ABSTRACT
Successful language learning depends crucially on a strong vocabulary. It is estimated that a vocabulary of 8000 word families is the necessary goal for English language learners who have to deal with unsimplified spoken and written texts. This study examined the English vocabulary size and vocabulary learning strategies of Mengubah Destini Anak Bangsa (MDAB) pre-diploma students at the Segamat Campus of Universiti Teknologi MARA (UiTM) Johor. While some of these students had intermediate English language proficiency, the majority of them had low language proficiency. Data for the study were collected in three stages. Firstly, a Vocabulary Size Test was conducted to determine the students’ vocabulary size. Secondly, a Vocabulary Learning Strategies Questionnaire was used to investigate the vocabulary learning strategies employed by the students. Finally, a semi-structured interview was carried out with eight students as a form of triangulation that supplements the findings gleaned from the Vocabulary Learning Strategies Questionnaire. The quantitative data from the first two instruments were analysed using descriptive statistics, while interview analysis was performed to analyse the qualitative data from the semi-structured interview. The findings revealed that only a very small number of the MDAB students had the desired vocabulary size of 8000 word families or more, and that they were medium users of vocabulary learning strategies. The findings on their vocabulary size suggest that many of them may not be lexically well equipped to successfully undertake a diploma programme of their choice at UiTM which uses English as its medium of instruction. In addition, the results of this study imply that some systematic vocabulary instructions and vocabulary learning strategies training should be introduced to help them improve their English vocabulary size.

Keywords: Intermediate language proficiency, low language proficiency, vocabulary, vocabulary learning strategies, vocabulary size

Introduction
English is arguably one of the most important languages in today’s era of Information and Communication Technology. As such, the continuing trend of declining English language proficiency among the Malaysian youths is worrisome (Murugesan 2003; Ling 2015). Teachers complain that the majority of the students are not able to master basic grammar, could not easily write 300-word essays and have to resort to colloquial speech when speaking the language (Ling 2015). In addition, lack of proficiency in English is one of the major reasons for jobless graduates in the
country (Teoh 2011). This is especially so when it comes to the bumiputera youths. Many of them have been unable to get into the private sector especially in multinational companies where English is the medium for communication (Gill 2006 & Puteh 2010, cited in Ha et al. 2013). As a matter of fact, more than 90% of the companies surveyed in 2009 acknowledged that English was their official business language (Teoh 2011).

This decline in the standard of English among the youths in the country may be due to several factors and one of them could be that they may have not been using the right language learning strategies which are “one of the most important individual difference factors in L2 acquisition” (Skehan 1989, cited in McMullen 2009: 419). For example, the use of language learning strategies has been found to facilitate second language acquisition, improve student performance and promote greater learner autonomy (McMullen 2009).

By extension, vocabulary learning strategies (henceforth VLS) are one of the important sub-categories of language learning strategies. In fact, it is one of the language learning strategies categories investigated in Cohen et al.’s Strategy Use Inventory (2005). This is because vocabulary learning, together with the learning of grammar, is one of the important aspects of language learning. Other than learning the four language skills of listening, speaking, reading and writing, a language learner needs also to learn the vocabulary and the grammar rules of the language. In fact, vocabulary knowledge is considered as one of the central aspects in second language learning (Ming 2007, cited in Zhang 2011). As such, the use (or the lack of use) of appropriate VLS might be one of the factors that differentiate the successful second language learners from the unsuccessful ones (Kalajahi & Pourshahian 2012).

However, although vocabulary learning has always been problematic for most students (Kalajahi & Poursahian 2012), the teaching of vocabulary has always been overshadowed by focus on reading and writing skills, and grammar (Bastanfar & Hashemi 2010). In fact, vocabulary instruction is often given little emphasis in the university curriculum in Asian countries (Fan 2003, cited in Kalajahi & Poursahian 2012). In addition, although research on both vocabulary and learner strategies had been widely investigated in the last forty years, until recently, research on VLS was rather limited (Schmitt 1997). Moreover, most of the research on VLS was conducted on successful language learners (Porte 1998). Finally, when it comes to research on vocabulary size, Nation (2012b: 1) asserted that it is “the worst researched area in applied linguistics”.

