

Applying the Structure of the Observed Learning Outcomes (SOLO) Taxonomy on Student's Learning Outcomes: an empirical study

CHARLES C. CHAN, M. S. TSUI & MANDY Y. C. CHAN, *Department of Applied Social Sciences, Hong Kong Polytechnic University, Hong Kong*

JOE H. HONG, *Education Office, City University of Hong Kong, Hong Kong*

ABSTRACT *This article explores the application of different educational taxonomies in measuring students' cognitive learning outcomes. The objectives were to compare three educational taxonomies—namely, the Structure of the Observed Learning Outcomes (SOLO) taxonomy, Bloom's taxonomy and reflective thinking measurement model—and to test the application value of these taxonomies. A comparative literature review was conducted to provide an underlying conceptual framework. Recommendations from this review were examined experimentally. Scripts of long essay papers and short classroom discussion responses were analyzed by the modified versions of the taxonomies. It was found that SOLO is suitable for measuring different kinds of learning outcomes. However, finer categorization of SOLO levels did not eradicate the problem of SOLO's conceptual ambiguity. It is suggested that the next step of research should be on setting up panels of judges to find out which taxonomy is suitable for measuring what learning outcomes under which contexts.*

Introduction

From the literature on the SOLO taxonomy, it was suggested that SOLO is a hierarchical model that is suitable for measuring learning outcomes of different subjects, levels and for all lengths of assignments (Biggs & Collis, 1982; Chan *et al.*, 2001). Several researchers who have applied SOLO credit its comprehensiveness in application, and its objective criteria provided for measuring students' cognitive attainment (Chick, 1998; Lake, 1999; Van Rossum & Schenk, 1984). They did try to make use of SOLO's

advantages and applied it to different subjects such as mathematics, biology and language studies (Chick, 1998; Campbell *et al.*, 1998; Lake, 1999). It was found that the relationships between SOLO and other learning factors such as learning motivations, writing styles (language), learning strategies, year of study, prior academic abilities, etc., were positive (reviewed and summarized in Chan *et al.*, 2001). At the same time, researchers and reviewers pointed out the loopholes of SOLO caused by its structure and the criteria for categorization (Chan *et al.*, 2001; Chick, 1998). The conceptual ambiguity of SOLO's structure makes categorization unstable along with the problem of low inter-rater reliability and the accident of mistaking an "extended abstract" (very good answer) as "prestructural" (very poor answer) (Chan *et al.*, 2001).

Most past reviews gave a general picture of SOLO and its application in assessing student's cognitive learning outcomes (Biggs & Collis, 1989; Boulton-Lewis, 1992, 1995; Chan *et al.*, 2001; Hattie & Purdie, 1998). A few reviews compared SOLO with other taxonomies but did not agree upon which taxonomy was better (Chan *et al.*, 2001; Hattie & Purdie, 1998). In order to address these current confusions, an empirical approach is used in this study. This article is a report of the experiments conducted for testing mainly the hypotheses suggested in the comparative review by Chan *et al.* (2001). For example, it was suggested that more sub-levels should be added into the SOLO categorization to enhance inter-rater reliability. The ultimate goal is to find out the optimal applications of educational taxonomies in higher education.

Hypotheses

H1: Adding sub-levels or sub-categories to the SOLO taxonomy will increase inter-rater reliability.

H2: SOLO is suitable for measuring all kinds of cognitive learning outcomes (different assignment lengths, subject areas, levels of study).

H3: Cross-validation with different educational taxonomies will make assessment of students' learning outcomes more reliable.

H4: Reflective thinking model performs better in measuring student's critical thinking levels than the other two taxonomies.

H5: The SOLO taxonomy compares more favourably among the three educational taxonomies discussed here (SOLO, Bloom's taxonomy and Reflective thinking model).

Methodology

Study 1

Subjects: The subjects were 17 postgraduate students who have taken an advanced practice subject in mental health during the Spring semester (February to May) in 1999. All of them were adult learners (registered social workers) who have enrolled in the part-time MA social work programme. Their ages ranged from 27–35 and the mean age was 30.85 years.

Procedure: As required by the subject, students had to hand in a term paper 10–15 pages in length on a topic they had chosen at the beginning of the semester. In that paper, they were required to present a case they have worked on and describe how they used the therapies taught in the subject to help their clients and resolve problems. Students were required to show that they had a basic knowledge of cognitive behavioral intervention

and the ability to apply the knowledge learnt (theories and therapies) in the real world situations. Evaluations and reflections on the application of therapies were encouraged. These reports were collected at the end of the semester (May 1999) and marked by two trained raters, using the three educational taxonomies mentioned above simultaneously. Inter-rater reliability was calculated for further analysis.

