JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

https://doi.org/10.33225/jbse/19.18.77

A FEASIBLE GUIDANCE FOR ORDERED MULTIPLE-CHOICE ITEMS IN STUDENTS' HIERARCHICAL UNDERSTANDING LEVELS

King-Dow Su

Abstract. *This research focuses on students' 5 hierarchical levels of Ordered Multiple-Choice (OMC) items for their extensive conceptualized understanding in the particulate nature of matter (PNM) chemistry. The basic framework for OMC items is to link students' conceptual understanding levels with possible cognitive responses. Developed as the substantial learning perspective, OMC item approaches benefit students with dominant formulations in directing a new impetus on individual mental constructions. This research attempts to establish the validity and reliability through assessing students' understanding levels. The main methodology of innovative OMC design manifests students' verified responsive accumulations; with 5 hierarchical mental conceptualizations from naive understanding up to complete systemic PNM understanding. OMC findings offer students more choices to determine the most appropriate corresponding answer towards different conceptualization levels and to set their individual responses at the medium understanding levels. All contribution of this research gives students' future perspectives in more collaborative engagements with further administering OMC items.*

Keywords: *Ordered Multiple-Choice (OMC), particulate nature of matter, hierarchical levels.*

King-Dow Su
Hungkuo Delin University of Technology, Taiwan & Chung Yuan Christian University, Taiwan

Introduction

Students' core conceptual understanding has become an important learning target for constructing scientific basic knowledge in the STEM program. Science educators have worried about how to evaluate students' understanding levels effectively with different alternative choices, by which they could facilitate cognitive performances and improve their learning effectiveness (Bell, 2007). The best encounter of conceptual understanding is in creating an environment of possible choice items via the major pathway to conduct students' different thinking levels (Neumann, Viering, & Boone, 2013). For example, there are some low performing students who rely on rote memorization without target conceptual understanding to comprehend abstract and complicated chemical phenomena (Grove, Hershberger, & Bretz, 2008). They may not develop mental cognition to understand chemical knowledge sufficiently and fail to pass university-level general chemistry exams (Bhattacharyya & Bodner, 2005; Ferguson & Bodner, 2008).

The use of ordered multiple-choice (OMC) was not a readily defined category of conceptualization, nor was its discussion and understanding limited to STEM program aspects. Serving initially as an assessment tool for models of students' cognitive development, the unique feature of OMC items could be linked to discover students' responses of core understanding levels in choice items. Researchers have observed students' understanding levels to build-up individual cognitive skills and construct multiple educational perspectives for effective test items in chemistry learning. Hadenfeldt et al. (2013) developed students' conceptual understanding levels as an approach for logical reasoning instrument in many university chemistry courses. Briggs and Alonzo (2009) integrated OMC items to elicit individual responses to appropriate understanding levels, and to analyze students' developmental insights for more elaborate perspective. From OMC item choices to individual cognitive performance, students' responses of concept understanding levels dominate the most crucial and substantial factors in evaluating students' concept responses, also provide an overview of corresponding particular models to investigate the validity of the OMC results (Briggs et al., 2006; Hadenfeldt et al., 2013; Özmen, 2013).

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Judged from OMC item choices to students' responses of concept understanding levels, instructors could follow a consistent design of high learning assessment to promote students' individual performance cognition. Selected as a starting model design for students' core conceptual understanding, OMC item choices reflect particular multiple-choice items which are different from the traditional multiple-choice items (TMC). The five hierarchical levels based on OMC instrument will require effective choices from which students make alternative information of a scientific concept they hold and retain more advantages on their way toward developing a deeper understanding of the respective target concept (Hadenfeldt et al., 2013). Students quickly and reliably facilitate their responsive options related to one level across a set of OMC items compared to their developments of typical TMC choice items. Thus, OMC item choices with hierarchical levels are featured by a more detailed assessment for students' cognitive coalition for reliable scientific concepts.

The present research of OMC item choices puts emphasis on students' concept understanding levels with regards to the potential basis in PNM chemistry learning design. It has been a new tactic study for university students to follow a complete OMC command in developing their responsive levels with cognitive learning performance (Yakmaci-Guzel, 2013). Since few researches offered satisfactory OMC items in formulating students' different understanding levels, the learning dilemma in need have existed to reassess students' enacting and reconstructing students' mental potentiality in chemistry problem-solving skills (Yakmaci-Guzel, 2013). The coalition of OMC items with PNM chemistry knowledge would inspire students' reconstruction of conceptual understanding levels for more particular scientific engagements (Bunce & Gabel, 2002; Sanger, Phelps, & Fienhold, 2000). In particular, this research sets up a functional assessment instrument for OMC items designed on students' responsive understanding levels of PNM chemistry.

Research Purposes

To fit the above learning requirement, this research developed newly designed OMC items to examine students' conceptual understanding levels in PNM chemistry learning. In response to students' scientific understanding levels of core concepts, this research proposed three fundamental research purposes, as follows:

1) To set up OMC items for students' assessment of conceptual understanding levels
2) To construct validity and reliability of OMC items as an effective diagnostic instrument
3) To evaluate students' PNM domains of understanding levels with OMC diagnostic instrument.

Research Questions

According to the above purposes, three research questions were designed in assessing students' conceptual performance as below:

1) What effective OMC items can be best designed for students' conceptual understanding?
2) To what extent of diagnostic function do OMC items fulfill the most practicable performance of validity and reliability?
3) What substantial PNM domains of students' understanding levels can be accessed with OMC diagnostic instrument?

