PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Volume 67, 2015

CLOSENESS OF FIT TO THE IDEAL: CLASSIFICATION AND SIMILARITY OF TEACHERS' REFLECTIONS USING MULTI-DIMENSIONAL SCALING

Thomas McCloughlin
Dublin City University-St. Patrick's Campus, Drumcondra, Dublin, Ireland
E-mail: tom.mccloughlin@dcu.ie

Abstract

The objectives of this research are to provide the theoretical basis for the analysis of qualitative data, such as teachers' responses to reflective questions using quantitative methods such as multidimensional scaling. The starting point is Estes' assumption that classification is basic to all intellectual activities, and since reflections are artefacts of intellectual activities, they should be amenable to classificatory methods such as multidimensional scaling. Similarity relates to distance between perceptions and MDS can utilise this feature. By way of example, N=29 teachers - graduates, experienced primary and experienced secondary - were probed for reflections on a science lesson and the results reflections analysed by the application of a scoring rubric and subsequent MDS - they were compared to a hypothetical 'ideal'. This method is discussed and future possibilities for analysing teachers' perceptions and reflections. The results of this method provide a valuable means to adapt a standard classification tool to involve a comparative element thus moving beyond mere representation of data.
Key words: *multidimensional scaling, teachers' reflections, measuring to the ideal.*

Introduction

Estes (1996) stated that "*classification is basic to all our intellectual activities"* and *"Similarity is at the very heart of a theory of cognition and memory*". These are bold statements, especially when taken outside the usual domains where classification and similarity are explicit, namely, biological classification; however, the thesis that is reexamined in this work is the applicability of classification and similarity to science education. As long ago as 1913, Thorndike (Thorndike, 1913) had assumed that the similarity between any two situations or things is measurable in terms of the proportion of common elements and that similarity is the basis for generalization, or trasnfer of learning, between tasks. Hahn and Chater (1997) went further to propose that "similarity is at the very centre [*sic*.] of a theory of concept and the theory of similarity would [*explain*] why people have the concepts that they do". A key question is: "*what is it that makes two objects similar*?" An obvious answer is that they have features in common, features that are perceptible and are concepts in themselves. A second research question is: "*what is the relationship between concepts and similarity*?" Hahn and Chater (1997) note that there is little research directly integrating the two terms. Similarity, although it may be frequently employed as an explanatory notion in the concepts literature is seldom given closer scrutiny. Likewise, models of similarity typically assume given properties and thus concepts as a starting point. If we emphasize that very close connection between concepts and similarity, it follows that categorizing an object will involve judging the similarity between that object and some other.

One of the many ways of calculating similarity is by the 'product rule' shown in Equation 1, which works well in the case of binary-valued attributes, *i.e.*, whether or not they are

ISSN 1822-7864

present. This rule was based on the multiplicative combination of cue and context and developed into a model of animal discrimination learning by Medin (1975) and then extended to the first effective formalization of the concept of instance-based categorization learning (Medin & Schaffer, 1978).

$$Sim\left(i,j\right)=s^{N-k}$$

Equation 1.

where N is the number of features; i,j are the patterns; k is the number of matches; Sim, is the pattern similarity, s is a parameter that is typically between 0 and 1 and can have its value assigned. Estes (1996) states that the product rule is widespread in models of classification due to its simple and useful properties such as the role of 'relational information' without the need for a special mechanism for the storage of such 'information'. Multidimensional scaling is a mathematical protocol that reduces data in a large set to only two 'dimensions'

The Problem Elicited

Historically, de Leeuw (2000) in a meta-review of the literature using MDS, claims that "MDS, as a set of data analysis techniques, clearly originates in psychology", and as such the early history starts with Carl Stumpf around 1880 (de Leeuw & Heiser, 1980).