As such, this study was undertaken to investigate the vocabulary size and VLS of ESL learners enrolled in a special UiTM pre-diploma programme called Program Mengubah Destini Anak Bangsa.

MDAB Programme

Mengubah Destini Anak Bangsa (hereafter MDAB) or ‘Changing the Destiny of our People’ programme is a special pre-diploma programme under the patronage of the Malaysian Prime Minister that was introduced at selected UiTM campuses in June 2010. This programme was an aspiration of the former Vice Chancellor of UiTM, Tan Sri Dato’ Sri Prof Ir Dr. Sahol Hamid Abu Bakar who believes in giving underprivileged bumiputera students who did not perform well in the Sijil Pelajaran Malaysia (SPM) examination a second chance to continue their studies at tertiary level. In other words, this programme was established to give bumiputera students who come from low-income family and at the same time have poor SPM results another opportunity at continuing their studies at UiTM (Sahol Hamid 2014;
Universiti Teknologi MARA n.d.). This programme is deemed as part of UiTM’s Corporate Social Responsibility (CSR) towards the education of the bumiputra people in the country (Universiti Teknologi MARA n.d.). There are two types of programmes that are offered under the MDAB programme which are pre-commerce and pre-science programmes.

The MDAB programme is only open to bumiputra students aged 16 and above whose parents earn less than RM3000 a month (regardless their home location and the number of dependents). Students who are accepted into the programme receive some advantages not offered to other UiTM students. These include waived course fees, free accommodation on campus and a monthly allowance (Sahol Hamid 2014; Universiti Teknologi MARA n.d.).

At the Segamat campus of UiTM Johor (henceforth UiTMJS), the programme began the first semester it was introduced to the whole UiTM system which was in June 2010. However, only one programme is offered at this campus which is the pre-commerce programme. Although management and religious study courses are also taught to these pre-commerce MDAB students, the two core courses the pre-commerce students have to focus on are Mathematics and English. They have 12 hours of English lessons and 10 hours of Mathematics every week. This is to ensure that they have a good solid foundation on both courses before they embark on their diploma level at the many different faculties at UiTM.

However, with regards to their English language proficiency, as the entry requirement for the MDAB programme – for both pre-science and pre-commerce programmes – is just a pass at the SPM level, many have observed that the majority of the MDAB students have generally low English language proficiency (Kamisah & Norzie Diana 2012; Wan Norliza et al. 2012; Metom et al. 2013). Indeed, a study by Rosnani and Nurul Hasni (2012) has identified lack of English language proficiency as a major factor in the poor academic achievement of the MDAB students at UiTMJS.

Purpose of the Study

This study aimed to establish the English vocabulary size and VLS of a group of MDAB pre-diploma students at UiTMJS. The research questions that guided the study were as follows:
1. What is the English vocabulary size of the MDAB pre-diploma students?
2. What are the types and frequency of VLS used by the MDAB students based on their vocabulary size?
3. What are the MDAB students’ most frequently used and least frequently used VLS based on their vocabulary size?

Literature Review

The importance of vocabulary learning could be encapsulated by a quote from Wilkins (1972: 110, cited in Dinh 2008: 1): “without grammar very little can be conveyed, without vocabulary nothing can be conveyed”. This suggests that “words are the building blocks of a successful communication” (Bastanfar & Hashemi 2010: 158). Furthermore, vocabulary learning is also important because in comparison to the grammar of a language that could largely be mastered by the time a child is ten years old, vocabulary learning continues for the rest of one’s life (Schmitt 2000). Moreover, other than for personal usage and communication, vocabulary learning is also important for academic study due to the fact that many standardised tests like TOEFL and IELTS incorporate tests of vocabulary knowledge (Dinh 2008).
Vocabulary Size of Language Learners

The term vocabulary size has been referred to as “the number of words a learner has in [his/her] mental lexicon” (Kalajahi & Pourshahian 2012: 141). Other than that, the term has also been defined as “the number of words of which the learner knows at least some significant aspects of the meaning” (Eyckmans 2004: 13). As such, a learner’s vocabulary size may be a good indicator as to whether he/she is a good reader or not because the number of words that he/she knows has a direct effect on his/her reading comprehension (Biemiller 2005, cited in Coxhead et al. 2015). Furthermore, vocabulary size has also been found to be strongly correlated with speaking, listening, grammar, and writing (Zimmerman 2004, cited in Engku Haliza et al. 2013).