Research Focus

There were three research focuses for students to go hand in hand with five hierarchical conceptual levels rendered as their target learning. After they have finished extensive understanding levels of PNM conceptual map, all students have a better participation in both OMC items and individual cognitive learning performance in class which could not be otherwise acquired. Since instructors' designation of OMC items is crucial to students' response and participation of conceptual understanding, the present research modifies Wilson's four building blocks of the construct map, items' design, outcome space, and measurement model (2005) to develop an instrument assessing students' chemistry equilibrium based on designs of OMC items. To consider the importance of diagnostic assessment for flexible OMC items could be explicated by an illustrated participation of students' individual conceptual understanding levels.

https://doi.org/10.33225/jbse/19.18.77

Methodology of Research

General Background

The primary research background enacted in this research specifically stressed students' conceptual understanding levels obtained through diagnostic assessment of OMC items. It suggested not only students' readiness and response of PNM pre-knowledge (Shepard et al., 2005) but also more possible diagnostic assessments for instructors to make further research requirements. To be different from both formative and summative assessment, diagnostic assessment provides the instructors with learning information about what students should learn and be able to do (Green & Johnson, 2010). For example, it often causes many students to get confused whenever they are confronted with complicated problems of PNM chemistry equilibrium in class. With the OMC diagnostic instrument, students could make correct choice items directly or indirectly for instructor's scaffolding learning framework. It would be hard for students to mediate the abstract behavior of the particulate of atoms or molecules in matters because of their poor visibility at macroscopic conditions. Before the class instructions, students are required to conduct a test in surf the net of OMC instrument during 2017/3~2018/5 to formulate their feedback and conceptual understanding levels represented by SPSS analysis.

Participants

All 548 participants were selected through two stages of qualification tests as research samples (Su, 2016). There were 173 university students who attended the author' course, Nature Science, at the first stage of qualification tests in October 2016. With the design of the full-scale experiment, the pilot test guided these 173 students (aged from 19 to 22 years) to gather validity and reliability data of the OMC items. The remaining 375 students were chosen for the normal participation of OMC diagnostic instrument at the second stage of qualification tests in the 2017 academic year. Table 1 shows students' background and characteristic analysis listed by characteristic, variance, number and percentage, such as enrollment (sport excellence, multi-star recommendation, learning achievement test, assigned test, uniform test and other tests). Since the participants of this research were limited to undergraduates in Taiwan, it should be careful that the approach might not go beyond the academic context within law inferences. All qualification tests at two stages would be carried out by participants for special precautions without violating local laws and legal agreement. It confirmed with the ethical procedures that all participants performed their publication in accordance with the full result understanding not to be anonymous and illegal as the ethical approval did (Taber, 2014) in 2017.

Instruments and Procedures

Each design of five hierarchical understanding levels was aligned with a corresponding level for students to assess their mental conception. The design framework of each corresponding understanding level can be a modified pattern of OMC instrument. The ordered multiple-choice (OMC) items of this research included four parts: major construct map, test items' design, outcome space, and measurement models. It is a constructed exercise of mental performances based on Wilson (2005). An exploration of the development of OMC diagnostic instrument with validity and reliability provides a substantial basis for designing students' corresponding level of mental framework.

Table 1. Students' background and characteristic analysis (N=375).

Characteristic	Variance	Number	Percentage (%)
School	1. CYCU	359	95.73
	2.TCU	16	4.27
Science	1. Major	24	6.40
	2. Non-major	351	93.60

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Characteristic	Variance	Number	Percentage (%)
Enrolment	1. Sport excellence	7	1.87
	2. Multi-star recommendation	39	10.40
	3. Learning achievement test	117	31.20
	4. Assigned test	146	38.93
	5. Uniform test	17	4.53
	6.Other tests	49	13.07
Gender	1. Male	136	49.45
	2. Female	239	50.55
Age	1. 18	81	21.60
	2. 19	225	60.00
	3. 20	43	11.47
	4. 21	14	3.73

To construct effective OMC diagnostic instrument, the lecturer first needs to define what proper instructional component should be accepted by students. Next, determine how to help students in getting more PNM progress to promote their understanding levels. In response to the initial requirements, Briggs et al. (2006) proposed the special test items as ordered multiple-choice (OMC), in discovering students' diagnostic assessment of five different understanding levels. One can approach five hierarchical OMC understanding levels, step by step from an ordered research finding (Hadenfeldt et al., 2013). There were five hierarchical conceptual levels for PNM understanding, indicated with scores and statements from level 1 to level 5 in Table 2. Starting from the naive concept of level 1 with basic PNM knowledge, students constructed their conceptual understanding in level 2 with the hybrid concept to explain PNM scientific phenomena. Equipped with both simple and differentiated particulate conception, students proceeded to the intermediate developments of level 3 and level 4. Finally, students can use the advanced particulate conception in level 5 to explain macroscopic PNM properties.

Table 2. Summary of five hierarchical OMC conceptual levels indicated with scores and statements.

Corresponding Level	Score Items	Participant Statement
1. Naive concept	1	The lowest level that students do not have much experience of how complicated PNM is.
2. Hybrid concept	2	Students have learned about PNM and explained scientific phenomena using a hybrid model.
3. Simple particulate concept	3	Using a simple particulate model to explain PNM phenomena.
4. Differentiated concept	4	Using a differentiated particulate model to understand particulate concept fine structure of sub-atoms.
5. Systemic particulate concept	5	Using an interaction of differentiated particulate to explain macroscopic properties of matter.