The stimulus for carrying out this research stems from a perceived contradiction between doing qualitative research and using quantitative methods of analysis. From a 'first principles' basis, the separation between qualitative and quantitative research is artificial and there should not be a conflict, but rather a difference in experience and therefore expertise. Quantitative methods such as factor analysis, principal components analysis, and so on, can be used by a single individual but with a given array of concepts. Because the research is based on individuals, the research is qualitative. However, despite such an understanding, there remains a certain hostility between qualitative and quantitative 'camps'. In this work, I aim to add to the mutual understanding of the qualitative-quantitative interplay. The concepts the author deals with are not (as yet) the topical concepts or factoids of the science syllabus, but rather the science teacher's self perception and reflections of engagement with science lessons as a means of evaluation of novel learning situations. Jaworska and Chupetlovska-Anastasova (2009) reviewed the possibility of using MDS in a range of psychological domains; while (Alt, 2015) assessed students' perceptions of their learning environment with respect to justice experiences. (Ding, 2012) used MDS in studying children's early literacy, while (Vitale, Williams, Kocsis, & Kosinski, 2015) used MDS to explore teacher perceptions in special education referral categories. Jaworska and Chupetlovska-Anastasova (2009) view MDS as an 'exploratory data analysis technique' but point out its key features, namely that MDS can 'handle nominal or ordinal data, and does not require multivariate normality" while (Li & Sireci, 2013) used MDS to analyze content validity. Furthermore, participants at different times may be included in the same data set. Thus, progression can be monitored in one visual representation e.g., pre/during/post and Ryan (2014) introduces the concept of a delimiting boundary within which the 'ideal' holds a particular influence on the points on the MDS plot which she terms the Zone of Acceptable Proximity (ZAP).

Specifically, the research problem is whether it is reasonable to use MDS to characterize perceptions of teachers and include a hypothetical 'ideal' with which to reference closeness of fit (Ding, 2015; Ding & Yang, 2012) in evaluating change in behavioural preferences and coping behaviours, respectively, employed a single 'ideal' point model in MDS.

Methodology of Research

Teachers carry out prescribed teaching sequences and complete reflections of the same, the reflections are then coded using a rubric protocol and a matrix of codes constructed. The matrix for each lesson of *x* teachers by *y* questions is then subjected to a scaling protocol such as MDS which generates a 'floorplan' or 'map' of the teachers' perceptions. Such floorplans demonstrate the 'closeness' (*i.e*., similarity) of teachers to each other within a 'mathematical space', and if the same teacher is represented more than once over several iterations, the teacher can be 'tracked' through the programme of lessons. A hypothetical or ideal teacher can also be included which allows a further comparison with a 'norm'. Typically, a richer and more dynamic appreciation of the teachers' interaction with their teaching is gained from such analysis. In terms of similarity to a 'norm', teachers 'ebb and floe' as many factors are at play to engage the teacher's attention or cognitive resources irrespective of the intrinsic qualities of a lesson. A final methodological point is that the overall method is qualitative but the analysis is quantitative.

People Understand, Believe and Behave

Educational research should try to work out what and why learners understand as they do, how their beliefs are formed and how both inform activity. Humans evaluate understandings, beliefs and behaviours and they can generate huge quantities of data - how to analyse such data becomes problematic, hence the need for data reduction, and traditionally this has been done by factor analysis. However, rather than suggesting factors for the spread of results in a plot of the factors extracted in a large dataset MDS was chosen as a data reduction AND a visualization tool which related directly to the original data. MDS, as with Factor Analysis, FA, and Principal Components Analysis, PCA, involve a continuous coordinate space. However, MDS assumes that most tasks vary on several features and all are taken into account: that individuals vary along each dimension according to their ideal point: factor analysis limited at the level of the individual, and it carries the assumption of a small number of factors that can represent difference.

Analysis of the Reflective Evaluations

The evaluations were ranked according to the 4 question-rubric presented in Table 1 and the Appendix. The ranks were tabulated with case ranks forming the rows and question ranks forming the columns, thus a matrix was produced for each lesson. Thus, all the mathematics discussed here is matrical. Non-metric multidimensional scaling, also known as principal coordinates analysis, was applied to this resulting matrix and plots were made of the emerging dimensions. The statistic allows any measure of similarity to be examined, MDS can use any distance matrix and as a result, the analysis focuses on the cases, i.e., the teachers: no information is provided about the contribution of individual questions answered (Fielding, 2007). Since the data are non-metric, they are interpreted as 'distance-like,' but not actual distance. Distance is a comparable metric to similarity, and is summarized in the following equation, Equation 2, which is effectively a variation of the Minkowski distance metric:

$$d_{ijk}^2 = \sum_{a=1}^{r} w_{ka} \left(x_{ia} - x_{ja}\right)^2$$

Equation 2.

where:

is the weight for subject *k* on a dimension *a*,
is the coordinate of stimulus *i* on dimension *a*,
is the coordinate of stimulus *j* on dimension *a*.