It has been suggested that 98% text coverage is the minimal optimal level for most learners to gain adequate comprehension of a text (Nation 2006; Coxhead et al. 2015). To have a 98% text coverage means one only encounters two unknown or unfamiliar words for every 100 words found in a text (Nation 2006). Using this yardstick, for a language learner to be able to read unsimplified reading texts that may include novels, newspapers and textbooks, he/she has to have a written receptive vocabulary of around 8,000-9,000 word families (Nation 2006; Nation 2012). In addition, it has been found that this vocabulary size (8,000 to 9,000 word families) was also the vocabulary size possessed by non-native speakers of English who were doing advanced degrees in institutions where English is the medium of instruction (Nation 2006).

Research on Vocabulary Size

Ahmad Azman et al. (2009) conducted a study on the vocabulary size of 360 diploma students at UiTM Perlis. He used Vocabulary Levels Tests that consisted of three different vocabulary tests (1) The Passive Vocabulary Test adapted from Nation (1990), (2) The Controlled Active Vocabulary Test adapted from Laufer and Nation (1995) and (3) The Free Active Vocabulary Test also adapted from Laufer and Nation (1995). Setting the desired scores for the three tests as 83% or more, students who scored less than 83% were considered to have weak vocabulary. Their findings revealed that these diploma university students had very limited receptive and productive vocabulary size because they failed to achieve the desired 83% level in the Passive Vocabulary Test and the Controlled Active Vocabulary Test, and because they mostly used high-frequency words for the Free Active Vocabulary Test.

In another research, Kafipour et al. (2011) used a Vocabulary Levels Test by Nation (2001) to measure the vocabulary level of 238 Iranian junior EFL students from five Semnan universities. They found that although the students knew sufficient vocabulary in 2000 and 3000 word levels, they did not know a large number of words in 5000, 10000 and academic levels.

Another study was conducted by Engku Haliza et al. (2013) on the English vocabulary size of Muslim pre-university students enrolled in the intensive English language programme at the Centre for Languages and Pre-University Academic Programme, International Islamic University Malaysia (IIUM). These students had to enrol in the intensive language programme because they had failed to meet a minimum band 6 for IIUM’ English Proficiency Test (EPT), a minimum band 6 for IELTS or a minimal score of 550 for TOEFL. The purpose of their study was to investigate the students’ receptive and productive vocabulary at the 1,000, 2,000, 3,000, 5,000, University Word List (UWL) and 10,000 levels. The study found that,
based on 87% mastery level, these students still had not achieved the “recommended threshold level of vocabulary” (p. 44).

Vocabulary Learning Strategies

Rubin’s definition of language learning strategies was adapted by Schmitt in his definition of VLS: “any [operations, steps, plans and routines] which affect the broadly-defined process [of obtaining, storage, retrieval and use of words]” (1997: 6). In addition, adapting Oxford’s definition of language learning strategies, Kafipour et al. (2011: 65) defined VLS as “any techniques or tools which can be used to learn vocabulary quickly, easily and independently”.

It is widely accepted that it is impossible for teachers to teach all the words their students need to know in a classroom (Sokmen 1997, cited in Kalajahi & Pourshahian 2012; Schmitt 2010, cited in Gusti 2015). In fact, the students need to acquire most of the words on their own. This can be made more feasible if they are familiar of the many different existing VLS that are actually available for them when they want to learn new words and improve their vocabulary. Arguably, one way to make sure they are familiar of these strategies is by exposing them to strategy training in vocabulary learning (Sokmen 1997, cited in Kalajahi & Pourshahian 2012). However, this is not the case as many students’ exposure to a systematic vocabulary learning strategy is rather lacking (Engku Haliza et al. 2013).

