The influence of working memory capacity on academic achievement of final year medical students

Sivalingam Nalliah

Abstract: This is a cross sectional study conducted in July 2010 at the International Medical University, Seremban, Malaysia. The objective of this study was to ascertain the relationship between working memory capacity of final MBBS medical students using the digit span backward test and their academic achievement based on the total score at the modified essay questions (MEQ) which was the principal component of the theory examination. Seventy eight final year medical students were recruited, 41 (52.6%) were females and remaining 37 (47.4%) were males. Working memory capacity was measured by digit span backward test (DSBT) which ranged from 3 to 8 digits. The mean digit score was 6.6 ± 1.1 falling under the category of 'above average' score. There was no significant difference between working memory capacity and gender (p>0.05). There was no significant difference in the MEQ mean score and the different categories of working memory capacity (p>0.05). The DBST shows uniformity in working memory adequate to pass the modified essay questions. Medical students appear to use encoding and retrieval process in problem solving based on functionality and pattern recognition in tackling the problems in the MEQ.

IeJSME 2012: 6 (Suppl 1): S87-S102

Key words: Working memory, digit span backward test, academic performance, modified essay question

Introduction

The pioneering work of Atkinson & Shiffrin (1968) on the human memory as a 'multistore' exemplified a logical three phase approach to information processing (IP) (i.e sensory, working and long term memory). Expounding on this model, the role of short term memory (STM) in perceiving and receiving information, attending to them and processing relevant information into chunks which eventually get encoded through effective cognitive functions into long term memory has been the focus of research. Information in sensory memory lasts for 250 milliseconds for visual memory and 1-2 seconds for auditory memory. Information that 'is attended to' is encoded into working memory (WM), the duration being limited to less than 18 seconds (Atkinson & Shiffrin, 1968).

Extensive scientific work in human cognitive psychology led to further understanding of 'short term' memory with the introduction of the alternative term 'working memory' (Daneman & Carpenter, 1980; Baddeley, 1986; Shute, 1991; Engle & Kane, 2004). Working memory (WM) and more specifically working memory capacity (WMC) is the focus for the differential characteristics seen among individuals in relation to their academic achievement (Hunt, 1999). This is especially so when domain-specific activities were evaluated in disciplines like mathematics, chemistry, agriculture, pharmacy and medicine. Some of these disciplines, especially medicine and pharmacy, invoke experiential learning which lend to adoption of learned principles and skills apart from involving the affective domain (Cassidy, 2004; Noble, Miller & Heckman, 2008).

Continuous challenges seen in problem analysis and problem solving are now dominant in the medical education curriculum as integrated spiral curricula enable learners to use problem based learning and task based learning in the clinical setting as the fore-front of both pedagogy and androgogy (Harden, 1999; Norman & Schidmt, 2000; Epstein, 2007).

As medical education involves a myriad of learning opportunities and related learning activities that sustain a combined contextual presentation that has all three components of cognition, psychomotor and affective domains, the learning is both analytical and authentic.

The aim of this study is to determine if working memory capacity (WMC) impacts on academic achievement of medical students at the end of 4-5 years of rigorous learning and clinical instructions. The principles of working memory (WM) as the processing unit will be tested using appropriate instruments and the findings

For Correspondence:

Professor Dato' Sivalingam Nalliah, Clinical School, International Medical University, Jalan Rasah,70300 Seremban, MALAYSIA

Email: sivalingam_nalliah@imu.edu.my or nelsonsivamani@gmail.com

will be correlated to suitable measures of academic achievement.

Materials and Methods

Final year medical students who had successfully completed part 1 of the Final MBBS examinations at the end of semester 9 and had entered the senior clerkship in semester 10 in the last six months of a five-year medical course at the International Medical University were recruited to participate in the study.

Research Questions

i. What is the working memory capacity of final year medical students?
ii. What is the working memory capacity of final year medical students according to gender?
iii. Is there a relationship between working memory capacity (WMC) of final year medical students and their academic achievement at the modified essay questions?

Research Hypothesis

Null Hypothesis and Alternate Hypothesis

Ho(1): There is no significant relationship between WMC and academic achievement at MEQ among final year medical students

Ha.(1): There is a significant relationship between WMC and academic achievement at MEQ among final year medical students.

Data on gender and their academic achievement at the MEQ were also obtained.

Sample

This quasi-experimental study was conducted at the Clinical School, International Medical University from 15^{th} – 31^{st} July 2010. All final year (semester 10) MBBS students who were resident at the Batu Pahat campus of the International Medical University who volunteered to participate in the study were recruited for this study. The reason for choosing this sample was the convenience of conducting the study, the accessibility to their modified essay question marks that they had taken in February 2010 at the end of the semester 9 examinations (Part 1, Final MBBS) and the possibility of having a rather homogenous population of medical students as far as cognitive functions are concerned. No change in cognitive behaviour and learning style was expected in the intervening period of five months (from February 2010 till July 2010). All volunteers signed an informed consent using standard consent forms available from the Research and Ethics Committee of the university.

Working memory capacity

Working memory capacity of final year medical students was determined by the reversed digit span backward test (DSBT) developed by Johnstone (2001). The administration of the DSBT has been found to be simple and allowed for determination of working memory span as a correlate of WMC. The DSBT has been validated for consistency and is applicable in subjects from different cultural backgrounds. It has also been shown that reversed digit span test would also be a measure of all the main components of WMC; Baddeley & Hitch (1974) attested to it being an appropriate tool for evaluating WMC.

All the instructions in the user manual were followed to minimize recency effects and working memory overload. Subjects were directly supervised by the researcher and care was taken to ensure that the answer matrix provided for entering the digits in reverse were only entered after 30 seconds of reading out each row of digits. This procedure was repeated for the entire test and this test took 12 minutes to administer.

Copies of the instrument were made available and subjects were informed about maintenance of anonymity by requesting only their identification number and gender to be placed on the front page of the response sheet. As the average age of the subjects was 24 (range 23-26 years) this data was not requested. Subjects were told about the basis of the WMC and the need to attend to the digit span read out by the researcher and to 'rehearse' in their memory before entering the columns provided for entry.