Instruments: Three educational taxonomies were used to measure students' learning outcomes as reflected in their term papers. First, the SOLO taxonomy was designed in accordance with Piaget's different cognitive stages of development by John Biggs in 1982 for the purpose of classifying students' responses. There were five categories: Prestructural (lowest level), Unistructural, Multistructural, Relational and Extended Abstract (highest level) (see Appendix A.) It is a hierarchical model of increasing structural complexity: increasing consistency, increasing number of organizing dimensions and increasing use of relating principles (Hattie & Purdie, 1998). Burnett (1999) and Trigwell & Prosser (1991) have added sub-levels to the SOLO taxonomy and used this new scaling to measure clients' response. This made the categorization more precise. This idea was modified and examined in this study. Three sub-levels were added to the Multistructural and Relational categories respectively, which resulted in a total of nine SOLO levels (see Appendix B).

Second, the cognitive aspect of the Bloom's taxonomy consists of six hierarchical categories: namely, knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom *et al.*, 1956). "Knowledge" is recollection or remembrance of previously learned materials. "Comprehension" is the ability to understand and interpret facts and principles. "Application" refers to the ability to apply learned materials in new and concrete situations. "Analysis" is the ability to break down materials into its component parts so that the relationship between parts can be understood. "Synthesis" brings together parts or elements to form a whole. "Evaluation" concerns the ability to judge the value of materials learned, including the ability to indicate logical fallacies in arguments (Hosseini, 1993) (see Appendix C for details).

Third, a reflective thinking instrument was developed in the 1990s to assess the level of reflection and critical thinking in students' written assignments (Kember *et al.*, 1999). This taxonomy divides students' written work into two categories: non-reflective and reflective. Within the non-reflective category, there are two sub-levels: habitual action (lowest level) and thoughtful action. "Habitual action" is performed automatically with little conscious thought. "Thoughtful action" makes use of existing knowledge, which is similar to the "application" category in Bloom's taxonomy. There are also two levels under the category of "reflective thinking": reflection and critical reflection, which is of a higher level. Kember followed Mezirow's work (1991) and subdivided "reflection" into content and process reflection. "Content reflection" is concerned with what we perceive, while "process reflection" concerns the methods or manner we used to perceive (see Appendix D for details).

Study 2

Subjects: The subjects were 11 postgraduate students who, during the Autumn semester (September to December 2000) took another postgraduate subject in mental health which has an emphasis on theories and concepts more than practice methods. They were also adult learners enrolled in the part-time MA social work programme, but the sample was

not the same as that of Study 1. All students were female and the age range was 26–40 years, with the mean of 31.50 years.

Procedure: A case study problem (3 sub-questions) was presented to students during the lecture on 24 October 2000, with another being presented one week later. There were five possible solutions for each problem presented (with a total of five questions). Students were required to choose the appropriate answer(s) and explained why they choose that answer(s) after discussion with group members. There was no limit on the length of responses, but students were advised to think carefully and considered comments from group members when giving written responses. Scripts of short responses (ranging from one sentence to around 120 words) were collected after lectures and analyzed with the educational taxonomies mentioned above (SOLO, Bloom's taxonomy and reflection model) separately by two trained raters. Inter-rater reliability was calculated for further analysis.

Instruments: The SOLO taxonomy, Bloom's taxonomy and Kemper's reflective model. (Note: the original five-level SOLO model was used here, as the responses were too short and did not need sub-categorization. For details, please refer to the previous section on Study 1.)

Results

Study 1

The average SOLO level was 4.76 (1 = prestructural ... 8 = extended abstracts, see Appendix B), which means students were between the multistructural/high and relational/low levels (SOLO level by rater 1 and 2 were 4.12 and 5.14, respectively.) The average reflective thinking level was 3.29 (1 = habitual action ... 5 = critical reflection, see Appendix D), which means students in general were reflective (they have reached the content reflection level). The average Bloom's level was 5.24 (1 = Knowledge, 2 = Comprehension, 3 = Application, 4 = Analysis, 5 = Synthesis and 6 = Evaluation, see Appendix C). Most students reached higher levels in the Bloom's hierarchy: 82.4% reached above "analysis" level. The inter-rater reliability was also quite high for all three taxonomies ($r=0.60$ for SOLO, $p<0.05$; $r=0.93$ for Bloom's and $r=0.87$ for reflective thinking model, $p<0.01$) (see Appendix G for details.)

All three taxonomies correlated with each other ($r=0.75$ for SOLO and reflective thinking model, $r=0.74$ for SOLO and Bloom's, $r=0.84$ for Bloom's and reflective thinking model; $p<0.01$).

Study 2

In Study 2, students' learning outcomes are being measured from short responses, which was not the same as the long essay papers in Study 1. Their responses ranged from one sentence to around 120 words, though most consisted of several sentences (around 40 words). The mean SOLO level of the first rater was 2.62 (1 = prestructural, 2 = unistructural, 3 = multistructural, 4 = relational, 5 = extended abstract; see Appendix A for details). The mean SOLO of second rater was 2.87. The grand average mean was 2.74, which means students in general were judged to be approaching the multistructural level of SOLO. The inter-rater reliability was 0.66, $p<0.05$.