Classification of Vocabulary Learning Strategies

VLS, as one of the sub-categories of language learning strategies, were one of the aspects investigated in Cohen et al.’s (2005) Language Strategy Use Inventory and Zamri et al.’s (2010) Language Learning Strategies Inventory. Cohen et al.’s (2005) Language Strategy Use Inventory covered six main language learning strategy categories of listening strategy use, vocabulary strategy use, speaking strategy use, reading strategy use, writing strategy use, and translation strategy use. Each of these main strategy categories was divided into sub-strategies. For vocabulary strategy use categories, the sub-categories were (1) strategies to learn new words, (2) strategies to review vocabulary, (3) strategies to recall vocabulary, and (4) strategies to make use of new vocabulary. Zamri et al.’s (2010) Language Learning Strategies Inventory, on the other hand, was adapted from Mohamed Amin’s (2000) and Zamri’s (2004) works. Their inventory investigated the language learning strategies for the four language skills of listening, speaking, reading and writing, and the three language aspects of vocabulary, grammar and comprehension. Each of these language skills and aspects were classified into classroom language learning strategies (CLLS), out-of-classroom language learning strategies (OLLS) and exam language learning strategies (ELLS). Meanwhile, Schmitt (1997) designed a taxonomy of VLS to solely scrutinise the VLS of 600 Japanese EFL learners. This taxonomy is discussed in the next section.

Schmitt’s Taxonomy of Vocabulary Learning Strategies

Schmitt’s (1997) taxonomy was modelled after Oxford’s (1990) Strategy Inventory for Language Learning (SILL). Schmitt categorised 58 VLS into two categories namely discovery and consolidation strategies. Discovery strategies involve strategies learners use to learn the meaning of new words, while consolidation strategies involve strategies learners use to remember the meaning of new words and to continuously use them in the long term (Schmitt 1997; Ahmad Iskandar 2008). The two sub-categories for the former are determination and social strategies, while the
four sub-categories for the latter are social, memory, cognitive and metacognitive strategies.

Determination strategies refer to the strategies where learners use their knowledge of the language, contextual clues or reference materials (Schmitt 1997). As such, the attempts to guess for meanings of unfamiliar words through analysis of parts of speech, the use of contextual clues and the use of dictionaries, thesauruses and word lists are examples of determination strategies. These strategies allow learners to discover the meanings of words without the help from another person (Ahmad Iskandar 2008). Social (discovery) strategies, on the other hand, refer to the strategies where learners interact with other people in order to discover the meanings of unfamiliar words (Schmitt 1997). Asking someone who knows the meaning of or the translation for an unfamiliar word are examples of social (discovery) strategies.

Like the social (discovery) strategies, social (consolidation) strategies also involve strategies that require learners to interact with other people. The difference is that while the social (discovery) strategies are employed to figure out the meaning of words that are encountered for the first time, the social (consolidation) strategies are utilised to strengthen the learners’ vocabulary knowledge (Schmitt 1997). Examples of social (consolidation) strategies include practising meanings of words in a group, asking teachers to check flash cards or word lists, and interacting with native speakers of English.

On the other hand, memory strategies involve strategies of storing and retrieving newly learnt words (Oxford 2001, cited in Tan et al. 2010) by relating them to previously learnt knowledge, by using some form of imagery or grouping (Schmitt 1997). Connecting words to their synonyms or antonyms, visualising the spelling of the words and grouping the words together to study them are among the examples of memory strategies.

In contrast, cognitive strategies comprise techniques and tools that can be manipulated in order to produce and understand words (Schmitt 1997). These techniques and tools include the usage of repetition and some mechanical means like word lists, flashcards and notebooks. Among the techniques and tools that fall under this sub-category are repeating words aloud or in written form, taking notes in class and keeping a vocabulary notebook.