The test paper had two sets of digits for each span (ranging from 3 to 8). Each digit span was read in sequence beginning from three digits and increasing the task to up to eight digits. After each digit span was read, thirty seconds elapsed before subjects entered their answers in reverse order. To avoid errors in the experiment (as subjects could enter the digits in the reverse as they were read), they were instructed not to make paper-pencil contact till thirty seconds of the completion of the reading of the digit span. A series of digits that were administered in the beginning of the test were for practice.

After completion of the task, scores for the DSBT were scored manually. Of the two spans for each set, any one correct entry (of digit span attempted) was accepted as representative of the subject's score. The maximum band score for the span test (at least one of the two sets for the band) was accepted as reflective of the WMC. Categorical data into four bands was derived based on the following: low score (<4), average (4-5), above average (6-7) and high score (>8).

The modified essay question and academic achievement

The modified essay question (MEQ), if well constructed could be used for evaluating problem solving skills. In medical undergraduate education this tool could be crafted so as to direct students to comprehend, reason out the responses in relation to the problem posed, analyze it and organize the answer after solving the problem before writing concisely the latter. While this could overcome the difficulties of some students who lack the skills of writing answers in medicine in narrative, they have been shown to be efficient in testing problem solving skills. The link to cognitive psychology in relating the function of working memory appears to be meaningful when MEQs are used to assess medical students.

The current study would attempt to explore the relation between the scores of working memory span (WMC) to academic achievement at MEQs in final year medical students. The dependent variable was academic achievement (AA) of the final year medical students. The MEQ scores were used as a measure of AA. The MEQ examination was conducted in February 2010 and all the respondents sat for that examination and had passed five months before the conduct of the study.

Permission was sought from the Academic Affairs of the Clinical School, International Medical University. The MEQ is the only theory paper taken at this high stake examination in addition to clinical examinations (OSCE) and the objective standardized practical examination (OSPE). The MEQ, a three hour theory paper, consisted of a total of six questions. Each question consisted of five parts of varying number of stems (with an average of 3-5 stems) to be completed in 30 minutes. Each part was allocated a pre-determined time duration by the examination committee. Each question tested students on one main domain (e.g. internal medicine, surgery, obstetric and gynecology, pediatrics and family medicine and psychiatry) with inclusion of other minor domains within each question. It consisted of 20 per cent of the total score for the MBBS (the other components being continuous assessment marks, the OSCE and OPSE). MEQ scores ranged from 0-20 with fail being less than 10/20 and distinction being >15/20. Scores at the MEQ (computed to a maximum of 20) were matched against the identification number of the respective subject obtained from the DSBT paper taken by each subject.

The researcher took cognizance of the value and limitations of the construct of the MEQ as to how it would reflect on higher cognitive functions as two studies mentioned the inclusion of up to 50% of questions that may be dependent on recall which effectively tests long term memory (LTM) and not WM (Knox, 1975; Felliti & Smith, 1986; Irwin & Bamber, 1982). The study design is shown in Fig. 1.

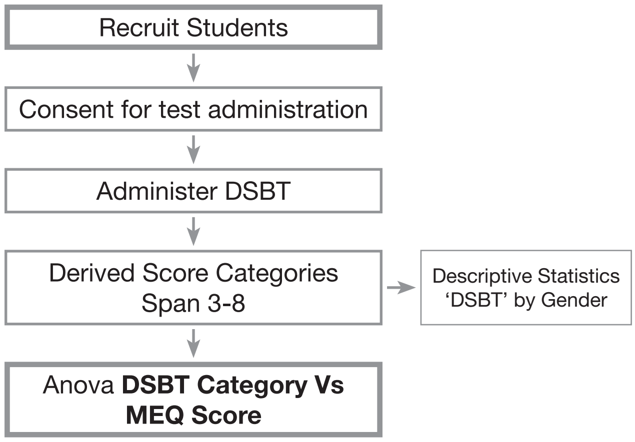

Figure 1 : Study Design

SPSS version 11.5 was used for statistical analysis. Appropriate tables and graphs were generated to illustrate the results derived from the study. The ANOVA statistical test of differences in means of more than 2 independent categories was used to calculate association of band achievement (low 3-4, average 6-7, high 8) at DSBT and academic achievement at the final part 1 MBBS modified essay questions examination (MEQ scores). The DSBT band scores were analyzed for relationship to gender using chi-square.

The research was approved by the Centre for Medical Education (Research and Ethics Committee) of the International Medical University, Kuala Lumpur. The project was funded by a grant from the International Medical University.

Results

There were 84 students in this cohort but only 78 (92.9%) consented to participate in the study. The mean score for DSBT was 6.6 (±1.1, range 3-8) (Figure 2).

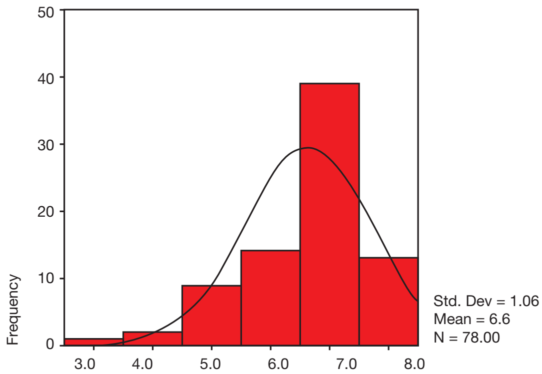

Figure 2: Distribution by DSBT scores

The DSBT frequency by span is shown in Table 1. Cumulative frequency of subjects with span 6 and above was 64/78. Three subjects had span of less than 4.

Table 1: Digital Span Backward Test Score

Digital Span Backwards Test Score	Frequency	Percent	Cumulative Percent
3	1	1.3	1.3
4	2	2.6	3.8
5	9	11.5	15.4
6	14	17.9	33.3
7	39	50.0	83.3
8	13	16.7	100.0
Total	78	100.0	

When distribution by gender was analyzed, (n=78; male=41, female=37) no significant differences were seen between digit span and gender (p=0.063; Table 2).