The average Bloom's level was 3.18, which means most students reached the "Application" level (see Appendixes C & G for details). The inter-rater reliability was 0.68, $p < 0.05$. For students' reflective thinking level, the mean averages by the first and second raters were 2 and 2.72 respectively (1 = habitual action, 2 = thoughtful action, 3 = content reflection, 4 = process reflection, 5 = critical reflection; see Appendix D for details.) The total average mean was 2.35. However, the inter-rater reliability for Kember's reflective thinking model was not significant.

The SOLO and reflective thinking model taxonomies correlated with each other ($r = 0.9$, $p < 0.05$) though the Bloom's taxonomy did not correlate with either of these taxonomies (see Appendix G for details).

Discussion

SOLO is claimed to be applicable in measuring cognitive learning outcomes in different kinds of subjects among different levels of students and on different types of assignments (Biggs & Collis, 1982; Hattie & Purdie, 1998). This study has found evidence to support such claims. SOLO did work for different subjects as students of Study 1 and Study 2 were taking practice oriented versus conceptual oriented studies. Content variations did not seem to reduce SOLO's functions. Also, it was found that SOLO could be applied to students of different levels as seen in the literature. The students in this study were postgraduates. Biggs & Collis (1982) have conducted studies among primary students while Chick (1998) has tried to use SOLO among undergraduates. Still others (Boulton-Lewis *et al.*, 1996) applied SOLO to teachers. For measuring written assignments, SOLO could be applied to different lengths of assignments. This study has shown that long essays allowed for the application of the modified SOLO version with sub-levels where raters could have a better focus on students' work. Hypothesis 2 is thus accepted based on the fact that SOLO can measure different kinds of cognitive learning outcomes.

It seems that the inclusion of sub-levels did help to reduce ambiguity. The inter-rater reliability for Study 2 (the one which applied the modified version of SOLO with sub-levels) was higher than that of Study 1. There was a clear difference in the distribution of students' SOLO level when comparing Study 1 and 2. In Study 1, the distribution was more even—for example, 5.9% of students reached the prestructural level, 5.9% reached the unistructural level, 5.9% reached the multistructural/moderate level, 17.6% reached the multistructural/high level, 29.4% reached the relational/moderate level, 11.8% reached the relational/high level and 23.5% reached the extended abstract level (marked by rater 2). In Study 2, 55.6% of students reached the prestructural level, while 44.4% reached the unistructural level for question 1 of Study 1. In addition, 9.1% of students reached the unistructural level, 45.5% reached the multistructural level and 45.5% reached the relational level for Study 2 (also marked by rater 2). Students' cognitive attainments concentrated in one or two levels (see Table 1), which made it difficult to distinguish the different cognitive levels of students.

Though some argue that ambiguity is caused by the short length of assignments (responses), the followings are some examples extracted from the short statements of students (in Study 2), showing that the ambiguity of measurement among short responses can be improved by adding sub-levels to the original SOLO scale. The examples were responses within Study 2, all getting the same SOLO level (see Appendix E for the sample of questions):

TABLE 1. Student's SOLO levels in Studies 1 and 2

SOLO levels (modified version)	Study 1 (%)	Study 2 (%)		SOLO levels
		Q1	Q2	
Prestructural	5.9	55.6	0	Prestructural
Unistructural	5.9	44.4	9.1	Unistructural
Multistructural/low	0	0	45.5	Multistructural
Multistructural/moderate	0			
Multistructural/high	17.6			
Relational/low	0	0	45.5	Relational
Relational/moderate	29.4			
Relational/high	11.8			
Extended abstract	23.5	0	0	Extended abstract

I think that option 2 is a better answer ... because support group can share experiences which is much more helpful. Other options are also good and helpful for the family as they offer help to the family on how to face difficulties. But option 2 is better as during the session of support group, information suggested in other options could be delivered and shared with the family concerned. (Student A)

I think option 1 is a better choice, as the family concerned may want to know more, assuming that they are not well educated and lack information on the illness that the child is suffering. Preparing them with more medical information is helpful and can act as the first step to building treatment relationship in the future. (Student B)

I think option 2 is good, as peer influence is very important. It helps to build trust and positive attitudes of clients towards the support group and later towards the whole treatment package. (Student C)

Although all the above responses got the same SOLO level (multistructural), there were actually differences in their content and presentations. The response given by Student C only listed two relevant points without further explanations, while Student B gave several reasons with the case context (client background) considered. Student A not only stated the reason for choosing option 2; she also compared her favoured option with other options and explained the advantage of her chosen option against the others. If there were sub-levels in the scale, these three responses would be divided into three different levels: low (Student C), moderate (Student B) and high (Student A), and this would give a fairer reflection of students' cognitive attainment levels.