Finally, metacognitive theory deals with what individuals know about their own thinking (Kellogg, 1994; cited in Gusti 2015). As such, it is always referred to as “thinking about thinking”. Oxford (2001, cited in Tan et al. 2010) considered metacognitive strategies as strategies that are consciously used to regulate language learning which Schmitt then surmised as strategies that are used to control and evaluate one’s own learning (Schmitt 1997). Thus, following these understanding of metacognitive theory, metacognitive VLS deal with strategies involve in planning, monitoring and evaluating one’s vocabulary learning through natural exposure and overview of the learning process in general (Oxford 1990; Schmitt 1997; Ahmad Azman et al. 2009). Watching movies, listening to songs and reading for pleasure are examples of naturally exposing oneself to more English words, while doing vocabulary exercises, and skipping or passing new words are examples of overviewing the learning process in general.

Research on Vocabulary Learning Strategies

Types and frequency. A study by Ahmad Iskandar (2008) on ESL students’ perception of their VLS involved 70 form two students from a school in Melaka. Descriptive statistics was used to analyse the level of practice of the VLS that the
students actually used when they encountered new words. It was found that the majority of the students fell under medium users.

Mohd Sahandri et al. (2009) conducted a study with 250 second year EFL students at Fars province in Iran, and reported that these Iranian EFL students were medium users of overall VLS and also of all the five sub-strategies of determination, social, memory, cognitive and metacognitive strategies. Similarly, a study by Kafipour et al. (2011) on 238 Iranian junior EFL students at Semnan universities found that they were medium VLS users for both the overall strategy and also for each of the five sub-categories.

However, a study conducted by Kalajahi and Poursahian (2012) on VLS of 125 undergraduate students at the Department of English Language Teaching of Education Faculty of Eastern Mediterranean University in North Cyprus found that the students were high users of overall VLS and also for both psycholinguistic and metacognitive strategies.

Finally, Lou’s (2014) study to investigate the VLS of 105 English-majors from five classes in the School of Foreign Studies at Yangtze University in China discovered that the students were high users of metacognitive strategies, and medium users of both cognitive and social/affective strategies.

Most and least used vocabulary learning strategies. Schmitt (1997) carried out a survey with 600 Japanese students who had taken or were taking English to study about the VLS that they frequently used. He found that the ten strategies most frequently used by these Japanese students (in descending order) were using a bilingual dictionary, using verbal and written repetitions, studying the spelling of words, guessing from textual context, asking classmates for meanings, saying new words aloud, taking notes in class, studying the sound of words and using word lists, while the five least used strategies were using physical actions, checking for and using L1 cognates, using semantic maps and asking teachers to check their flash cards for accuracy.

Mohd Sahandri et al. (2011) found that the eight most frequently used strategies among the second year EFL students at Fars province were using monolingual dictionary, guessing meaning from context, studying new words many times, connecting words to synonyms and antonyms, using new words in sentences, using English language media, taking notes or highlighting, and studying the sound of a word. Meanwhile, the eight least frequently used strategies were making lists of new words, using new words in a paragraph, studying the words with classmates, asking classmates for meanings of words, checking for L1 cognates, using physical actions when learning a word, talking with native speakers and asking teachers to check their definitions of words.

Zamri et al. (2010) studied the language learning strategies of successful English language learners in Malaysia. Their subjects were 637 form four students who had scored A in their PMR English for the year 2007. They used a survey questionnaire that was adapted from Mohamed Amin (2000), Zamri (2004) and Zamri et al. (2006) which categorised the language learning strategies into in-class (CLLS), out-of-class (OLLS) and for-examination strategies (ELLS). VLS, as a sub-strategy for language learning strategies, were one of the focuses of their study. They found the strategy with the highest mean for CLLS VLS was “I ask my friends when I do not understand the meaning of a word”; and the strategy with the lowest mean was “I do a revision on vocabulary before class”. For OLLS VLS, the strategy with the highest mean was “I ask my friends when I do not understand the meaning of a word”, and the strategies with the lowest mean were “I plan the time for vocabulary exercises.
ENGLISH VOCABULARY SIZE AND VOCABULARY LEARNING

every week” and “I buy vocabulary exercise books so that I can do vocabulary exercises”. Finally, with regards to ELLS VLS, the strategy with the highest mean was “I buy revision books so that I can do the vocabulary exercises before the examination”.