The aim of the MDS is to transform matrical data into a set of genuine Euclidean distances. The outcome consists of an arrangement of points in a small number of dimensions, usually two, and located in such a way that the distances between the points relate as closely as possible to the dissimilarities between the objects. In non-metric MDS, only the rank order of entries in the data matrix - not the actual dissimilarities - is assumed to contain the significant information. Hence, the distances of the final configuration should as far as possible be in the same rank order as the original data. Thus, the purpose of the non-metric MDS algorithm is to find a configuration of points whose distances reflect as closely as possible the rank order of the data while demanding a less rigid relationship between the dissimilarities and the distances.

Table 1. A sample of the rubric for ranking responses to the first question given to the teachers.

Question	Descriptor	Type of reflection	Rank
Comment on the dialogue itself	The dialogue provided examples of questions that can be used again. A general comment on its ease of use. The teacher misunderstood the question or the function of the dialogue.	Reflection for action	3
	The dialogue provided the opportunity to discover how much the children knew or the teacher expresses surprise by the quality of the dialogue. A general evaluation of the dialogue	Reflection-on-action	2
	The teacher intends to use dialogue as a learning tool in the future or describes fully why the dialogue is worthwhile.	Reflection-in-action	1

It is important to note that a perfect ordinal re-scaling of the data into distances is usually not possible, rather, what is required is 'optimal scaling'. This involves finding a series of configurations, in which the inter-point distances have the closest conformity with the data. The question is, therefore, how to construct the best possible re-scaling? This can be done through *monotonic regression*, which involves the calculation of a new set of distances, usually called pseudo-distances, since they are not actual distances corresponding to any real configuration. They are also called fitted distances or disparities. The given dissimilarities δ_{ij} are used to generate a set of distances d_{ij}, which are approximately related to the given dissimilarities δ_{ij} by a monotonic increasing function *f*. Note that only the rank order is important, and the scaling is ordinal. The most common approach to determine the elements d_{ij} and the underlying configuration is an iterative process, commonly referred to as the Sheppard-Kruskal algorithm. A simplified view of the MDS algorithm is as follows (Borgatti, 1997; Cízek, Härdle, & Weron, 2005; UNESCO, n.d.), but see Fahrmeir and Hamerle (1984):

1. Assign points to arbitrary coordinates in a p-dimensional space.
2. Compute Euclidean distances among all pairs of points, to form the DHAT matrix.
3. Compare the DHAT matrix with the input D matrix by evaluating the stress function.
4. Adjust coordinates of each point in the direction that maximally reduces the stress.

After determining the dissimilarity matrix *D* and the corresponding scaling matrix *A*, an iterative process begins, which successively revises the dissimilarities and object coordinates until an adequate fit is obtained. The objective of the iterative process is to obtain a spatial representation in *p* dimensions, such that the Euclidean distances among the objects are monotonically related to the original dissimilarities.

The iterative process comprises four steps:

1. Step 1 (Initial phase) selects the dimensions (p) and determines the initial configuration and the resulting distances.
2. Step 2 (Non-metric phase) uses monotone regression to relate the to d_{ij}. The estimated regression produces a new set of dissimilarities, called disparities that are monotonically related to the d_{ij}.
3. Step 3 (Metric phase) revises the spatial configuration to obtain new distances, which are more closely related to the disparities generated in step 2.
4. Step 4 (Evaluation phase) determines the goodness of fit of the distances and the disparities

If the fit is not adequate, Steps 2 and 3 are repeated.

Example of Quantifying Teachers' Reflections

In the example following, a Euclidean distance model was used which produces a projection similar in form and function to principal components analysis (PCA); primary teachers were coded p1 – p11, and secondary teachers were coded s1 – s10, and recent graduates grad1 – 20. MDS was carried out using the ASCAL protocol in PASW/SPSS v21.0, which uses the Takane-Young-de Leeuw S-stress – formula 1 - Equation 3, which is a variation of Kruskal's stress function.

$$SSTRESS(1) = S = \left[\frac{1}{m}\sum_{k=1}^{m}\left[\frac{\sum_i\sum_j\left(d_{ijk}^{2}-d_{ijk}^{*2}\right)2}{\sum_i\sum_j d_{ijk}^{*4}}\right]\right]^{\frac{1}{2}}$$

Equation 3.

Kruskal's stress function, and its variants, Kruskal (1964) is the most commonly used measure in determining a model's goodness of fit.