Table 2: Gender distribution and Digital Span Backward Test Score

Gender	Digital Span Backward Test Score						Total Number	P Value
	3	4	5	6	7	8		
Male	1	2	2	11	18	7	41	
Female	0	0	7	3	21	6	37	***0.063***
Total	1	2	9	14	39	13	78	

The mean MEQ marks for all spans were between 12.1-12.2. As shown in Table 3, no significant difference was seen in mean MEQ scores among the four DSBT categories (ANOVA f=1.354, p=0.917).

Table 3: Digital Span Backward Test Score and Performance at Modified Essay Question (MEQ)

Digital Span Reversed Test Score	Mean MEQ marks	Std. Deviation	N	P value
Low score	12.2	1.3	12	
Average score	12.1	1.4	14	0.917
Above average score	12.1	1.2	39	
High score	12.2	1.3	13	
Total			78	(Anova test)

(Anova test: Test of difference in means of more than 2 independent category)

The box plot in Figure 3 shows the distribution of subjects by span against MEQ. The mean hovered around 12 marks.

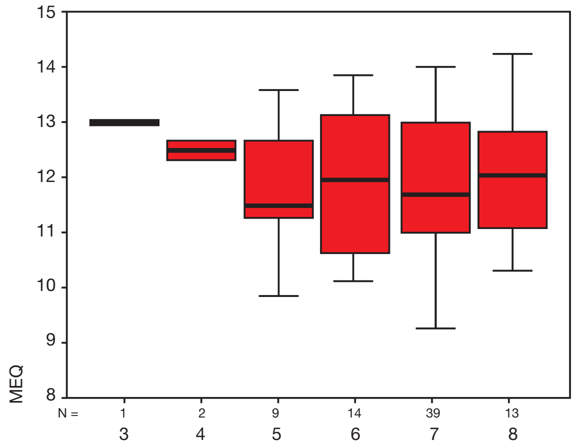

Figure 3: Box plot of MEQ scores by DSBT Score

When mean overall marks were analyzed in relation to DSBT, again no significant difference was seen (ANOVA f=1.56, p=0.864; Table 4). This result is exemplified in the box plot (Figure 4). The frequency distribution of overall marks is shown in Table 4 & Figure 5. The overall mean score for this cohort was 62.5 (± 4.7).

Table 4: Digital Span Backward Test Score and Overall Final Marks

Digital Span Backwards Test Score	Mean Overall marks	Std. Deviation	N	P value
Low score	62.5	3.6	12	
Average score	62.7	6.0	14	0.864
Above average score	62.2	4.5	39	
High score	63.0	5.3	13	
Total	62.5	4.7	78	(Anova test)

(Anova test: Test of difference in means of more than 2 independent category)

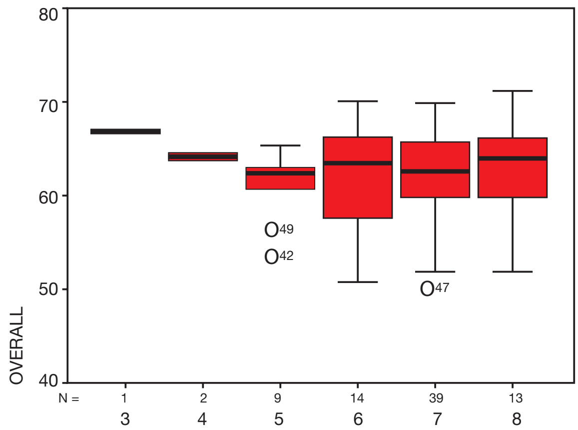

Figure 4: Box plot of Overall Score by DSBT

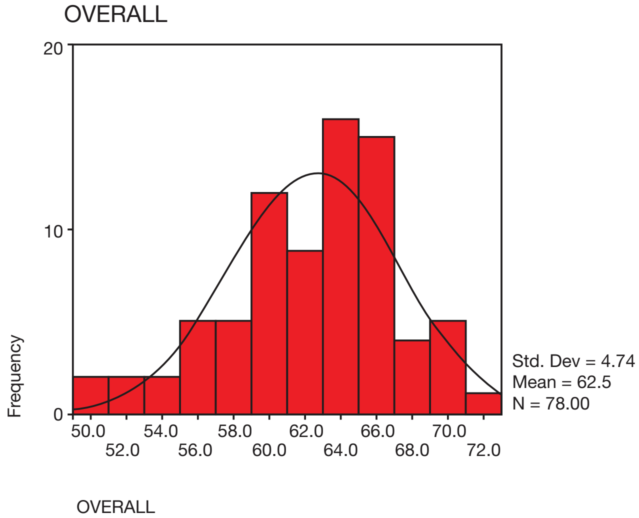

Figure 5: Frequency Distribution of Overall Marks

Discussion

WM is a key cognitive function that is used in daily work that helps selected information to be held for brief periods of time (less than 20 seconds). It develops during childhood, confirming Piaget's developmental learning theory. The capacity of WM grows about one unit for every two years of life declining with age (>30 years) (Hunt, 1999). Very simply put, it is the mental work space that has been central to various intellectual functions. Although controversial, 50% of variance in intelligence is said to be related to differences in WMC.

Baddeley & Hitch, 1971; Baddeley, (2000); Johnstone & El-Banna, 1986; Johnstone & El-Banna (1989); Pascual-Leone, 1987 have suggested the presence of components and sub-components of WM and their cognitive functions to explain the processes that come into play. The model of Baddeley and Hitch, 1974 and Hitch, 1978 is referred to in this study. A central executive control is present in this model with two major slave systems i.e. an articulatory (phonological) loop and a visual-spatial sketchpad component (Fig. 5).