With the above evidence, Hypothesis 1 is accepted (i.e., adding sub-levels to the original SOLO scale will reduce ambiguity and increase assessment reliability). Also, if the other taxonomies such as the Bloom's taxonomy were used as cross-validation measures, there could be a fairer reflection of students' learning outcomes. For example, the response given by Student C seemed to be taken from personal experience and textbook theories, so it would be classified as the "Knowledge" level in Bloom's taxonomy. The response given by Student B showed that client's background was taken into account and was a case of careful application of theories learnt, which would be classified as reaching the level of "Application". The response given by Student A was a clear case of analyzing through comparison, which would be classified as "Analysis".

Students' responses would be classified into categories appropriately described students' cognitive abilities if Bloom's taxonomy were also used here. Thus, Hypothesis 3 is accepted as cross-validation measures can improve assessment accuracy.

However, in some cases, the problem of ambiguity persists even with the implementation of the sub-level system in the SOLO taxonomy. The comparative review (Chan *et al.*, 2001) conducted earlier suggested that there were possibilities to misjudge low SOLO level as high cognitive attainment or vice versa. There were cases where raters gave very different marks to the same student, indicating that there were ambiguities. There follows some extracts from students' essays (Study 1) and their marks given by raters (see Appendix F for the essay topics).

Under the cognitive-behavioral approach, the client's problem was the result of interactions between his distorted thoughts, dysfunctional behaviors and his unstable mood ... Cognitive restructuring was one of the strategies used. Worker collaborated with the client to identify those distorted thoughts ... Worker helped the client to recognize his own efforts and successful experience. Cognitive restructuring was effective ... though it was not the only strategy ... Emotional skills were also taught to the client ... facilitating client's self-exploration of his own feelings. Worker also tried to expand client's ways to handle critical situations ... Crisis intervention and problem-solving strategy were adopted ... Worker can use different techniques in different cases to change behaviors and emotional states. (Student D: rater 1 considered it to be multistructural/low level, but rater 2 considered it to be relational/high level.)

Client Mrs. A's emotional problem stems from her husband's over suspicious attitude ... besides her baby is suffering from autism and mental retardation ... she feels despair and disappointed to her marriage and family. She dares not to turn her emotion to her husband as she still loves him ... We grow up in a patriarchy society ... men hold superior power ... men are expected to direct family and shoulder financial burden according to sex-role stereotyping. Women were taught ... showing anger and rage is unfeminine and undesirable ... learned to turn emotion inward. Mrs. A ... pressure from her husband ... external oppression ... feel inferior internally for not fulfilling societal expectation to give birth of a healthy baby. Her self-identity is shaking ... despair, anger and frustration turns inward ... depression is the consequence. (Student E: rater 1 considered it to be multistructural/low level, while rater 2 considered it to be relational/moderate level.)

The above extracts showed that different raters did have different interpretations on what is a listing of ideas and what is generalization of ideas. Though criteria were set for each SOLO level (listing of ideas is multistructural, while generalization and integration of ideas is relational) and raters agreed with each other most of the time, there were still cases where large discrepancies occurred and raised the problem of accuracy. It is not fair to the student concerned if raters average out the ratings as it would put the student on a middle level which may not represent his/her actual cognitive attainment. The same case applies if raters take either the highest or lowest ratings as the final result.

The above discussion reveals the fact that there were still some ambiguities and application problems when applying SOLO, even where modifications were applied. At the same time, other taxonomies (such as the Bloom's) have a clearer division of categories, which are mutually exclusive (Chan *et al.*, 2001). Furthermore, it was found

that the reflective thinking model was more suitable for assessing students' ability in critical thinking and evaluation. It has a special function in terms of assessing student's learning outcomes. Thus, we were not able to decide whether SOLO is the best of all. Hypothesis 5 is therefore rejected.

For the remaining Hypothesis 4, the reflective thinking model could divide students' ability in critical thinking into evaluating the content of knowledge learnt and the process of the application of this knowledge, and even questioning the basic propositions of the discipline concerned. For example,

... I find that building a rapport and collaborative relationship with clients is of paramount important ... In this cognitive approach, I do not find it highlights the importance of inviting the clients' family to be involved in the therapeutic intervention. This approach is directive and worker has to be aware of avoiding dispensing quick-cure methods ... providing solutions too quickly ... could not enhance clients' coping ability but increased his/her dependence when facing difficulties. (An extract of the self-reflection part by Student F: rater 2 considered it to be critical reflection and belonging to the "evaluation" category of Bloom's taxonomy.)

The above example shows that the reflective model measured more specifically students' evaluative abilities as it could identify the extract as having reflected on both the content and the process of the knowledge learnt, and the student did question some aspects of the therapy learnt and suggested improvement. The rating given by Bloom's taxonomy could only show that the student had high evaluation ability, which could not be as specific as Kember's model. Thus, Hypothesis 4 is accepted with the support of the above evidence.