Methodology
This study is a case study that employed a mixed methods research design that comprised a combination of quantitative and qualitative data. It is a case study because it involved students from one single group only. In fact, this study involved only a small number of MDAB pre-diploma students from one UiTM branch campus only. As a form of research design, case study is extensively used in many areas and fields such as education, sociology, management, law and medicine (Zaidah 2007).

The type of mixed method design used for this study was triangulation where quantitative data and qualitative data were simultaneously collected and then merged in order to understand the research problem (Creswell 2005). With triangulation, a research problem is investigated from multiple data collection and analysis which allows for the confirmation of the accuracy of the findings of a study (Neuman 2006; Zaidah 2007). One rationale for triangulation design is that it combines the best of both quantitative and qualitative research methods where the former makes it possible to gather data from a large number of people and the latter enables an in-depth exploration of a small number of individuals (Creswell 2005). In addition, the use of triangulation design can also increase the credibility of a study (Zaidah 2007; Hussein 2009) as “one data collection form supplies strengths to offset the weaknesses of the other form” (Creswell 2005: 514). For this study, triangulation was used for confirmatory purposes that were to confirm the research results and conclusion (Hussein 2009).

Respondents
The respondents of this study were all students who were enrolled in the pre-diploma MDAB programme at UiTMJS for Session 2 2013/2014 semester. In total, there were 36 students enrolled in the programme for the semester. However, on the prearranged date, only 31 students came and completed the Vocabulary Size Test and the Vocabulary Learning Strategy Questionnaire. Of the 31, 10 were males (32.3%) and 21 were females (67.7%). These students were between 18 to 21 years of age. For qualitative data collection, only eight MDAB students were interviewed as they were the only respondents who agreed to take part in the semi-structured interview.

Research Instruments
**Vocabulary size test.** The first instrument for this study, the Vocabulary Size Test (henceforth VST), was developed by Nation and Beglar (2007) to measure both first language and second language learners’ total written receptive vocabulary size in English. The term written receptive vocabulary had been defined as vocabulary knowledge required for reading (Nation & Beglar 2007; Nation 2012). It has been claimed that written receptive vocabulary size or reading vocabulary size is generally considered as the broadest vocabulary size a second language learner possesses because learning words receptively, for example when reading, is easier than learning to produce them at suitable times (Nation 1990).

The VST is available in monolingual and bilingual versions. There are currently three versions of the VST available: the 14,000 version and the 20,000 versions A and B (Nation 2012). The online adaptations of these three versions had been made
available recently on vocabularysize.com which was developed in association with Victoria University of Wellington, New Zealand.

The VST version used for this study was the 14,000 monolingual version. It was used to estimate the MDAB pre-diploma students’ total written receptive vocabulary size or their total reading vocabulary size. The 14,000 version of the test was chosen instead of one of the 20,000 versions because the English language proficiency of the majority of the respondents of the study was gauged to be low. This appraisal was made based on the entry requirement into the MDAB programme (just a pass for English at SPM level) and observations made by the lecturers who have taught the MDAB students for a few semesters (Kamisah & Norzie Diana 2012; Wan Norliza et al. 2012; Metom et al. 2013). As such, the use of the 14,000 version was considered adequate as it already encompassed frequency levels beyond their likely vocabulary size (Nation 2012).

The VST is different from Schmitt’s Vocabulary Levels Test (Schmitt, Schmitt & Clapham 2001) which had been used to measure vocabulary size in some studies like Kafipour et al. (2011) and Kalajahi and Poursahian (2012). While Nation and Beglar’s (2007) VST was specifically designed to measure the total written receptive vocabulary size in English, Schmitt’s Vocabulary Levels Test was designed to measure vocabulary size in different levels such as the 1000-word, 2000-word, 3000-word and academic levels. As such, Schmitt’s Vocabulary Levels Test was not selected as the research instrument to determine the respondents’ vocabulary size because it could not provide the respondents’ total vocabulary size that is essential for the purpose of this study. One study that had used the 14,000 version of VST was a study by Mohd Sahandri et al. (2009) that investigated the vocabulary size of Iranian undergraduate EFL students.