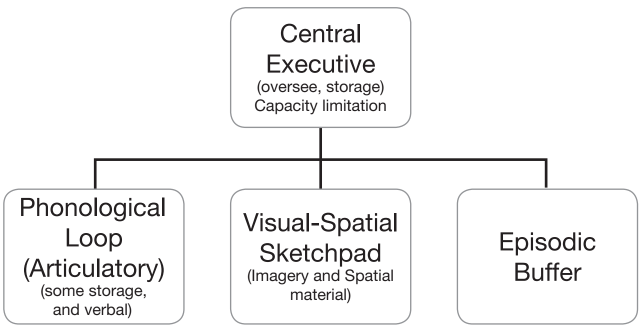

Figure 5: Working Memory Model of Baddeley & Hitch (1974)

This study adopted the DSBT (Johnestone, 2001) to evaluate WMC in final MBBS students. As shown in Table 2, the digit span ranged from 3-8 with a majority having a digit span of 6 (68/78) and nine having a digit span of 5. The mean score was 6.6 (± 1.06) skewing to the left (Fig. 2). When evaluating for gender differences, no significant difference was seen between males and females (Table 4.5, p=0.063). As shown in Table 3 there was no significant difference in DSBT and MEQ scores (p=0.917). In fact the mean scores in MEQ within the DSBT band were almost similar. The mean was either 12.1 or 12.2 of a maximum score of 20 marks in each category. This was reflective of the overall marks scored by this cohort where the mean was 12.1 ± 1.20, (n=78). When DSBT was evaluated against overall marks, no significant differences were seen (p=0.864).

As WM is a multi-component functional unit of information processing (IP) it is prudent to explain the use of the DSBT as a test of WMC and to attempt to answer the research question if MEQ is a suitable model for meeting the criteria of problem solving and being analytical in answering the MEQs (Opdenackera, Fierensa, Van Brabanta, Sevenantsa, Spruyta, Slootmaekersa & Johnstone, 1990; Palmer & Devitt, 2007).

Conway, Kane, Bunting, Hamrick & Engle (2005), Bunting, Conway & Heitz (2004), Cantor & Engel (1993) and May, Hasher, & Kane (1999) used psychometric tests as a measure of WM span (WMC). They concurred that 'WMC is an important individual–differences variable and accounts for a significant portion of variance in general intellectual ability'. Our study showed that the mean DSBT (as a measure of WMC) is 6.6 ± .06. There was little variation in the cohort studied and one needs to wonder if final MBBS students tend to have an adequate intellectual capability to meet assessment standards in this vocation with such scores in DSBT. Excluding the small majority with low scores (12/78, <5 band) and assuming that the mean MEQ of 12.1 is sufficient to meet academic standards to pass the theory paper in the final medical examination, one can conclude that in a field like medicine, at high stake examinations, where problem solving and higher order thinking is required, a digit span of six (6) and above is adequate. Individual variations are probably related to more efficient chunking before encoding. The other factors that support this proposal are discussed below when novice-intermediate–expert model is elaborated.

Multiple assessment tools are used to evaluate cognition, affective domain and skills training (procedural medicine). It would be incumbent on teachers to use a multitude of evaluation methods that would effectively assess students so that they would meet the requirements of clinical practice (Howe, Campion, Searle & Smith, 2004).

For the cognitive evaluation tools, a paradigm change from tests that were based on low level of the Bloom's taxonomy (Bloom, 1956) to higher order cognition tests have been introduced. Questions are now rated based on clinical situations which exhibit realism rather than test factual knowledge through recall. While multiple choice questions can be designed for this purpose, research in assessment and evaluation methods in medical education has changed over the years. Long essays have been known to have its own limitations as it is often focused on a limited domain for testing and suffers from the ability to cover a broad range of subjects.

Short answer questions were thought to be important to overcome this problem but again there were limitations and the lack of purpose in not testing higher order thinking.

Applying Miller's chunking theory, DSBT scores exceeding six (6) appeared to be adequate to deal with the WMC's ability to 'work' with the requirements of stems presented at the MEQ. The MEQ was designed to test the students' ability to use prior knowledge (from LTM) to solve current problems. In each of the MEQ set taken by the students, a clinical scenario was posed followed by short answer questions. Subsequent stems in the question evaluated different aspects of the original problem (clinical trigger) posed to assess the student's competence in managing the case with increasing difficulty. The subsequent questions (of the five parts of each question) would begin by providing the answers to the original case scenario, so as not to jeopardize answering the new case scenario posed, so that yet another complication that arose from the original case scenario was tested. At least five such questions appeared in one set for each of the six disciplines of medicine examined and students were asked to answer each set within stipulated times. Each question that had been answered was taken away as soon as they were answered, so the student would not have access to the previous question.

The internal consistency and the construct of the MEQ had been evaluated by the university examination vetting committee. The committee had ensured that the time-duration stipulated to answer each component was fair and sufficient time was given for the student to use the WM and LTM. There was also a need to ensure that there was no cognitive overload as this would have severely affected the performance (Sweller, 1994; Kun, Steedle, Shavelson, Alzonzzo & Oppezzo, 2006). This probably did not occur in this examination as the mean score of MEQ was 12.1 (SD 1.2) with a DSBT mean of 6.6 with a narrow dispersion in both variables. Kun *et al.* (2006) in reviewing fluid intelligence and WMC had discussed the value of WMC overload and the preparation of instructional materials to improve WM function in teaching science. The findings would be applicable here.

The design of the MEQ appeared ideal to test WM and WMC as there was a need to read the question, analyse the problem, synthesize and organize the answer (maintenance and rehearsal of information) and write them down using previously acquired knowledge and skills (LTM) within a stipulated time.

It would not be wrong to infer that digit span exceeding six (6) should be sufficient to adequately answer MEQs due to efficient chunking at information processing (IP) that may have taken place over the years of study in medicine. A confounding factor that had not been addressed in the 'Methods' in administration of the MEQ was the 'unfair advantage' that the subjects may have had when answering the questions. They could use the question paper to write the salient points derived from the question to work on and organize the responses (answers) that were to be written. This activity could draw on the visio-spatial slave system of the WMC and also LTM (for application of prior knowledge). Such a function is not the basis of most tests of WMC whether it is word span or digit span where information attended to in the WM is maintained and rehearsed in the presence of interference. However, it can also be argued that more complex cognitive functions come into play in answering the MEQ (as a test of higher order thinking) as the elements of clinical reasoning and experiential learning are drawn into IP.