Besides being specific in measuring evaluative ability, the reflective thinking model can be used as a guide for both instructors and students for further improvements. Instructors could re-design the curriculum and class activities according to students' level of critical and reflective thinking in order to enhance their evaluative ability.

In summary, it is recommended that the reflective thinking model be used specifically for measuring students' critical and evaluative thinking abilities. Also, it has been shown that adding sub-levels to SOLO would reduce ambiguity and increase agreement between raters (inter-rater reliability) when applying it. Moreover, SOLO was shown to be useful in measuring cognitive attainment of students of different classes, subjects, levels and with different assignment requirements. However, further studies will have to be conducted to see if it is better to use the new version of Bloom's taxonomy (two-dimensional and not cumulative hierarchical) (Anderson & Krathwohl, 2001) when instructors and researchers have to assess a variety of students' cognitive abilities and to test the use of SOLO on assignments of other formats such as graphics, calculations, etc.

Conclusion

Looking at the findings of this study, the three educational taxonomies seemed to be closely related to each other, each could complement the weaknesses of the others. The hypotheses set in the previous comparative review (Chan *et al.*, 2001) were tested out in this study and it was suggested that the problem of conceptual ambiguity of SOLO could be improved but not totally resolved by adding sub-levels to the scale. Further studies could be employed to test out whether adding a third rater or improving raters' experience would help to improve agreement among raters. In addition, further studies

could explore whether using the new version of Bloom's taxonomy could improve the accuracy of assessing cognitive learning outcomes.

Acknowledgments

This study is under the auspices of the "Transforming Learning in the Quality Reform—A Model of Lifelong Learning in the Human Services" project. It is a part of the "Tackling the Challenges to Professional Education Arising from the Pace of Technological Development and the Need for Lifelong Learning" project, which is funded by the University Grants Committee, Hong Kong Special Administrative Region, People's Republic of China.

Notes on Contributors

CHARLES C. CHAN is an Associate Professor at the Hong Kong Polytechnic University with teaching responsibility in social and clinical psychology subjects. He is the president of Hong Kong Professional Counseling Association and is also advisor to a number of social service organizations and schools in Hong Kong. His major specialty areas focus on outcome evaluation of health, education and social care programmes. *Correspondence:* Dr Charles C. Chan, Department of Applied Social Sciences, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. E-mail: ssshchan@polyu.edu.hk.

MING-SUM TSUI is Senior Lecturer, Department of Applied Social Sciences, the Hong Kong Polytechnic University. His research interests include supervision, mentorship, human service management and substance abuse.

MANDY YUEN-CHIN CHAN is a Research Assistant at the Department of Applied Social Sciences, the Hong Kong Polytechnic University. Her research interests include computer-mediated communication, educational information policy and outcome evaluation in higher education.

JOE H. HONG is a Senior Education Development Officer at the Education Office at the City University of Hong Kong. He has developed expertise in research methods, learning and teaching strategies, and evaluation of the use of technology in learning and teaching in higher education.

REFERENCES

- ANDERSON, L. & KRATHWOHL, D. R. (Eds) (2001) *A Taxonomy for Learning, Teaching and Assessing* (New York, Addison Wesley Longman, Inc.).
- BIGGS, J. B. (1999) Formulating and clarifying curriculum objectives, in: J. B. BIGGS *Teaching for Quality Learning at University* (Buckingham/Philadelphia, PA, Open University Press/Society for Research into Higher Education).
- BIGGS, J. B. & COLLIS, K. F. (1982) *Evaluating the Quality of Learning: the SOLO taxonomy* (New York, Academic Press).
- BIGGS, J. B. & COLLIS, K. F. (1989) Towards a model of school-based curriculum development and assessment using the SOLO taxonomy, *Australian Journal of Education*, 33 (2), pp. 151–163.
- BLOOM, B. S., ENGELHART, M. D., FURST, E. J., HILL, W. H. & KRATHWOHL, D. (1956) *Taxonomy of educational objectives: the cognitive domain* (New York, David McKay Co.).