Reliability

Although there is no strict rule regarding how much stress is tolerable, and thus rendering the procedure reliable, the rule of thumb is that a value ≤0.1 is excellent and anything ≥0.15 is not tolerable (Kruskal & Wish, 1978). In the following example, 3 iterations were conducted which were stopped when the improvement between iterations was less than 0.001; the S-stress in this example was reduced from 0.13247 to 0.10282 - both within the critical limit adopted by the researcher.

Results of Research

As previously stated, the evaluations were ranked according to the rubric presented in Table 1. and the Appendix. The ranks were tabulated with case ranks forming the rows and question ranks forming the columns, thus a matrix was produced for each lesson. The descriptive statistics are represented in Table 2. When the mean value is '3', there is a standard deviation of .0000 and this is the 'IDEAL' value. It also happens to the values for 'grad1'. When the scores are represented in a MDS plot, Figure 1. In this task, noting the position of the 'ideal', seven of the nine primary teachers appear in the same half of the plot as the 'ideal'. The graduate teachers had a much broader spread of locations and a small cluster of grad1, grad2, grad10, and grad13 appeared in the extreme right-hand side appearing to be 'opposite' in their perceptions

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Volume 67, 2015

of the 'ideal'. The plot therefore is a visual representation of the perceptual 'distance' between people and this is essentially its use. MDS helps the research identify clusters or groupings of people based on their perceptions and extremes such as outliers are easily identified. Once outliers are identified, the researcher refers back to the original data to ascertain why such divergence did become evident, or in this case, which question or questions the teachers concerned differ from the ideal.

Table 2. Descriptive statistics of the dataset.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
grad1	4	1.00	1.00	1.0000	.00000
grad2	4	1.00	2.00	1.5000	.57735
grad3	4	1.00	3.00	1.7500	.95743
grad4	4	1.00	3.00	2.0000	.81650
grad5	4	1.00	3.00	2.0000	.81650
grad6	4	1.00	2.00	1.5000	.57735
grad7	4	1.00	2.00	1.7500	.50000
grad8	4	2.00	3.00	2.2500	.50000
grad9	4	1.00	3.00	2.0000	.81650
grad10	4	1.00	2.00	1.5000	.57735
grad11	4	1.00	3.00	2.0000	.81650
grad12	4	2.00	3.00	2.2500	.50000
grad13	4	1.00	2.00	1.2500	.50000
grad14	4	2.00	3.00	2.5000	.57735
grad15	4	1.00	3.00	1.7500	.95743
grad16	4	1.00	3.00	2.0000	.81650
grad17	4	3.00	3.00	3.0000	.00000
p8	4	1.00	3.00	2.5000	1.00000
p5	4	1.00	3.00	1.7500	.95743
p3	4	2.00	3.00	2.7500	.50000
p7	4	3.00	3.00	3.0000	.00000
p6	4	1.00	2.00	1.7500	.50000
p1	4	2.00	3.00	2.5000	.57735
p10	4	2.00	3.00	2.7500	.50000
p9	4	1.00	3.00	2.0000	.81650
s5	4	1.00	3.00	2.5000	1.00000
s6	4	2.00	2.00	2.0000	.00000
s7	4	1.00	3.00	2.5000	1.00000
s8	4	1.00	3.00	1.7500	.95743
IDEAL	4	3.00	3.00	3.0000	.00000
Valid N (listwise)	4				