In a high stake examination like final year MBBS, the MEQ is designed to be fair and relevant to future clinical practice. It has to be realistically designed so the average student would be able to pass the examination. It has to reflect on all three domains the graduate should be competent in i.e. affective, psychomotor and cognitive. As was derived from the results of the study, the skewed distribution of DSBT with a mean of 6.6 (the majority had scores exceeding 7) appears representative of mature university students' WMC. While the mean score at the MEQ of the whole cohort was 12.1 (±, 1.20), interestingly the overall score marks for this cohort was 62.5 (± 4.7) which was not very different from the MEQ scores. Both these examination marks allude to the standard required for exhibiting competency to be around just above 60 per cent. The researchers were comfortable with not seeing a significant difference between DSBT and MEQ scores (p=0.92) nor a linear relationship between these variables, as the purpose of the evaluation of final MBBS was to ensure set standards were met and clinical competency for final MBBS students had been achieved. One should consider that a two-year internship follows after this evaluation at final MBBS in Malaysia and the outcome based assessment is a reflection of 'preparedness for internship' after graduation rather than an 'exit' professional examination.

On the other hand, how is that those who were in the low score on DSB were also able to score a mean of about 12.1? This could be addressed by test administration and consideration of personal factors of the subjects concerned. As this was a voluntary exercise and anonymity was assured, the respondents who scored low in DSB may not have been equally serious about attempting to complete the test with the same vigour as the others. It was not possible to prove this notion; however, at best those who scored less than 4 should be treated as outliers.

The study assists one to apply the theoretical underpinnings of the WMC to reflect on the design of the MEQ. Although individual differences in WM have been demonstrated for language comprehension and reasoning (Hansen, 1995) this may not be applicable when a mean DSBT for final year medical students (6.6) in a relatively homogenous group are being evaluated.

The mean score for MEQ of 12.1 (of 20 marks) drives home the point that when final year medical students are assessed, they have had a training that involved a myriad of instructional and experiential experiences which are more realistic than that used in pilot training using a flight simulator or a graduate programme in computer science and physics. The spiral outcome based curriculum which is spread over five years of training culminating in an 'apprenticeship' model in the clinical years of the programme presents unique learning experiences that prepares him/her to attain the required standards of competence in tests at theory and clinical settings (Harden, 1999). The continuous exposure to clinical disease and need for developing clinical diagnosis on day-to-day learning based on disease pattern permits pattern recognition and appropriate encoding in WM and LTM. This forms a broad knowledge-based schemata that, if organized in a meaningful way, becomes easy for subsequent retrieval when a task demands for its use.

To enable the medical student to perform complex cognitive tasks required of MEQ, mention has been made of both cognitive and experiential learning impacting on successful execution of the tasks. In order to integrate coherently, information and data derived from the stems of the MEQ contextual information would be required. As mentioned above, this is acquired over a period of learning in the medical school. The student appears now to process the new information derived from the questions and while maintaining them, has to draw on contextual information for both appropriateness and applicability. Herein lies the value of construct validity of the MEQ. As is noted under limitations of the study, the researcher did not do any factor analysis of the MEQ administered to the cohort of students studied to detect the percentage of questions that tested lower order cognitive functions (i.e. recall questions) that would effectively evaluate LTM. On the other hand, if questions are poorly constructed, cueing effect may lessen the 'burden' on working memory functions.

Although the MEQ was utilized to explain the possible factors that could confound the inferences from the study, cognitive scientists feel that the model proposed by Baddeley & Hitch (1974) would not be able to explain all the issues related to IP. Ericsson & Kintsch (2000) had reviewed the WM model and suggested alternatives to simply explaining WM as a sequence of stable states of 'end products' processing. Others had alluded to a seamless IP model (Ormrod, 2008).

Tasker & Dalton (2006) in researching instructional materials for teaching chemistry emphasized the value of animations to improve performance at understanding chemistry based on models developed by Johnstone & El-Banna (1986), cognitive load theory and the IP model described above. They inferred the importance of three dimensional models to assist learners to move seamlessly through the functional components of WM and LTM. Using the classroom research results by Tasker and Dalton (2006), medical education presents students to the real world of health and disease from the onset. The value of learning in the real environment is realized as it enables them to acquire both skills and knowledge so as utilize the functional units of both WM and LTM in a seamless manner.

Whether training in medical sciences through integration with the real world introduces a higher level of cognition (mental activity and metacogniton) that permits larger chunks to be encoded which can be retrieved at a faster rate, needs to be evaluated. Experiential learning in medicine invokes relevance to future practice and learning can be through automation when similar principles of clinical management are taught and utilized through problem based learning, a teaching activity that is in vogue in the researchers' university. This would enable the skilled learner to easily access desired information from LTM increasing retrieval rates and making WM function effective to complete the task. Retrieval cues contribute to ease of access and this would come into play with the type of training given to medical students. It is clear that memory performance increases after specific learning tasks and familiarity with both material and information (Ericsson, 1985). This would explain Miller's chunking theory where familiarity and experience with a particular stimuli over time would lead to a set of complex patterns in the LTM. Hence when a retrieval cue is presented as the student analyses the problem in the MEQ, he would have an efficient system of retrieving acquired 'pattern' or chunks. This proposal supports the argument that learning in medical sciences using a myriad of teaching–learning activities (which also involves a large portion of experience-based learning) perhaps has attuned the final year medical student to adopt more efficient retrieval systems in the WM without having to suffer from WMC overload. The results of the current study support this notion as there was very little variation in the MEQ score (Mean =12.1 ±.2 marks) and the DSBT ranged from 6-8 in the majority of respondents. As the MEQ was based on previous experience and learned information, perhaps the WMC was adequate and students did not have to suffer from effects of WMC overload.

Proactive interference has been mentioned as affecting retrieval speed (Underwood & Ekstrand, 1967). In the context of answering the MEQ, the questions were administered based on six major disciplines in medicine. Each component of the question in each discipline was based on one theme before it moved to another sequentially. Such a process appeared not to work the WM maximally and the student was able to focus on a separate task at any one point in time. Clearly no interference was anticipated as each task in the stem of each question was focused on a particular competency to be tested. This could be another reason why the DSBT did not show any significant association with MEQ scores (Patel & Groen, 1991) .

Proponents of the IP model of Atkinson and Shiffrin (1971) maintain the view that information that is attended to would be processed through the model as proposed by Baddeley & Hitch (1974). This model has been challenged by many who question if working memory and long term memory are really different and if conscious thought processes are a pre-requisite for encoding in LTM. In fact some of these views are applicable to medical training. Although WM is an active and conscious mechanism, some kinds of information may be automatically stored in LTM (Omrod , Frensch & Runger, 2003; Unsworth & Engle (2005); Unsworth & Engle, 2007a, Unsworth & Engle, 2007b). This non-conscious knowledge very often occurs in the apprenticeship model that is used in the clinical training of the medical students. This approach to explain why medical education may be different compared to conventional teaching needs consideration when one tries to derive inferences from tests of WM and WMC.