- BOULTON-LEWIS, G. M. (1992) The SOLO taxonomy and levels of knowledge of learning, *Research and Development in Higher Education*, 15, pp. 482–489.
- BOULTON-LEWIS, G. M. (1995) The SOLO taxonomy as a means of shaping and assessing learning in higher education, *Higher Education Research and Development*, 14 (2), pp. 143–154.
- BOULTON-LEWIS, G. M., WILSS, L. & MUTCH, S. (1996) Teachers as adult learners: their knowledge of their own learning and implications for teaching, *Higher Education*, 32, pp. 89–108.
- BURNETT, P. C. (1999) Assessing the structure of learning outcomes from counselling using the SOLO taxonomy: an exploratory study, *British Journal of Guidance & Counselling*, 27 (4), pp. 567–581.
- CAMPBELL, J., SMITH, D. & BROOKER R. (1998) From conception to performance: how undergraduate students conceptualise and construct essays, *Higher Education*, 34, pp. 449–469.
- CHAN, C. C., HONG, J. H. & CHAN, M. Y. C. (2001) Applying the Structure of the Observed Learning Outcomes (SOLO) taxonomy on student's learning outcomes: a comparative review (Unpublished manuscript, Hong Kong, Hong Kong Polytechnic University).
- CHICK, H. (1998) Cognition in the formal modes: research mathematics and the SOLO taxonomy, *Mathematics Education Research Journal*, 10 (2), pp. 4–26.
- COURTNEY, T. D. (1986) The significance of the SOLO taxonomy for learning and teaching in geography, *Geographical Education*, 5 (2), pp. 47–50.
- HAMBLEN, K. (1984) An art criticism questioning strategy within the framework of the Bloom's taxonomy, *Studies in Art Education: A Journal of Issues and Research*, 26, pp. 41–50.
- HATTIE, J. & PURDIE, N. (1998) The SOLO model: addressing fundamental measurement issues, in: B. DART & G. M. BOULTON-LEWIS (Eds) *Teaching and Learning in Higher Education* (Camberwell, Australian Council for Educational Research).
- HOSSEINI, J. (1993) Application of Bloom's taxonomy and Piaget's model of cognitive processes to teaching of management information systems concepts, *Journal of Information Systems Education*, 5 (3), pp. 1–12. Retrieved 28 May 2002 from www.gise.org/JISE/Vol1-5/APPLICAT.htm.
- KEMBER, D., JONES, A., LOKE, A., MCKAY, J., SINCLAIR, K., TSE, H., WEBB, C., WONG, F., WONG, M. & YEUNG, E. (1999) Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow, *International Journal of Lifelong Education*, 18 (1), pp. 18–30.
- LAKE, D. (1999) Helping students to go SOLO: teaching critical numeracy in the biological sciences, *Journal of Biological Education*, 33 (4), pp. 191–198.
- MEZIROW, J. (1991) *Transforming Dimensions of Adult Learning* (San Francisco, CA, Jossey-Bass).
- TRIGWELL, K. & PROSSER M. (1991) Relating approaches to study and quality of learning outcomes at the course level, *British Journal of Educational Psychology*, 61, pp. 265–275.
- VAN ROSSUM, E. J. & SCHENK, S. M. (1984) The relationship between learning conception, study strategy and learning outcome, *British Journal of Educational Psychology*, 54, pp. 73–83.

Appendix A. The SOLO Taxonomy

SOLO Level	Coding Description	examples*
Prestructural [1]	The task itself is not attacked in an appropriate way such as using Tautology, or just repeating the question. The student hasn't understood the point	"Teaching is a matter of getting students to approach their learning."
Unistructural [2]	One relevant aspect of task is picked up and there is no relationship between facts or ideas. Understanding is nominal.	"Approaches to learning are of two kinds, surface, which is inappropriate for the task at hand, and deep, which is appropriate. Teachers need to take this into account."

Appendix A.—Continued

Multistructural [3]	Several (two or more) independent aspects are picked up or understood serially, but are not interrelated.	“Approaches to learning are of two kinds, surface, which is inappropriate for the task at hand, and deep, which is appropriate. Students using a surface approach try to fool us into believing that they understand by rote learning and quoting back to us, sometimes in great detail. Students using a deep approach try to get at the underlying meaning of their learning tasks. Teaching is about getting students to learn appropriately, not getting by with short cuts. We should therefore teach for meaning and understanding, which means encouraging them to adopt a deep approach.”
Relational [4]	Relevant aspects are integrated into an overall coherent structure.	“Approaches to learning are of two kinds ... [etc.]. The approaches come about partly because of student characteristics, but also because students react differently to their teaching environment in ways that lead them into surface or deep learning. The teaching environment is a system, a resolution of all the factors, curriculum, assessment, teaching methods and students' own characteristics. If there is imbalance in the environment—for example, a test that allows students to respond in a way that does not do justice to the curriculum, or a classroom climate that scares the hell out of them—the resolution is in favour of a surface approach. What this means is that we should be consistent ...”
Extended abstract [5]	The coherent whole is generalized or re-conceptualized to a higher level of abstraction.	

* Examples are quoted from Biggs's (1999) article and were illustrations answering the following questions “What are approaches to learning? How can knowledge of approaches to learning enhance university teaching?”