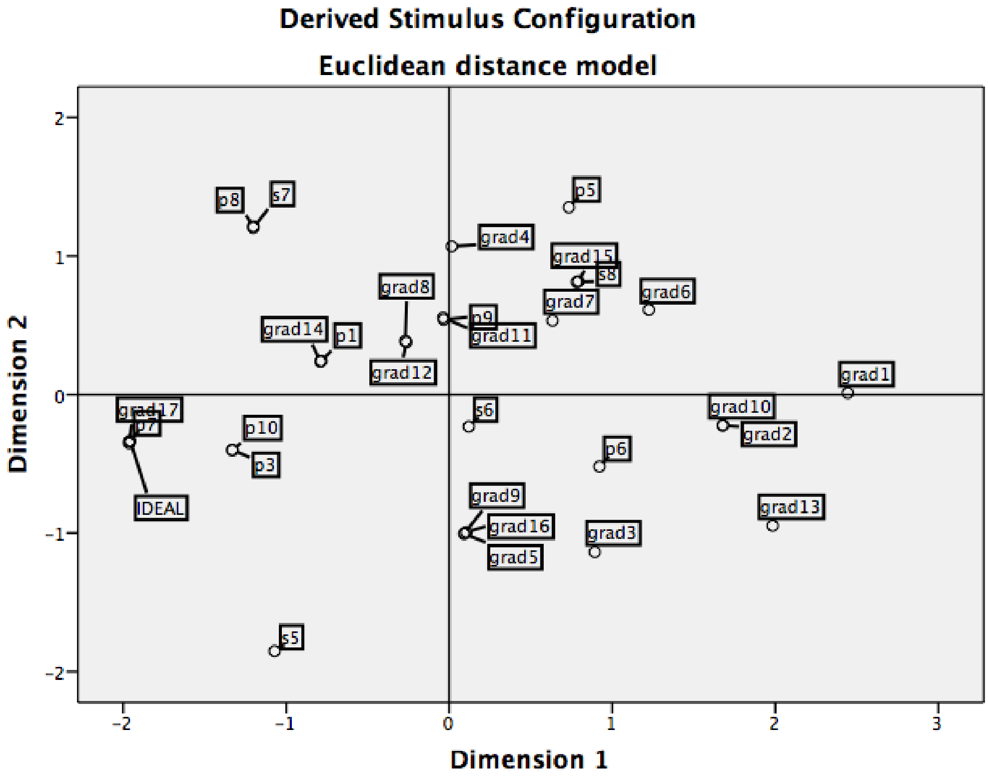

Figure 1: Output plot of the teachers' dimensions and 'ideal' based on their initial reflections.

Discussion

MDS provides a useful visualization tool based on the principle of data reduction. The calculations involve in the case of non-metric MDS of calculating distances between people and then the calculation of dimensions subject to SSTRESS iterations. Interesting questions in ongoing work will provide fruitful sources of further research. As we have seen, an hypothetical 'ideal' can be included in the data to see if there has been 'movement' towards, the 'ideal' or not - the question of effect of the ideal on the data is relevant - a case of uncertainty principle? When 'gathering' plot points into clusters where to make the 'cut' with respect to proximity to an ideal is relevant. Furthermore, two time captures as in a pre-/post-test scenario can be included in the same plot which permits comparison between 'how I was' and 'how I am' and 'movement' in the structure of perceptions discerned as used by (Ryan, 2014). Three-dimensional plots are possible, which, given the possible role in analysing large numbers (n>30), suitable software in viewing such plots are desirable. A simple practical detail is that a plot on paper of 30+ individuals becomes problematic to interpret. Further work is needed in order to align MDS with other methods such as that done by (Li & Sireci, 2013) who used MDS to analyze content validity, and (Alt, 2015) who combined MDS with structural equation modelling (SEM) which has the possibility of transforming MDS into a deeper analytical tool beyond visualisation and assessment.

Conclusion

The use of MDS for visualizing and assessing teachers' development in an in-service programme is a viable research method and it opens a wide variety of applications in primary, secondary and tertiary education including teacher education.

Any perceptions or responses to stimuli if they are ameliorable to assigning an integer in order to rank, score or place on a scale over a set of questions for a group of individuals can have MDS applied to them. The resultant plot provides a direct visual representation of the original integers without loss of 'richness' while 'reducing' the data to two 'dimensions'.

The MDS plots may be used to provide an overall visualisation of the data and when an ideal is included in the data, assessment can become a feature of the procedure. Combination of MDS with SEM can provide a powerful analytical method which go beyond traditional methods such as factor analysis.

References

Alt, D. (2015). Using structural equation modeling and multidimensional scaling to assess students' perceptions of the learning environment and justice experiences. *International Journal of Educational Research, 69*, 38-49. doi: http://dx.doi.org/10.1016/j.ijer.2014.10.001.

Borgatti, S. P. (1997). Multidimensional scaling. Retrieved from http://www.analytictech.com/borgatti/mds.htm.

Cízek, P., Härdle, W., & Weron, R. (2005). Multidimensional scaling. Retrieved from http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/tutorials/mvahtmlframe131.html.

de Leeuw, J. (2000). Multidimensional scaling. *Department of Statistics Papers*. Retrieved from eScholarship - University of California website: http://escholarship.org/uc/item/83p9h12v

de Leeuw, J., & Heiser, W. (1980). Theory of multidimensional scaling. In P. Krishnaiah (Ed.), *Handbook of Statistics* (Vol. II). Amsterdam: North Holland Publishing Company.