As for the validity and reliability of OMC test items, the difficulty index of test items *p* value was determined by students' fractional answering rate. The descriptive statistics of *p* values were indicated in four difficulty indices: (1) $p < 30\%$ to be regarded as the difficulty test item, (2) $30\% \leq p < 50\%$ as the difficulty towards easy test item, (3) $50\% \leq p < 70\%$ as the easy towards difficult test item, and (4) $70\% \leq p$ as the easy test item. Each separate *p* value was measured as an average percentage of cumulative scores for test items in students' conceptual understanding level. The function of *p* value was aimed at understanding students' distribution of difficulty index in OMC test items. All 21 OMC draft test items were adapted from research resources of Su (2015), Nyachwaya et al. (2011), Margel, Eylon and Scherz (2008). It was demonstrated that these draft test items had been already scrutinized, deleted and revised by four renowned chemistry professors and two science professors to provide the content validity of students' achievement test items. Furthermore, the corresponding OMC test items were answered by three senior

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

chemistry professors with more exact analyses of inter-rater reliability as Kendall's coefficient of concordance. The analyses opened up four possibilities for the following statistic coefficient: (1) When the coefficient was below 0.4, the analyses were developed from slight to fair agreement, (2) the coefficient between 0.4-0.6, developed to be substantial agreement, (3) the coefficient between 0.6-0.8, developed to be consistent agreement, and (4) the coefficient above 0.8, developed to be almost perfect agreement (Marozzi, 2014).

Data Analysis

From the analyses of Table 2, 19 OMC test items could be scored in 5 assessed statements according to students' conceptual understanding levels. The total score items covered the statements from level 1 to level 5 with different points from corresponding participants. All university participants were scored for assigning 19 OMC test items with 95 points for the total score. They could answer 19 test items within the time span of 45 minutes in class. It was available for participants to administer their OMC instrument via Wi-Fi from their cell phones from Wi-Fi or with a computer provided in the classroom. The aggregate of participants' responses to set up statistical information would be stored in an excel file, which eventually had been transmuted into the file of SPSS 22.0 Windows software to analyze each open-ended choice item.

Results of Research

Hierarchical Correspondence of OMC Instrument

In response to students' OMC options, three renowned chemistry professors were independently charged to examine the fulfillment of assignments and the differentiated validity of PNM understanding levels. Kendall's coefficient of concordance (W) was used to determine

statistic agreement by the three professors. The research finding W value 0.613 (c^2=33.107, p=.016) indicated that the substantial responses of OMC options in understanding levels were agreed upon by three professors. The OMC instrument included a set of 19 test items which were assigned to assess university students' PNM understanding level. Figure 1 was a detailed demonstration for No. 17 test item as one of the correct choices in OMC options. As seen in

Figure 1 from the summary of No. 17 test item, choice item A was a simple description of particulate model for PNM phenomena in conjunction with Level 3; choice item B was a hybrid model for PNM scientific phenomena with Level 2; choice item C was a naive model for not complicated PNM phenomena with Level 1; choice item D was an interaction of differentiated particulate model for macroscopic properties of matter with Level 5. Students could make out their own subsequent choice items developed from their understanding levels of hierarchical correspondence. These 19 test items were all terminological usages for OMC instrument which were published at Chinese Internet station (web address: https://goo.gl/forms/LUPm5mAgcn4xfzMv2). The aforementioned 19 test items provided a well-designed and profound instruction for students to participate actively OMC options for guidance.

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

17. Consider chemistry equilibrium reaction equation between dinitrogen tetroxide and nitrogen dioxide,

$N_2O_4(g) \rightleftharpoons 2NO_2(g)$, in Figure (a), at 100°C the flask is definitely reddish-brown owing to a large amount of $NO_2(g)$ present; and Figure (b), at 0°C the equilibrium is shifted toward colorless $N_2O_4(g)$. Which of the following chemistry equilibrium reaction between 100°C and 0°C changes will be the correct choice item?

○A. At 100°C Figure (a) products a large amount of NO_2 (g) present. (Level 3)

○B. At 0°C equilibrium shifting to left due to colorless in Figure (b). (Level 2)

○C. At both Figure (a) and Figure (b) finds it is a endothermic reaction. (Level 1)

○D. The chemistry equilibrium reaction relates to chemical bond formation and breaking, so higher temperature will favor chemical bond breaking. (Level 5)

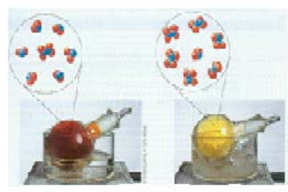

Figure 1. Summary of OMC instrument in No. 17 test item.

Conceptual Response and Analyses

In order to give more assessments for inspiring students' conceptual response, this research carried out the pilot test for 173 university students in the first stage. The existing OMC qualities and functions justified students' major conceptual response within different understanding levels. All 19 test items were designed for constructing the OMC structure and composition in PNM. Both the function and choices of 19 test items described in Table 3 characterized students' conceptual response of understanding levels within two major corresponding distribution areas. The corresponding distribution of *p* values to difficulty index has been analyzed in the following four statistic items: (1) $p < 30\%$ in reference to 1 difficulty test item; (2) $30\% \le p < 50$ reference to

Table 3. 173 students' responses of choice item mean score and fraction percentage for OMC test items.