Appendix B. Sub-levels of the SOLO Taxonomy (Modified Version)

Sub-levels	Descriptions	Coding
Prestructural	Don't understand question; wild guessing and tautology.	0
Unistructural	Only mention one relevant piece of information.	1
Multistructural/low	Pick up two to three independent aspects that are related to the question asked but without further elaboration.	2
Multistructural/moderate	Pick up quite a number of related information but without further elaboration.	3
Multistructural/high	Pick up many related aspects and elaborate each point with illustrations.	4
Relational/low	Generalize ideas and form paragraphs in one or two parts of the essay.	5
Relational/moderate	Generalize ideas in many different parts of the essay.	6
Relational/high	Overall generalization of major concepts in the whole essay.	7
Extended abstract	Consistently generalize ideas all over the essay, and question or criticize conventional practices and/or underlying principles of the discipline.	8

Appendix C. The Bloom's Taxonomy

Bloom's taxonomy	Description	Examples*
Knowledge	Remembering of previously learned material, including facts, conventions, principles and theories. It represents the lower level of learning outcomes in the cognitive domain.	"Who painted Guernica?"
Comprehension	The ability to grasp the meaning of material. It involves translation and interpretation (explaining and summarizing) of materials.	"Describe the subject matter of Guernica."

Appendix C.—Continued

Application	The ability to use learned material in new and concrete situations. This may include the application of rules, methods, concepts, principles, laws and theories.	“Relate the theme of Guernica to a current event.”
Analysis	The ability to break down materials into its component parts so that the organizational structure may be understood. Skill in comprehending the interrelationships among the ideas and recognize unstated assumptions.	“What compositional principles did Picasso use in Guernica?”
Synthesis	The ability to put parts together to form a new whole. Learning outcomes in this area stress creative behaviors, with major emphasis on the formulation of new patterns or structures.	“Imagine yourself as one of the figures in Guernica and describe your life history.”
Evaluation	The ability to judge the value of material. Judgement criteria may be from internal or external.	“What is your opinion of Picasso's Guernica?”

* Examples were drawn from a series of art questions by Hamben (1984), cited in Hattie & Purdie (1998).

Appendix D. Reflective Thinking Model

Level of reflective thinking/coding	Description	Examples*
Habitual action [1]	Conventional practice, repetition, with little conscious thought	“Appropriate communication skill is that the attitude of the radiographer should be devoted to communicate with the patient without preception or discrimination. [...] The radiographer should pay attention to his voice, speech, intonation, wording of his speech. [...] The radiographer should try to have better language [...] so that the radiographer can communicate with the foreign patients.”
Thoughtful action (or called Understanding) [2]	Apply existing knowledge, e.g., comprehension, application, analysis, and synthesis information.	“Communication and interaction with patients are a type of concern for the patients. [...] I care for them and try my best to fulfil their needs. [...] During leisure times and breaks, I should communicate and talk to patients and therefore I can know what they need and what their feelings are. [...] For some young patients, the communication skill is different from adult patients. [...] I first will make friends with them and treat them as my brothers and sisters. If they are not so frightened after a period of time, the course will be much easier.”

Appendix D.—Continued

Reflection	Content (what we perceive) [3]	“Appropriate communication interaction with patients is very important and essential for the health professionals because patients come to hospital and expect to receive appropriate health service and care [...] Interaction with the patients is important to the patient and the radiographer [...] therefore the patient will be willing to disclose their information [...] Is it necessary to have a radiographic examination? The radiographer can know more about the patient so that the radiographer can modify the technique to the patient’s condition. [...] Therefore it is important to the radiographer to acquire and develop a good communication skill so that they can communicate with the patient without any problem.”
	Process (how we perceive) [4]	
Critical reflection [5]	Critical review of presuppositions of subject disciplines and existing knowledge.	

* Examples were drawn from the materials from the workshop on measuring reflective thinking held by David Kember at Educational Development Centre, the Hong Kong Polytechnic University in November 2000.

Appendix E. Case Study Questions

Case Study 1, sub-question 1

After listening to the case presentation by XX on a client referred for school social work service suspected of eating problem and the details from the initial and subsequent interviews with the worker, what is a suitable diagnosis for the client

1. Anorexia nervosa
2. Bulimia nervosa
3. Major depressive disorder
4. Adjustment disorder
5. Others

Sub-question 2

Recommendation regarding management plan after initial intake:

1. Special attention should be provided by special duty teacher.
2. Exposure and response prevention advice through interpersonal counselling.

Appendix E.—Continued

3. CBI through teaching of self-monitoring and self-regulatory strategies aiming at managing purging and reestablishment of control over eating.
4. Teach the mother to supervise client for at least two hours after every meal.
5. Other management plans.

Sub-question 3

The written report/face-to-face communication with the school principal/parent after six months of intervention should include:

1. Encourage parents/teachers to allow/encourage client to participate in drawing up the meals, exercise, hygiene arrangement and learning schedule.
2. Teach/demonstrate to parents/teachers to respond to the client in a supportive way (e.g. I know you feel fat), and end with a reality-based comment (e.g. but others still see you as very thin).
3. To warn the client that if her weight comes down below a certain level, mother will report to worker and forced hospitalization may be the result.
4. Encourage making friends with normal eating habits.

Case study 2

In the past two decades, we witnessed a shift from child-centered, medically based model to family-centered, health and quality-of-life focussed intervention in the West.

