of all kinds (Johnson, 1965; Shavelson, 1972; Shavelson, 1975; Geeslin & Shavelson, 1975; Thro, 1978; Bahar, Johnstone & Sutcliffe, 1999), and several studies have been performed in chemistry (Johnstone & Moynihan, 1985; Gorodetsky & Hoz 1985; Cachapuz & Maskill, 1987; Gussarsky & Gorodetsky, 1988; Maskill & Cachapuz, 1989; Cachapuz & Maskill, 1989; Cardellini & Johnstone, 2000).

THE CALCULATION OF THE RELATEDNESS COEFFICIENT

A relative index of the overlaps between pairs of primers can be obtained using the Garskof and Houston's formula (Garskof & Houston, 1963). The Relatedness Coefficient (RC) obtained can range from 0 (totally unrelated) to 1 (perfect relatedness; the same word) and measures the number of identical words given as responses to two key words and their rank order. The ranks of the words are the numbers beside them in Table 1. In a word association test, the degree of overlap between the response hierarchies is a measure of the semantic proximity, or the "strength" of links between concepts between the primers in memory (Deese, 1962; Rips, Shoben & Smith, 1973).

Now we consider some examples of word association and we calculate the RC values.

$$RC = \frac{\bar{A}\cdot\bar{B}}{(A\cdot B) - [n^P - (n-1)^P]^2}$$

where:

- $\bar{A}$ represents the rank order of occurrence of words under A which are shared in common with $\bar{B}$ and $\bar{B}$ represents the rank order of words under $\bar{B}$ which are shared in A.
- $\bar{A} \cdot \bar{B}$ represents the sum of products of the rank order of every word in $\bar{A}$ multiplied by the rank order of the same word in $\bar{B}$.
- $A\cdot B$ represents the rank order of words in A multiplied by the rank order of words in B, considering the longer of the two lists.
- n represents the number of words in the longer list.
- P represents a fixed number greater than zero which may be determined from the shape of the probability distribution of the responses. P is usually set equal to 1 (Geeslin & Shavelson, 1975); in this way all portions of the student's responses distribution received equal weight. In this case the formula becomes:

$$RC = \frac{\bar{A}\cdot\bar{B}}{(A\cdot B) - 1}$$

The denominator can be indicated as $\Sigma n^2 - 1$. The formula can also be written the form:

$$RC = \frac{\text{Sum of products of ranks of common words}}{\Sigma n^2 - 1}$$

White and Gunstone (1992, p. 145) present an example of calculation of RC; using their example:

Associations	Rank	Associations	Rank
BASEBALL	7	CRICKET	6
run	6	run	5
bat	5	wicket	4
field	4	bowl	3
throw	3	field	2
base	2	baseball	1
bleachers	1		

From the calculation they obtain: RC = 0.32. They rank the words starting from the last word in the list; the usual way of forming ranks in general. But, with verbal associative data, Garskof and Houston (Garskof & Houston, 1963), basing on extensive psychological research, gave priority to words associated with a key word and to rank order starting from the first word (primer). According to Garskof and Houston: "The procedure of giving equal weight to the first associate is based upon the assumption that the importance of the first associate in a small hierarchy is at least as great as the importance of the first associate in a larger hierarchy" (p. 281).

The different method of ranking the associated words brings to different RC values: let us consider an example using primers from the domain of General Chemistry. A student response to the primer 'Reaction' was the following sequence of associations: 'reactants', 'products', 'exothermic', 'balancing'. To the primer 'Oxidation-reduction': 'oxidant', 'balancing', 'reaction'. We have:

Associations	Rank	Associations	Rank
reaction	5	oxidation-reduction	5
reactants	4	oxidant	4
products	3	balancing	3
exothermic	2	reaction	2
balancing	1		

In the two lists, there are two common words: 'reaction' with rank 5 when associated to 'Reaction' and rank 2 when associated to 'Oxidation-reduction' and 'balancing', with rank 1 when associated to 'Reaction' and rank 3 when associated to 'Oxidation-reduction'.

Words overlapping: 'reaction', 'balancing'.

$\bar{A}$ = [5 1], i.e. the rank order of the common words in the 'Reaction' list.
$\bar{B}$ = [2 3], i.e. the rank order of the common words in the 'Oxidation-reduction' list.
A = [5 4 3 2 1] = B

$$RC = \frac{[5\ 1]\begin{bmatrix}2\\3\end{bmatrix}}{[5\ 4\ 3\ 2\ 1]\begin{bmatrix}5\\4\\3\\2\\1\end{bmatrix} - 1} = \frac{5\cdot2 + 1\cdot3}{5\cdot5 + 4\cdot4 + 3\cdot3 + 2\cdot2} = 0.24$$

Instead, using the White and Gunstone ranking method, we have:

Associations	Rank	Associations	Rank
reaction	5	oxidation-reduction	4
reactants	4	oxidant	3
products	3	balancing	2
exothermic	2	reaction	1
balancing	1		

$\bar{A}$= [5 1], $\bar{B}$ = [1 2]; RC = 0.13.