Test Item Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Mean	2.78	2.01	2.80	2.72	2.42	3.29	1.68	2.19	1.40	4.28	2.77	4.17	3.60	2.79	3.55	2.39	2.42	3.40	3.42
P(%)	55.6	41.5	56.0	54.3	48.4	65.9	33.6	43.8	28.0	85.7	55.4	83.4	72.0	55.7	71.0	47.7	48.3	68.1	68.3

6 difficulties towards easy test items; (3) $50\% \le p < 70\%$ to 8 easy towards difficult test items; (4) $70\% \le p$ to 4 easy test items. In particular, all *p* values highlighted two statistic distribution areas, which ranged from difficulty towards easy test items to easy towards difficult test items. Table 3 illustrates what responsive difficulty index for 173 university students' engagement would set up the relevant distribution of *p* values with their separate test items.

The second stage gave an impetus for 375 undergraduate students to participate their authentic OMC instrument diagnoses of conceptual understanding level during 2017 academic year. While students were not expected to give a response of the same conceptual understanding level, each OMC instrument diagnosis would be considered to mark students' major response numbers of each test item in Table 4. As the second stage was administered, each choice item

https://doi.org/10.33225/jbse/19.18.77

Table 4. Choice item analysis for 375 students' response numbers of each test item.

Choice Item	1	2	3	4	5	6	7	8	Test 9	Item 10	11	12	13	14	15	16	17	18	19
1	121	52	101	46	62	19	78	69	105	19	47	112	139	22	85	28	102	33	57
2	38	146	108	128	225	115	139	47	139	145	131	72	104	87	88	184	84	72	84
3	52	126	88	118	42	129	72	209	49	165	116	77	110	147	121	73	112	96	79
4	164	51	78	83	25	20	62	23	82	46	81	114	22	88	81	91	77	91	155
5					21	92	24	27						31				83	

required students to mark their separate responses of conceptual understanding level. For example, most students inclined toward No. 5 test item 2 in their choice items. Meanwhile students chose relatedly few choice items with the disliked No. 6 test item 1 and No. 10 test item 1. Because differentiated phenomena illustrated that students could make their refinement of OMC instrument for conceptual response, their different understanding levels in assigned OMC test items would detect the core cognition of scientific matter within the whole progresses of Table 4.

With regard to 375 students' understanding levels, their distribution in Table 5 illustrates that the hierarchical levels of the L4 and L5 were sparser than those of the L1-L2 and L3 for PNMCE OMC. Accordingly, students' identification toward higher hierarchical levels remained in a more difficult position to attain than toward lower hierarchical levels in the OMC options. Students were required to make their responses of OMC 19 test items related to one level across a set of the scattering plot in Figure 2. Thus, each response option gave students to contemplate their thinking for a particular choice of hierarchical understanding levels. Students' speculation on respective OMC items manifested that L3 hierarchical understanding levels accumulated 47.4% from the total OMC options in conjunction with students' medium choice of understanding levels.

Table 5. Distribution of understanding levels for 375 students' responses of PNMCE OMC choice items.

Level Item	1	2	3	4	5	6	7	8	Test 9	Item 10	11	12	13	14	15	16	17	18	19
L1		103	78	46	129	154	158	92	188	65	47			22		101	112		
L2	121	146		128			139		187					147	85		84		
L3	254	126	196	118	246	129	78	283			328	184	236	87	88	274	102	292	220
L4			101	83		92								31	121				
L5										310		191	139	88	81		77	83	155

To obtain additional OMC options, L1 and L2 hierarchical understanding levels received 19.0% and 13.6% from the total OMC items, in conjunction with low-understanding level students. For students with L1 and L2 hierarchical understanding levels, we need to strengthen their PNM construction to avoid their particulate confusion with hybrid conceptualization. It was calculated that L4 and L5 hierarchical understanding levels were considered to have 6.5% and 13.5% from the total OMC items, in conjunction with high-understanding level. The immediate improved remedy for L4 and L5 hierarchical understanding levels was to assign students with more instructions for particulate concept structure of sub-atoms and macroscopic properties of matter. Of all 5 respective understanding levels, students with L4 and L5 hierarchical understanding were ranked as the advanced group who administered OMC item as specified and prominent participation in the scattering plot of Figure 2.

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

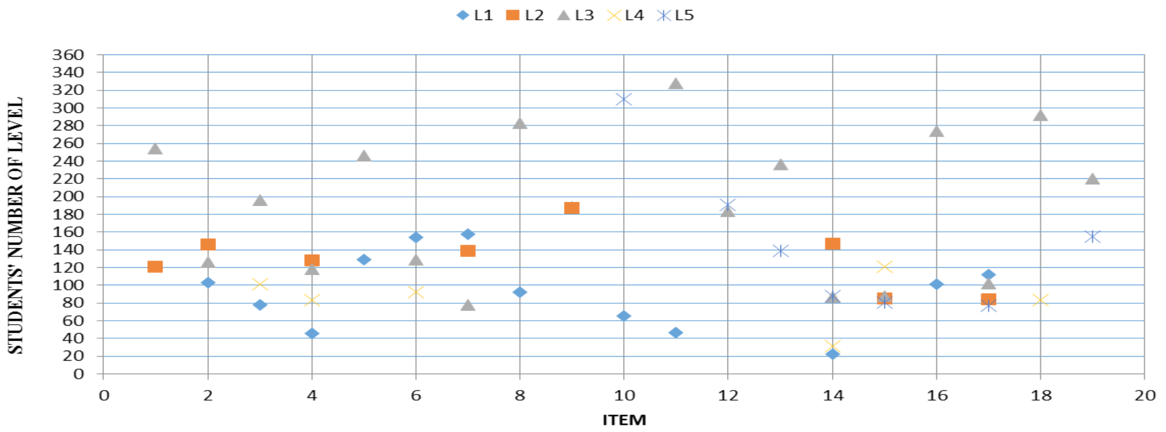

Figure 2. The scattering plot of understanding levels for 375 students' responses of 19 test items.