Ding, C. (2012). Studying children's early literacy development: Confirmatory multidimensional scaling growth modeling. *International Journal of Educational Research, 53*, 278-288. doi:http://dx.doi.org/10.1016/j.ijer.2012.04.002

Ding, C. (2015). Evaluating change in behavioural preferences: Multidimensional scaling single-ideal point model. *Measurement and Evaluation in Counceling and Development, 18*, 165-185.

Ding, C., & Yang, D. (2012). Assessment of grade-level differences in coping behaviour among adolescents using multidimensional scaling single-ideal-point model. *Measurement and Evaluation in Counceling and Development, 46* (2), 101-113.

Estes, W. K. (1996). *Classification and Cognition* (Vol. 22). Oxford: Oxford University Press.

Fahrmeir, L., & Hamerle, A. (1984). *Multivariate statistische verfahren*. Berlin: De Gruyter.

Fielding, A. H. (2007). *Cluster and Classification techniques for the biosciences*. Cambridge: Cambridge University Press.

Hahn, U., & Chater, N. (1997). Concepts and similarity. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (1st Ed., Vol. 1, pp. 43-92). London: Psychology Press.

Jaworska, N., & Chupetlovska-Anastasova, A. (2009). A review of multidimensional scaling (MDS) and its utility in various psychological domains. *Tutorials in Quantitative Methods for Psychology, 5* (1), 1-10.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika, 29*, 1-28.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling* (Vol. 07-011). Newbury Park, CA.: Sage Publications.

Li, X., & Sireci, S. G. (2013). A new method for analyzing content validity data using multidimensional scaling. *Educational and Psychological Measurement*. doi:10.1177/0013164412473825.

Medin, D. L. (1975). A theory of context in discrimination learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 263-314). New York: Academic Press.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Ryan, S. (2014). *Transforming Science in the Irish Primary School – CASE and Metacognition.* (PhD), Dublin City University, Dublin.

Thorndike, E. L. (1913). *Educational psychology* (Vol. II). New York: Teachers' College, Columbia University.

UNESCO (n.d., 30th September 2008). Non-metric multidimensional scaling. Retrieved from http://www.unesco.org/webworld/idams/advguide/Chapt8_1.htm.

Vitale, M. R., Williams, P. B., Kocsis, J. J., & Kosinski, S. (2015). Using multidimensional scaling and unfolding to explore teacher perceptual biases for selected special education referral categories. *Journal of Experimental Education, 51* (3), 134-141.

Appendix. Complete Rubric.

Question	Descriptor	Type of reflection	Rank
Comment on the dialogue itself	The dialogue provided examples of questions that can be used again. A general comment on its ease of use. The teacher misunderstood the question or the function of the dialogue.	Reflection for action	3
	The dialogue provided the opportunity to discover how much the children knew or the teacher expresses surprise by the quality of the dialogue. A general evaluation of the dialogue	Reflection-on-action	2
	The teacher intends to use dialogue as a learning tool in the future or describes fully why the dialogue is worthwhile.	Reflection-in-action	1
Comment on recommendations regarding methodologies	General dismissal / disengagement of methodologies. Methodology viewed in a very general way.	Evaluation	3
	Some engagement with the description of methodologies / 'enjoyment' of specific methodologies.	Reflection-on-action	2
	The teacher attempts to weave-in the methodologies into their practice / state why a certain methodology was 'good'.	Reflection-in-action	1
What seemed to go well and what seemed to go badly?	Teacher focuses on the children's discipline or whether the practical worked / mechanics of the lesson.	Not applicable	3
	The teacher focuses on the children's motivation / enjoyment levels. Anecdotes of who did what.	Not applicable	2
	The teacher focuses on the children's (group – but not individuals) positive quality of dialogue or interaction. Lesson led to discovery / exploration by the students	Not applicable	1
Comment on how your own teaching experience fitted-in to the lesson plans.	Nothing to offer, no effect. Misunderstood the question.	Not applicable	3

Advised by Vincentas Lamanauskas, University of Siauliai, Lithuania

Received: *October 02, 2015* Accepted: *October 25, 2015*

Thomas McCloughlin PhD, Lecturer, Dublin City University-St. Patrick's Campus, Drumcondra, D09 DY00 Dublin, Ireland.
E-mail: tom.mccloughlin@dcu.ie
Website: http://www.spd.dcu.ie/site/education/staff_details/mccloughlin_tom.shtml