Is the DSBT a robust test to evaluate WMC? The WM is required to actively maintain information in the presence of on-going interference or distraction. Baddeley and Hitch's model places emphasis on the domain 'general executive attention'. When administering the DSBT, one is expected to involve the individual's ability to address the task of receiving the digits read out (auditory-phonological loop) and to be able to write the digits in reverse. The latter requires the maintenance of information and to 'work' on it so as to able to write them down in a reversed fashion inculcating the visuo-spatial slave system against other interferences like environment where the study is conducted. Background 'noise', the anxiety of submitting oneself to go through a test that has no direct relevance to their daily work and the rapid progression to another set of digits after attempting one digit span test can impact negatively on volunteers who consent to the study but were not cooperative in completing the task. This may account for the few outliers in this study. Johnstone's DSBT has a preliminary test incorporated to 'pace the mind' before the formal test to overcome some of such effects.

However, some of the nuances mentioned above could not be totally excluded in confounding the final results. The researcher could not control for such confounders. The basic principle of a WM span tasks was to engage executive attention processes whether one uses reading, operation or counting span (Oberauer, Heinz-Martin-Sub., Wilhelm & Wittman, 2000; Oberauer, Heinz-Martin-Sub, Wilhelmo & Wittman, 2003). A critical component that needs to be borne in mind is interference with rehearsal. It is well known that if there is a significant delay between reading the digit span (in the administration of the command), the test would deviate from meeting its objectives and may measure short term memory storage (Conway, Kane, Bunting, Hamrick & Engle, 2005). The test administration adhered closely to the recommendations made by Johnestone & El-Banna (1986) in the current study. One of the concerns about digit span tests is whether it evaluates only the storage component and not the control function. Oberauer, Heinz-Martin-Sub, Wilhelmo & Wittman (2000) attested to the DSBT to measure both storage and transformation.

A large volume of data, now available with regards to cognitive ability of the medical expert does not explain the way clinical diagnosis and treatment strategies are suggested based on proposed principles of WMC and IP. Basic medical facts can be encoded in the conventional way but both clinical experience and clinical procedural competency permit more complex cognitive sub-sumption that would enable both top-down and bottom-up reasoning strategies (Patel & Groen, 1991; Schmidt & Boshuzien, 1993). As has been alluded to the above with regards to IP and chunking among medical graduates, medical experts (specialists) and medical students who are no longer novices (pre-clinical year medical students) appear to store information in memory in an orderly manner as initially presented. When tasks are presented in a scrambled manner, they are able to re-analyze and categorize in manner that would permit meaningful diagnosis of the clinical condition.

Although concepts of the WM on medical expertise has not been extensively researched, it appears that clinical training and clinical experience impact on the way information is encoded. Retrieval of information in the WM may not follow the sequential end-of task processing. Automation and retrieval cues may suggest increased speed of retrieval due to more efficient chunking and effective sub-sumption (Ericsson, 1985; Patel & Groen, 1991; Groen & Patel, 1988). More data would be required to explain the cognitive differences between medical education and other disciplines. However, these findings on medical expertise and IP confirms our contention that when evaluating final year medical students using summative examinations like the Final MBBS examinations, the aim is to achieve academic standards based on conventional methods like the borderline regression method rather than absolute scores. Moreover, the use of several assessment tools apart from theory examinations is indicated.

Patel, Groen & Frederiksen (2009), Brailovsky, Charlin, Beausoleil, Cote & Van der Vleuten (2001), Norman & Schidmt (2000) elaborated eloquently the differences between medical students and doctors in memory for clinical cases and also the training medical students go through that impacted on their IP. Although there are distinct differences in the way how relevant information is isolated by medical experts compared to medical students in higher level of training, they follow a similar pattern in using a highly developed knowledge (cognition) base.

The expert physician reasons in a more efficient way in representing information in working memory. In the LTM knowledge acquired by experts is ordered as 'rules', specific for each action. A combination of both process oriented cognition and normative oriented decision making is clearly discriminated. Patel, Groen & Frederiksen (2009) concluded that although there were few differences in recall (LTM) between novice and expert, those who were in higher level of medical training (residents and senior medical students-intermediates) use more information in WM that is irrelevant. This U-shaped findings (with novice and experts not using more basic science knowledge for clinical reasoning, but intermediates using more of such) alludes to significant elaboration of information that goes on in WM of the intermediates. Expert physicians tend to skip the fundamentals of basic sciences in decision making, hence accounting for a more efficient and rapid retrieval ability. This could be due to 'early hypothesis assumption' and clinical experience.

It becomes more clear now that there is a difference in chunking between experts and novices in the medical discipline; a causal network is evident in disease pattern recognition (representations and encoding in WM and LTM) unlike in other vocations. Clinical training enables chunking to occur as rules rather than informal 'textbook' knowledge with both forward and backward reasoning. The 'expert physician' appears to use forward hypothetical-deductive approaches without having to retrieve basic science knowledge because of pattern recognition and utilization of pathophysiology, an ability acquired from experience and skill acquisition. This would allude to more effective chunking and encoding (Patel & Groen, 1986; Groen & Patel, 1988) and subscribes to meaningful learning.

Medical students in their senior years of training, who have limited clinical experience but have developed similar diagnostic reasoning skills, may be using a more 'inefficient' backward and forward (multidirectional) reasoning in a seamless manner between the components of the IP model, contributing to the notion that a more pliable or fluid model of Baddeley and Hitch (1974) may be necessary to explain the IP of medical students. This explanation is also in line with the thoughts expressed by Ericsson, Krampe & Tecsh-Romer (1993), Omrod, Frensch, & Runger (2003), Arocha & Patel (1995), Omrod (2008) and Patel& Groen (1986) that experts acquire a set of retrieval cues at time of encoding that is associated with function, that which is relevant to clinical practice. This enables the expert physician rapid access to LTM relevant data when problem solving occurs in WM. The efficient 'chunking' also allows IP in WM to function without WMC overload.