By conducting the previous 375 students' choices across 19 test items, 5 major percentages of hierarchical understanding levels were carried out to identify their responsive item numbers, with the differentiated order mark L3(3241, 45.5%)> L1(1295, 18.2%)> L5(1041, 14.6%) >L2(1037, 14.5%)>L4(511, 7.2%) in Figure 3. The authentic distribution of 375 students' responsive choices goes deep into 5 order marks of hierarchical understanding levels. It deserved careful attention that the order mark L3 occupied the most dominant percentage (45.5%) of total 375 students' responsive choices. The distributive test items of L3 understanding levels demonstrated students' familiar knowledge for the simple particulate model of PNM phenomena. Figure 3 also indicated other responsive item numbers of the lower students' 32.7% understanding levels for L1 and L2might be compared with those of the higher students' 21.8% understanding levels for L4 and L5.

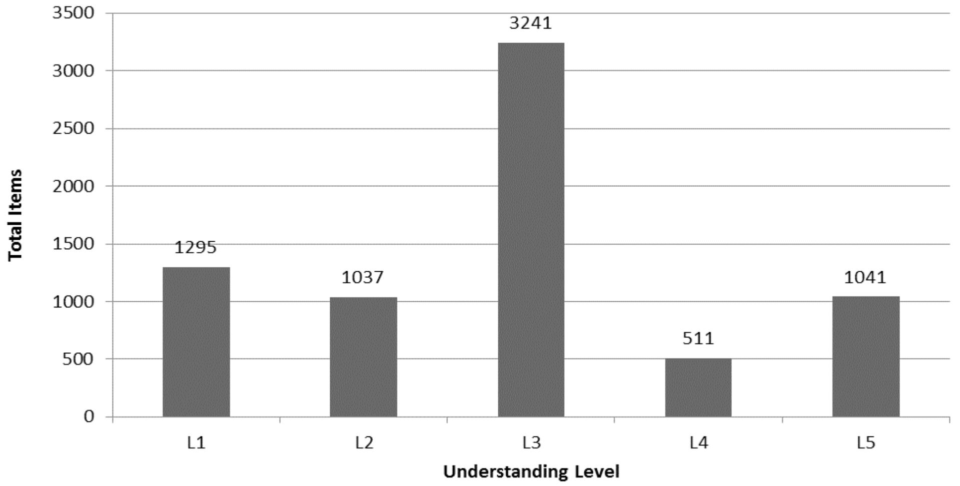

Figure 3. Distribution of understanding levels for 375 students' responses across 19 test items.

In order to find out students' divided distribution of item aspects, this study conducted the descriptive statistics of mean scores and standard deviation to measure four subscales QD, QMD, QME and QE in Table 6. It was available to demonstrate four subscales QD (difficulty test item), QMD (difficulty towards easy), QME (easy towards difficult) and QE (easy) with the total mean score of 2.94 and the standard deviation of 0.69 measured students' four different item aspects from difficult to easy. Students' participation of OMC test items validated a successful design to assess their authentic understanding levels and learning achievements from basic broad degrees to

https://doi.org/10.33225/jbse/19.18.77

the final depth of item aspects. The employment of mean scores (*M*) and standard deviations (*SD*) offered an important instrument to inspect students' learning effectiveness and understanding levels for their item aspects in class. Instructors would likewise apply all results of OMC item aspects to improve students' learning of the textual structure and meaningful contents.

Students' Divided Distribution of Item Aspects

Table 6. Mean scores (*M*), standard deviations (*SD*), and aspects for students' item tests.

Item Aspect	*M*	*SD*	Number	Test Item
QD	1.64	.47	2	7, 9
QMD	2.52	.30	9	1, 2, 3, 5, 6, 8, 11, 16, 17
QME	3.42	.40	6	4, 13, 14, 15, 18, 19
QE	4.16	.92	2	10, 12

Students' ANOVA and Relative Analysis

Table 7 would convey more accounts of OMC blocking variable with the guidance of a series of ANOVAs for the multi-variants of the Wilks' Lambda parameter in four subscales of item aspects. Accordingly, Table 7 suggested a brief summary of individual OMC item aspects to go with the discussion of the *F*-ratios, *p*-values and relative analysis (*r*). To regard the school as an example of the independent variable in Table 7, there appeared a significant difference (F=6.742, p=.010) of dependent variable QD (difficulty test item) which illustrated that QD of CYCU school students was superior to that of TCU school students in Taiwan, and the grades of CYCU (M=1.66, SD=.48) got higher scores than those of TCU school students (M=1.34, SD=.30). In contrast to QD of CYCU' school students' grades, the data shows that all dependent variables QMD, QME and QE were estimated without significant differences in Table 7. In developing more accurate evaluations of OMC item aspects, this research conducted Pearson's differentiate relative analysis (*r*) with students' four item aspects of dependent variable, represented the relative analysis range between .044 and .133 in Table 7. Judged from four dependent variables, only the QD dependent variable received significant negative relation (r=-.133, p<.01) as QD towards difficulty item aspect administered TCU school students to have an inclination for reducing consciousness learning achievement. Other three dependent variables, QMD, QME and QE did not provide significant relative analysis.