If the difference in the length between the two lists is greater, the difference between the two methods of calculation is greater. For example, if a student's response to the primer 'Reaction' was the following sequence of associations: 'equilibrium', 'reactants', 'products', 'balancing', 'exothermic', 'limiting reactant', 'moles', 'molecules', 'atoms'. To the primer 'Equilibrium'; 'reaction'. We have:

Associations	Rank	Associations	Rank
reaction	10	equilibrium	10
equilibrium	9	reaction	9
reactants	8		
products	7		
balancing	6		
exothermic	5		
limiting reactant	4		
moles	3		
molecules	2		
atoms	1		

$\bar{A}$ = [10 9], $\bar{B}$ = [9 10]; RC = 0.46.
With the White and Gunstone ranking method, $\bar{A}$ = [10 9], $\bar{B}$ = [1 2]; RC = 0.071.
In the above White and Gunstone example, using the Garskof and Houston method, we obtain RC = 0.45.

CONCLUSIONS

The Word Association Test is an instrument that can be used by teachers for assessing student learning or for formative purposes in ordinary classrooms. From the continuous word association hierarchies, the cognitive structure is usually inferred using Garskof and Houston's formula, that allows the calculation of the Relatedness Coefficient, which is assumed as a proximity index and a measure of the overlap between two key words. As a consequence, it is important that the calculations are performed in a correct way in order to reveal the interconnectedness between the concepts as they are in the student's mind.

ACKNOWLEDGEMENTS

I would like to thank Richard J. Shavelson, Stanford University, USA, and Richard T. White, Monash University, Australia, for commenting this paper.

BIBLIOGRAPHY

Bahar, M.; Johnstone, A.H.; Sutcliffe, R.G. Investigation of students' cognitive structure in elementary genetics through word association tests. *Journal of Biological Education*, 33 (3), 134-141, 1999.

Cachapuz, A.F.C.; Maskill, R. Detecting changes with learning in the organization of knowledge: Use of word association tests to follow the learning of collision theory. *International Journal of Science Education*, 9 (4), 491-504, 1987.

Cachapuz, A.F.C.; Maskill, R. Using word association in formative classroom tests: following the learning of Le Chatelier's Principle. *International Journal of Science Education*, 11 (2), 235-246, 1989.

Cardellini, L.; Bahar, M. Monitoring the learning of chemistry through word association tests, *Australian Chemistry Resource Book*, 19, 59-69, 2000.

Cardellini, L.; Johnstone, A.H. Il test sull'associazione di parole nelle scienze. *La Chimica nella Scuola*, 22 (1), 9-16, 2000.

Deese, J.: On the structure of associative meaning. *Psychological Review*, 69 (3), 161-175, 1962.

Garskof, B.E.; Houston, J.P. Measurement of verbal relatedness: An idiographic approach. *Psychological Review*, 70 (3), 277-288, 1963.

Geeslin, W.E.; Shavelson, R.J. An Exploratory Analysis of the Representation of a Mathematical Structure in Students' Cognitive Structure. *American Educational Research Journal*, 12 (1), 21-39, 1975.

Gorodetsky, M.; Hoz, R. Changes in the Group Cognitive Structure of Some Chemical Equilibrium Concepts Following a University Course in General Chemistry. *Science Education,* 69 (2), 185-199, 1985.

Gussarsky, E.; Gorodetsky, M. On the chemical equilibrium concept: Constrained word association and conception. *Journal of Research in Science Teaching,* 25 (5), 319-333, 1988.

Johnson, P.E. Word relatedness and problem solving in high-school physics. *Journal of Educational Psychology*, 56 (4), 217-224, 1965.

Johnstone, A.H.; Moynihan, T.F. The relationship between performances in word association tests and achievement in chemistry. *European Journal of Science Education*, 7 (1), 57-66, 1985

Maskill, R.; Cachapuz, A.F.C. Learning about the chemistry topic of equilibrium: the use of word association tests to detect developing conceptualizations. *International Journal of Science Education*, 11 (1), 57-69, 1989.

McNamara, T.P. Theories of Priming: I. Associative Distance and Lag, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (6), 1173-1190, 1992.

Rips, L.J.; Shoben, E.J.; Smith, E.E. Semantic Distance and the Verification of Semantic Relations. *Journal of Verbal Learning and Verbal Behaviour*, 12, 1-20, 1973.

Shavelson, R.J. Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63 (3), 225-234, 1972.

Shavelson, R.J. Learning from physics instruction. *Journal of Research in Science Teaching*, 10 (2), 101-111, 1973.

Shavelson, R.J. Methods for examining representations of a subject-matter structure in student's memory. *Journal of Research in Science Teaching*, 11 (3), 231-249, 1974.

Thro, M.P. Relationship Between Associative and Content Structure of Physics Concepts. *Journal of Educational Psychology*, 70 (6), 971-978, 1978.

Waern, Y. Structure in similarity matrices, *The Scandinavian Journal of Psychology*, 13, 5-16, 1972.

White, R.; Gunstone, R. *Probing Understanding*, The Falmer Press, London, 1992, Chapter 9.

Received: 28.12.2006
Approved: 29.09.2007