The medical students in this study achieved academic marks which were above the pass mark of 50 percent and the dispersion of marks around the mean for the MEQ was narrow (12.1 ± 1.2). Clinical experience and a large knowledge base impacted on WM elaboration and efficient clinical reasoning abilities, deduced by the clear, but not high pass mark (>15/20 marks) for the cohort. This fact points to final year medical students adopting both forward and backward reasoning that did not reflect as efficient, leading to delay in retrieval related to increased elaboration of irrelevant information in the WM. Although the WMC was presumably not overloaded, they lacked experience like the experts (specialists) and 'settled' to achieve academic standards that enabled them to pass the examination comfortably at a mean of about 60%. Only one student attained the distinction mark (70%).

The results of the study permitted the researcher to make a few observations that may be relevant in cognitive psychology. There were differences in cognitive function especially in IP among final year medical students in this cohort. The mean score of the WMC of 6.6 was adequate to achieve pass marks in end of course assessment with a mean of about 60 % (mean overall score 62% and mean MEQ score 12.1/20 marks).

The learning experiences through the years in medical schools using varying teaching methods, including clinical experiential learning in an authentic practice environment enabled most subjects to pass the MEQ meeting the standards set by the examination committee. Unlike medical experts (specialists), they had to take a broader field in their assessment tests (six different disciplines). The findings are in line with others who had reported a difference in the way the novice medical student and intermediate (final year medical students and residents) use retrieval cues and encoding (Patel, Groen & Fredericksen, 2009; Ericsson, Krampe & Tecsh-Romer, 1993).

Conclusions

Final year medical students had a mean score of 6.6/8.0 on the DSB. WMC (as determined by the DSBT) showed narrow dispersion about the mean indicating a rather homogenous population. There was no significant difference between academic achievement and WMC. Clinical experience and experiential learning permitted medical students in this study to use the novice-intermediate-expert model to solve clinical problems through 'set of rules' and pattern recognition given clinical scenarios in MEQ.

The WMC as measured by the DSBT appears to average 6.6/8.0. Clearly the working memory span is adequate to be successful at the MEQ, the majority attaining a mean of 21.1/20.0. This finding could be interpreted as sufficient for the WM to 'process and rehearse' information derived from the MEQ without WMC overload.

This study supports the proposal made by Patel & Groen (1991) that medical education that incorporates a myriad of teaching learning activities including experiential learning must take cognizance of the differing ways cognitive functions operate in problem solving, adopting the novice intermediate-expertise paradigm. The final year medical student probably retrieved information from schemata in a back and forth fashion which was not as efficient as the expert physician. The current observation that a majority of the final year medical students attain a common average of

12.1/20 marks in the MEQ indicate this fact. WM cognition appears to be more complicated in medical students than those in other disciplines of higher education.

Limitations

The finding of this study is applicable to only final year medical students who appeared to have attained a set standard to pass the MEQ paper. This finding should not be extrapolated to other cohorts or those in different disciplines. Working memory capacity using more complex psychometric tests could be done at different stages of medical education to see if there is change in WMC as students advance through medicals school based on differing ways of encoding and retrieval (based on disease pattern recognition) for problem solving in the medical field.

Clinical clerkship assessment in semester 10 and internship years should analyzed to determine if novices, intermediate and experts in the local population from the medical fraternity exhibit different problem solving approaches in the local context.

Aknowledgement

I wish thank Dr K Y Loh, Family Physician at the International Medical University for statistical assistance and my deep gratitude to Dr Nagalingam Karuppiah of the Ministry of Education for kindly overseeing this project.

REFERENCES

1. Arocha JF & Patel VL. Novice diagnostic reasoning in medicine: Accounting for evidence. The Journal of Learning Sciences, 1995; 4(4): 355-358.
2. Atkinson R. & Shiffrin RM. Human memory: A proposed system and its control processes. In K.W. Spence and J.T. Spence (eds). The psychology of learning and motivation: Advances in research and theory, 1968; Vol. 12, pp 772-775, New York Academic Press.
3. Atkinson RC & Shiffrin RM, The control of short term memory. Scientific American, 1971; 225: 82-90.
4. Baddeley AD & Hitch G. Short term forgetting in absence of proactive interference. Quarterly Journal of Experimental Psychology 1971; 3(3):275-283.
5. Baddeley AD.. & Hitch G. Working memory. In G.A. Bower (Ed.), The psychology of learning and motivation, 1974; Vol.8, pp.47-89, New York: Academic Press.
6. Baddeley AD. Working memory. Oxford: Oxford University Press, 1986.
7. Baddeley AD. The episodic buffer: a new component of working memory? Trends in Cognitive Sciences., 2000;4: 417-423.
8. Bloom BS. Taxonomy of Educational Objectives Handbook I: The Cognitive Domain. New York: David McKay Co. Inc., 1956.
9. Brailovsky C, Charlin B, Beausoleil S, Cote S, & Van der Vleuten, C. Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: an experimental study on script concordance test. Medical Education 2001;, 35; 430-436.
10. Bunting MF, Conway ARA & Heitz RP. Individual differences in fan effect and working memory capacity. Journal of Memory and Language, 2004: 51: 6040-6062.
11. Cantor J & Engel RW. Working memory as long term memory activation: an individual's different approach. Journal of Experimental Psychology, Learning , Memory and Cognition, 1993; 19: 1110-1114.
12. Cassidy S. Learning styles: An overview of theories, models and measures. Educational Psychology, 2001; 24(4): 419-444.
13. Conway ARA, Kane MJ, Bunting MF Hamrick DZ & Engle RW. Working memory span tasks: a methodological review and users' guide. Psychonomic Bulletin Review, 2005; 12(5): 769-786.
14. Daneman M & Carpenter PA. Individual differences in working memory and reading, Journal of Verbal Learning and Verbal Behaviour, 1980; 19: 450-466.
15. Engle RW & Kane MJ. Executive attention, working memory and two factor theory of cognitive control: In B. Ross (Ed.), The Psychology of Learning and Motivation, 2004; Vol 44 ,pp. 145-199, New York: Elservier.
16. Epstein RM. Assessment in Medical Education. New England Journal of Medicine 2007; 356: 387-396 .
17. Ericsson KA. Memory skill. Canadian Journal of Psychology, 1985; 39: 188-231.
18. Ericsson KA, Kintsch W. Long term working memory. Psychological Review, 1995; 102: 221-245.
19. Ericsson KA, Krampe R & Tecsh-Romer C. The role of deliberate practice in acquisition of expert performance. Psychological Review,1993; 100(3): 363-406.
20. Ericsson KA & Kintsch W. Shortcomings of generic retrieval structures with slots of the type that Gobet (1993) proposed and modeled. British Journal of Psychology, 2000; 91: 571-588.
21. Feletti GI & Smith EK. Modified questions: are they worth the effort? Medical Education, 1986; 20; 126-132.
22. Groen GJ & Patel VL. The relationship between comprehension and reasoning in medical expertise. In M.T.H.Chi, R. Glaser, & M.J. Farr (Eds.), The nature of expertise, 1998; pp.287-310. Hillsdale, N.J: Erlbaum.
23. Harden RM. What is a spiral curriculum? Medical Education, 1999; 21(2): 141-143.
24. Hartley J. Learning and Studying: A research perspective. London: Routledge.