Table 7. 375 students' understanding level of item aspects for one-way ANOVAs in PNMCE.

Experimental Course	Blocking Variable	Analysis of Variance	Item Aspect			
			QD	QMD	QME	QE
PNMCE	School	*F-ratio*	6.742	.710	3.614	2.426
	(1.CYCU	*p-value*	.010*	.400	.058	.120
	2. TCU)	*r*	-.133**	.044	-.098	.080
	Science	*F-ratio*	.169	3.231	.600	4.198
	(1. Major	*p-value*	.682	.073	.439	.041*
	2.Nonmajor)	*r*	.021	-.093	-.040	.106
	Enrollment	*F-ratio*	2.006	.990	.730	2.226
	(1.Sport test	*p-value*	.007	.424	.599	.051
	2. Star test	*r*	.044	-.075	.070	.021
	3.School test					
	4.Indicate test					
	5. Present test					

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Experimental Course	Blocking Variable	Analysis of Variance	Item Aspect			
			QD	QMD	QME	QE
	6. Other)					
	Gender	*F-ratio*	.182	.000	1.541	8.068
	(1. Male	*p-value*	.667	.990.	.215	.005**
	2. Female)	*r*	.022	.001	.064	.146**
	Age	*F-ratio*	.151	1.520	1.925	.546
	(1.18 2.19	*p-value*	.963	.196	.106	.702
	3. 20 4.21	*r*	-.008	.011	.039	-.034
	5. 22)					

Note: $^{*}p< .05$; $^{**}p< .01$.

In the area of science, science major students (*M*=4.19, *SD*=.91) received higher scores than those of non-science major students (*M*=3.79, *SD*=.98). Taking science as an independent variable, there appeared a significant difference (*F*=4.198, *p*=.041) of dependent variable QE (easy test items) in Table 7. Four dependent variables marked students' science range of relative analysis between .021 and .106 in Table 7. Only the QE dependent variable received significant positive relation (*r*=.106, *p*< .05) which meant major science students to have more inclinations for good consciousness learning achievement than those of non-major science students. There were no more significant relative analyses for other three dependent variables, QMD, QME and QD.

In the area of gender, male students (*M*=4.26, *SD*=.86) scored higher than female students (*M*=3.99, *SD*=1.00). Taking gender as an independent variable, a significant difference (*F*=8.068, *p*=.005) of dependent variable QE (easy test items) is shown in Table 7. There were four dependent variables for students' gender range of relative analysis between .001 and .146 in Table 7. The only manifestation might be attributed to QE dependent variable of significant positive relation (*r*= .146, *p*< .01) for requiring more male students to perform inclined perceptions of learning achievement than female students did. No more significant relative analysis existed for other three dependent variables, QMD, QME and QD. In both the enrollment and age areas, there did not appear any significant differences (*p*>.05) for ANOVAs and relative analysis accorded with four dependent variables QMD, QME, QE and QD in Table 7.

Accordingly, in the case of five blocking variables with four dependent variables, the final statistical analysis appeared to indicate that students' good command of performed perceptions would affect their learning achievement. It is yet to be determined whether there were not any significant differences within the relative analyses of four dependent variables. More science major and male gender students would only take the easy test items in the performed assumption of dependent variable for both significant differences and significant relation. CYCU school students had a good command of the difficulty of test items which illustrated more performances of beneficial scores than those of TCU school students. A continuous research confirmed that TCU school students made up a significant negative relation in QD dependent variable towards difficulty item aspect.

Discussion

A critical problem for this research is to adapt and validate students' consistent agreement by three senior chemistry professors. Through the designed content of inter-rater reliability for Kendall's coefficient of concordance, it has been analyzed that there are four agreements to be accorded with the statistic coefficients: (1) from slight to fair agreement, (2) substantial agreement, (3) consistent agreement, and (4) almost perfect agreement (Marozzi, 2014). A comparison has been made for this study to focus upon the similarity of the consistent agreement under the four agreements of corresponding OMC instrument with the functional validity and reliability in this study. Students' response of the consistent agreement requires a subsequent self-development of choice items from their understanding levels of hierarchical correspondence within 19 test items. Further researches on the hierarchical correspondence of OMC instrument would clarify students' five understanding levels starting from (1) naive concept (2) hybrid concept (3) simple particulate concept (4) differentiated concept and (5) systemic particulate concept up to the most perfect level (Hadenfeldt et al., 2013).