The former intervention typically takes the child as the patient, the professional as the expert and the parents were passive observers. A lot of information and explanation on the illness and related conditions were given to the parents and so expectedly, concerned parents were found explaining, reassuring and even admonishing the child in taking care of their body and the illness conditions.

The more family-centered intervention recognized families as important resources and see families as part of the solutions. They are kept in the forefront of the children's intervention needs with the professional being more of a consultant, seeking a collaborative relationship that support the parents in meeting their goals.

Coming back to the local scene, a 14 year old boy Ah B, youngest of 6 siblings of non-skilled working parents in their 50s, suffers from leukemia.

Ah B has begun to complain about nightmares and loss of appetite. The mother was visibly distress and getting folklore advice.

In preparation for a bone marrow transplant from one of his brothers and the first few of a series of invasive, sometimes painful medical procedures, what would you be doing?

Option 1.Explaining the overall effect of the illness to the family and answering questions concerning the procedures.

Option 2.Helping the family to connect with support groups to model positive attitudes and values.

Option 3.Teaching Ah B various coping strategies such as deep breathing, attention distraction, relaxing imagery, behavioral rehearsal etc.

Option 4.Teaching the parents to use distraction, contingent praise, and active directive to coping (“take a deep breath now”).

Option 5.Teaching the mother, in particular, similar cognitive skills to increase her ability in successfully coping with the traumatic experience of having to see two of her children undergoing an important surgical procedure.

Appendix F. Topics of Students' Term Papers (Whose Extracts being Quoted)

Student D: Application of Cognitive Behavioral Intervention (CBI) on substance abuser case.

Student E: Applying CBI on women having depression.

Student F: Application of CBI on a depressive woman with child care problem.

Note: The term end exercise mentioned here is the final exercise required by an advanced practice subject in mental health.

Appendix G. Descriptive Statistics of Study 1

Constructs	Mean	Standard deviation
SOLO level by rater 1	4.12	1.90
SOLO level by rater 2	5.41	2.43
Average SOLO level	4.7647	1.9374
Reflective level by rater 1	3.29	1.21
Reflective level by rater 2	3.29	1.57
Average reflective level	3.2941	1.3470
Bloom's level by rater 1	1.8529	.6505
Bloom's level by rater 2	1.8039	1.0278
Average Bloom's level	1.8922	.7915

Correlation among educational taxonomies in Study 1

	Average SOLO level	Average reflective thinking level	Average Bloom's level
Average SOLO level	—	.747**	.740**
Average reflective thinking level	.747**	—	.825**
Average Bloom's level	.740**	.825**	—

** $p < 0.01$ (2-tailed)

Inter-rater reliabilities in Study 1

	SOLO level by rater 1	SOLO level by rater 2	Reflective level by rater 1	Reflective level by rater 2	Bloom's level by rater 1	Bloom's level by rater 2
SOLO level by rater 1	—	.599*	.228	.511*	.481	.520*
SOLO level by rater 2	.599**	—	.785**	.901**	.748**	.780**
Reflective level by rater 1	.228	.785**	—	.870**	.743**	.777**
Reflective level by rater 2	.511*	.901**	.870**	—	.742**	.842**
Bloom's level by rater 1	.748**	.793**	.639**	.809**	—	.928**
Bloom's level by rater 2	.642**	.853**	.726**	.947**	.928**	—

$p < 0.05$ (2-tailed).

** $p < 0.01$ (2-tailed).

Descriptive statistics of Study 2

Constructs	Mean	Standard deviation
SOLO level by rater 1	2.6136	.5404
SOLO level by rater 2	3.50	.5730
Average SOLO level	3.46	.5079
Bloom's level by rater 1	3.2882	.5319
Bloom's level by rater 2	3.1736	.7474
Average Bloom's level	3.18	.6349
Reflective level by rater 1	2.50	.3051
Reflective level by rater 2	3.50	.7539
Average reflective level	2.92	.4180

Correlation among educational taxonomies in Study 2

	Average SOLO level	Average Bloom's level	Average reflective level
Average SOLO level	—	.359	.904**
Average Bloom's level	.359	—	.377
Average reflective level	.904**	.377	—

** $p < 0.01$ (2-tailed).

Inter-rater reliabilities in Study 2

	SOLO level by rater 1	SOLO level by rater 2	Bloom's level by rater 1	Bloom's level by rater 2	Reflective level by rater 1	Reflective level by rater 2
SOLO level by rater 1	—	.664*	.432	.452	.522	.706*
SOLO level by rater 2	.664*	—	.350	.131	.083	.877*
Bloom's level by rater 1	.432	.350	—	.682*	.338	.491
Bloom's level by rater 2	.452	.131	.682*	—	.603*	.110
Reflective level by rater 1	.522	.083	.338	.603*	—	.082
Reflective level by rater 2	.706*	.877*	.491	.110	.082	—

* $p < 0.05$ (2-tailed).

Copyright of Assessment & Evaluation in Higher Education is the property of Carfax Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.