25. Hansen, J.W. (1995). Student cognition styles in postsecondary technology programs. Journal of Technology Education, 1998; 692: 23-28.
26. Hitch G. The role of short term memory in mental arithmetic. Cognitive Psychology. 1978; 10: 302-323.
27. Howe A, Campion P, Searle J & Smith H. Learning in practice: new perspective-approaches to medical education at four new medical schools. British Medical Journal, 2004; 329: 327-331.
28. Hunt E. What is a theory of thought. In (Ed) . Robert J. Sternberg: The Nature of Cognition. MIT Press, Cambridge, Massachusetts, 1999.
29. Irwin WG & Bamber JH. The cognitive structure of the modified essay question. Medical Education, 1982; 16: 326-331.
30. Johnstone AH & El-Banna H. Capabilities, demands and processes: a predictive model for science education. Educational Chemistry, 1986; 23: 80-84.
31. Johnstone AH & El-Banna H. Understanding learning difficulties : a predictive model. Studies in Higher Education, 1989; 14: 159-68.
32. Johnstone AH. Digit Span Backwards Test, Centre for Science Education, University of Glasgow, United Kingdom, 2001.
33. Knox J. The Modified Essay Question Booklet No.5 .Dundee; Allied Health Medical Education (ASME), 1975.
34. Kun Y, Steedle J, Shavelson R, Alzonzo A & Oppezzo M. Working memory, fluid intelligence, and science learning. Educational ResearchRreview, 2006; 4: 83-98.
35. May CP, Hasher L & Kane MJ. Role of interference in memory span, Memory and Cognition, 1999; 27(5): 759-767.
36. Noble KA, Miller SM & Heckman J (2006). The cognitive style of nursing: educational implications for teaching and learning. J Nurs Educ 2008; 47(6): 245-253.
37. Norman GP & Schidmt HG. Effectiveness of the problem based curricula ; theory, practice and paper darts. Medical Education, 2000; 34: 721-728.
38. Oberauer K, Heinz-Martin-Sub, Wilhelmo O & Wittman W.W. Working memory capacity-facets of cognitive ability construct. Personality and Individual Differences, 2000; 29: 1017-1045.
39. Oberauer K, Heinz-Martin-Sub, WilhelmO & Wittman MW. The multiple faces of working memory: storage, processing, supervision, and coordination. Intelligence, 2003; 31: 167-193.
40. OmrodJL, Frensch PA & Runger D. Implicit learning. Current Directions in Psychological Science, 2003; 12: 13-18.
41. Opdenackera C, Fierensa H,. Van Brabanta H, Sevenantsa J, Spruyta J, Slootmaekersa PJ & Johnstone AH. Academic performance in solving chemistry problem related to student memory capacity . International Journal of Science Education, 1990; 12 (2): 177 – 185.
42. Ormrod JE. Introduction to Cognition and Memory. In. J.E. Ormrod, (Ed.), In Human Learning , 2008; pp.149-193 New Jersey: Pearson Prentice Hill.
43. Palmer JP & Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? BMC Medical Education ,2007; 7: 49.
44. Patel VL & Groen GJ. Knowledge-based solution strategies in medical reasoning. Cognitive Science, 1986; 10: 91-116.
45. Patel V.L & Groen, GJ. The general and specific nature of medical expertise : a critical look. In K.A. Ericsson & J. Smith (Eds.) , Towards a general theory of expertise 1991; pp.93-125. Cambridge, M.A.: Cambridge University Press.
46. Patel VL, Groen GJ & Frederiksen C.H. Differences between medical students and doctors in memory for clinical cases. Medical Education, 2009; 20 (1): 3-9.
47. Pascual-Leone J. Organismic processes of neo-Piagetian theories: a dialectical causal account of cognitive development. International Journal of Psychology, 1987; 22: 531-570.
48. Schmidt HG & Boshuizen H.A. On the origin of intermediate effects in clinical recall. Memory & Cognition, 1993; 21: 338-351.
49. Shute V. Who is to acquire programming skills? Journal of Educational Computing Research. 1991; 7: 1-24.
50. Sweller J. Cognitive overload theory, learning differences and instructional design. Learning and Instruction, 1994; 4: 295-312.
51. Tasker R & Dalton R. Research into practice: visualization of the molecular world using animations. Chemistry Education Research, 2006; 7(2): 141-159.
52. Underwood BJ & Ekstrand BR. Effect of distributed practice on paired associated learning. J Expt Psychol 1967; 73: 1-21.
53. Unsworth N & Engle RW. Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. Memory & Cognition, 2005; 33: 213-220.
54. Unsworth N & Engle RW. Individual differences in working memory capacity and retrieval: a cue dependent search approach. In J.S.Nairne (Ed.), Foundations of Remembering, 2007a; pp,241-258. New York: Psychology Press .
55. Unsworth N & Engle RW (2007b). On the decision of short term and working memory: an examination of simple and complex span test and their relation to higher order abilities. Psychological Bulletin, 2007b; 133(6): 1038-1066.