https://doi.org/10.33225/jbse/19.18.77

Most college students, imbued with the cognitive performances of conceptual response, have better diagnostic responses of ordered multiple-choice (OMC) instruments than those of traditional multiple-choice (TMC) items in their substantial PNM domains of understanding levels (Hadenfeldt et al., 2013; Othman, Treagust, & Chandrasegaran, 2008; Su, 2015). It has been accepted that students could make assigned choice items in different OMC understanding levels as their conceptual response of non-linear scientific learning (Neumann, Viering, Boone, & Fischer, 2013). In fact, students' responsive distribution of choice items varies with their different conceptual performances of assigned OMC test items in core scientific structure and composition of matter. The responsive focus of students' conceptual analyses makes a foremost concentration on L3 understanding levels, to be the highest percentage for their entire engagement in the simple particulate model of PNM phenomena. For other analyses of students' conceptual response, only 21.8% students give their responsive distribution on both higher understanding levels of L4 and L5, and 32.7% students occupy their responsive distribution on both lower understanding levels of L1 and L2. More endeavor for students of lower understanding levels would be needed in their refinement of PNM conceptual response to accumulate their potential impact of higher understanding levels by flexible test item designs (Wilson, 2005; Yakmaci-Guzel, 2013).

Students' divided distribution of item aspects provides an intrinsic factor to assess their authentic understanding levels in four subscales starting from QD, QMD, QME and mediating up to QE. This research makes more applications in detail for students to participate their improved learning of the textual structure and meaningful contents. Several scholars have explored how the most complicated and abstract nature of atoms and molecule gives different relevant bearing for students PNM misconceptions in their understanding levels (Othman et al., 2008; Özmen, 2013). Additionally, students' responsive distribution of OMC test items often misplaced their understanding performances either stuck in fragmented mental models or in confused obstacles without cognitive development of chemistry structures and behavior of matter (Margel, Eylon, & Scherz, 2008; Nakhleh, Samarapungavan, & Saglam, 2005; Yezierski & Birk, 2006). They also advocated many disciplinary methods for students to take favorable advantages of OMC items in their developed response of diagnostic interpretations (Briggs, Alonzo, Schwab, & Wilson, 2006; Alonzo & Steedle, 2009).

In the corresponding case of the above four dependent variables with five blocking variables, students' command of difficulty test items reflects a dominant dependent variable in both significant differences and significant relation. For example, science major and male students are good at the easy test items, and the students of CYCU school make higher scores than those of TCU school towards difficulty test items. As Yakmaci-Guzel (2013) emphasized this assertion, more flexible discourses of consistent dependent variable should be discovered in developing students' problem-solving performances. A close coalition of dependent variables with students' performances is related to the correspondent fulfillment of responsive analysis. There would be more engagements in students' manipulating the functional skills of ordered multiple-choice (OMC) items, which include the cases of semi-structure interviews (Alonzo & Steedle, 2009; Coll & Treagust, 2003), diagnostic assessment (Treagust & Chiu, 2011) and traditional multiple-choice (TMC) items (Briggs et al., 2006; Othman et al., 2008).

A favorable goal for five hierarchical correspondence of OMC instrument allows students to accumulate their PNM understanding level, set up by instructors' flexible OMC textual designs. Students are encouraged to have direct control of increasing learning response and assess their potential impact of complicated and abstract concepts in different understanding levels. For further framework of the functional skills, the research has to mediate the performance of lower understanding level students in differentiating the correct answer for particulate concept structure of sub-atoms and macroscopic properties of matter as higher understanding level students do in OMC instrument.

Conclusions

The key function of OMC instrument serves to facilitate students' correct manipulation in PNM conceptual understanding levels. At first, this research set up the disciplinary design of students' mental development based on OMC choice items with respect to PNM conceptual understanding. The full-fledged design of five hierarchical conceptual levels for OMC items led students to follow a noticeable mental conceptualization on a convenient scale. A close look at students' OMC items would elucidate to a large degree in this research that their cognitive response of conceptual understanding was different from one matrix level of TMC items. The whole procedure also provided an effective assessment, in which senior chemistry professors examined the corresponding validity and inter-rater reliability to reflect students' understanding levels in detail.

https://doi.org/10.33225/jbse/19.18.77

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

The three findings of OMC instrument research offer several benefits. First, the implementation of OMC functioned successfully in class to assess and detect different levels of students' PNM understanding and learning progress. Second, as students fulfilled individual performances of conceptual understanding, they could accumulate more effective core chemical concept. Third, for further impact of advanced OMC items, this study offered instructors an easy- to-check tool in scrutinizing students' previous selections of OMC items which mostly had been hidden within their perceptive organization for interactive potentials. More flexibilities of OMC instrument would be taken into alternative measurements for both students of lower and higher understanding levels. The remaining finding of OMC instrument would not be restricted to the single assumption in adjusting to only one approach of students' understanding levels.

The survey of students' PNM understanding levels made up more extended approaches for further OMC instruments. Scholars have agreed that many disciplinary designs would be explored for students to take favorable advantages of OMC items in their utmost development of individual responsive understanding levels. Although students' conceptual responses were apt to act out for non-linear scientific learning, more adjustments of OMC items would be in great need to revise and adapt students' misfitting breakdown in overlapping concepts in the future. An important indicator of OMC items would be needed in the extended study to link more students' advanced performances with different understanding levels to carry out the feasible guidance as the fundamental framework of the functional skills.

Acknowledgements

A short but sincere thank must also be given to the patronage of the Ministry of Science and Technology, MOST in Taiwan (under Grant No. MOST 105-2511-S-237-002-MY2). Without their help and financial support, this research could not have been completed in the present form.

References

Alonzo, A. C., & Steedle, J. T. (2009). Developing and accessing a force and motion learning progression. *Science Education, 93* (3), 389–421.
Bhattacharyya, G. & Bodner, G. M. (2005). It gets me to the product: how students propose organic mechanisms. *Journal of Chemical Education, 82*, 1402-1407.
Bell, B. (2007). Classroom assessment of science learning. In S. K. Abell and N. G. Lederman (Eds.), *Handbook of research on science education*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 965-1006.
Briggs, D. C., & Alonzo, A. C. (2009). The psychometric modeling of ordered multiple-choice item responses for diagnostic assessment with a learning progression. *Learning Progressions in Science* Conference*: Current Challenges and Future Directions*, Iowa City, IA, June, pp.293-316.
Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, *11*(1), 33-63.
Bunce, D. M., & Gabel, D. (2002). Differential effects of males and females of teaching the particulate nature of chemistry. *Journal of Research in Science Teaching, 39*, 911–927.
Coll, R. K., & Treagust, D. E. (2003). Learners' mental models of metallic bonding: A cross-age study. *Science Education, 87*, 685–707.
Ferguson, R. L., & Bodner, G. M. (2008). Making sense of arrow pushing formalism by chemistry majors enrolled in organic chemistry. *Chemistry Education Research and Practice*, *8*, 99-110.
Green, S. K., & Johnson, R. L. (2010). *Assessment is essential*, New York: McGraw-Hill.
Grove, N. P., Hershberger, J. W., & Bretz, S. L. (2008). Impact of a spiral organic curriculum on student attrition and learning. *Chemical Education Research and Practice, 9*, 157–162.
Hadenfeldt, J. C., Bernholt, S., Liu, X., Neumann, K., & Parchmann, I. (2013). Using ordered multiple-choice items to assess students' understanding of the structure and composition of matter. *Journal of Chemical Education*, *90*, 1602-1608.
Margel, H., Eylon, B. S., & Scherz, Z. (2008). A longitudinal study of junior high school students' conceptions of the structure of materials. *Journal of Research in Science Teaching, 45*(1), 132–152.
Marozzi, M. (2014). Testing for concordance between several criteria. *Journal of Statistical Computation and Simulation*, *84* (9), 1843-1850.
Nakhleh, M. B., Samarapungavan, A., & Saglam, Y. (2005). Middle school students' beliefs about matter. *Journal of Research in Science Teaching, 42*, 581–612.
Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, *50*, 162–188.
Othman J., Treagust, D. F., & Chandrasegaran, A. L. (2008). An investigation into the relationship between students' conceptions of the particulate nature of matter and their understanding of chemical bonding. *International Journal of Science Education, 30*(11), 1531–1550.

https://doi.org/10.33225/jbse/19.18.77

Özmen, H. (2013). A cross-national review of the studies on the particulate nature of matter and related concepts. *Eurasian Journal of Physic and Chemistry Education, 5*(2), 81-110.
Potgieter, M., & Davidowitz, B. (2011). Preparedness for tertiary chemistry: Multiple applications of the Chemistry Competence Test for diagnostic and prediction purposes. *Chemistry Education Research and Practice*, *12*, 193–204.
Sanger, M., Phelps, A., & Fienhold, J. (2000). Using a computer animation to improve students' conceptual understanding of a can-crushing demonstration. *Journal of Chemical Education*, *77*(11), 1517–1520.
Shepard, L., Hammerness, K., Darling-Hammond, L., & Rust, F. (2005). *Assessment*. In L. Darling-Hammond and J. Bransford (Eds.), *Preparing teachers for a changing world: what teachers should learn and be able to do*, San Francisco, CA: Jossey-Bass, pp. 275-326.
Su, K. D. (2015). The development and application for a three-tier diagnostic instrument to assess college students' misconceptions in chemistry equilibrium. *Chinese Journal of Science Education, 23*(4), 321-352.
Su, K. D. (2016). Strengthening strategic applications of problem-solving skills for Taiwan students' chemistry understanding. *Journal of Baltic Science Education, 15* (6), 662-679.
Taber, K. S. (2014). Ethical considerations of chemistry education research involving 'human subjects.' *Chemistry Education Research and Practice*, *15*, 109-113.
Treagust, D. F., & Chiu, M. H. (2011). Diagnostic assessment in chemistry. *Chemistry Education Research and Practice*, *12*(2), 119-120.
Wilson, M. Constructing Measures: An Item Response Modeling Approach; Lawrence Erlbaum: Mahwah, NJ, 2005.
Yakmaci-Guzel, B. (2013). Preservice chemistry teachers in action: An evaluation of attempts for changing high school students' chemistry misconceptions into more scientific conceptions. *Chemistry Education Research and Practice*, *14*, 95-104.
Yezierski, E. J., & Birk, J. P. (2006). Misconceptions about the particulate nature of matter. Using animations to close the gender gap. *Journal of Chemical Education, 83*, 954-960.

Received: *November 07, 2018*   Accepted: *January 16, 2019*

King-Dow Su	Chemistry PhD, Professor, Department of Hospitality Management and Center for General Education, Hungkuo Delin University of Technology, NO.1, Lane 380, Ching-Yun Road, Tu-Cheng District., New Taipei City, Taiwan 23646, R.O.C. & Center for General Education, Chung Yuan Christian University, 200 Chung Pei Road, Chung Li District, Taoyuan City, Taiwan 32023, R.O.C. E-mail: su-87168@mail. hdut.edu.tw

https://doi.org/10.33225/jbse/19.18.77