**Vocabulary learning strategies questionnaire.** The second instrument used for data collection purposes was a Vocabulary Learning Strategies Questionnaire (hereafter VLSQ) which was adapted from Schmitt (1997) who, in part, adapted it from Oxford’s Strategy Inventory for Language Learning (SILL) (1990). Schmitt’s VLSQ had been adapted by many researchers such as Kudo (1999), Dinh (2008), Riankamol (2008), Mohd Sahandri et al. (2009), Kafipour et al. (2011), and Harun and Zawawi (2014).

The five-point Likert scale VLSQ was divided into two sections. Section I consists of 20 strategies for the discovery of the meaning of a new English word, and Section II consists of 47 strategies for consolidating a word once it has been encountered. The data from the VLSQ provided information about the type of VLS that the participants use and the frequency for which they are used. To measure the frequency of use, the level of frequency rating as suggested by Oxford (1990) was used.

The VLSQ was also translated into Bahasa Malaysia. In addition, some examples and explanations were also given for some items. That the VLSQ was translated into Bahasa Malaysia was consistent with the fact that many other questionnaires on language learning strategies and VLS had been translated into other languages (Zamri 2004, Kamarul Shukri 2009; Harun & Zawawi 2014; Lou 2014). Translation of the VLSQ was needed to ensure that any misunderstanding or misinterpretation could be avoided (Gu 2010). In other words, the translation allowed the MDAB students to understand every single aspect of the VLSQ well. This in return warranted that their responses would be as truthful as they could possibly be. Besides, examples and explanations were given following the steps taken by Kamarul Shukri (2009), and Harun and Zawawi (2014) who gave some examples and
explanations to their respondents of Arabic learners in order to ensure that they could understand each every item clearly. Giving such examples and explanations were necessary due to the fact that the majority of the participants had low English language proficiency.

Still another, out of the 58 original items in Schmitt’s taxonomy of VLS, a few were not included in the VLSQ because they were considered to be unsuitable for the students’ background knowledge, competence level and learning environment (Riankamol 2008). Among the strategies that were removed were “I check for L1 cognates”, “I use semantic maps”, “I use ‘scales’ for gradable adjectives”, “I use peg method”, “I use loci method” and “I learn the words of an idiom together”. However, some other strategies originally not in Schmitt’s (1997) taxonomy of VLS were included in the VLSQ. They were technologically driven strategies such as “I use an online translation to translate a new English word”, “I use a mobile dictionary to the meaning of a new English word”, and “I use online games to test my vocabulary”. They were included because although the use of technology such as online translation, mobile dictionary and online games may not be common strategies to learn and reinforce unfamiliar words in 1990s, they have become more common now. As such, Ahmad Iskandar (2008), for example, included a few of these technologically driven VLS in his study.

Finally, some of the original items from Schmitt’s taxonomy were broken into several parts as they were considered to be phrased as double-barrelled items. Double-barrelled items are items that cannot be accurately answered as they contain two or more questions which would lead to more than one possible favourable response (Del Greco and Walop 1987) that can sometimes result in a non-response (National EMSC Data Analysis Resource Center n.d.). One such item was on the usage of English-language media. Instead of just replicating that item, it was broken to several distinct strategies such as “I listen to English songs”, “I watch English movies” and “I read English books”. This was done to ensure that the respondents would not be undecided in giving their responses because they may have used one strategy frequently while they have never used another strategy altogether.

To determine whether the VLSQ is a reliable instrument for this study, it was tested for internal consistency reliability. It was found that the Cronbach’s alpha coefficient for the six sub-categories of determination, social (discovery), social (consolidation), memory, cognitive and metacognitive strategies ranged between 0.774 and 0.872. This indicates that this instrument has a good internal consistency level and thus is a suitable instrument for this study.

Semi-structured interview. The third instrument for this study was a semi-structured interview. A series of interviews were carried out with eight pre-diploma MDAB students for the purpose of triangulation that has been argued can improve the accuracy of a research’s findings (Neuman 2006) and can also increase the credibility of the study (Hussein 2009) as “one data collection form supplies strengths to offset the weaknesses of the other form (Creswell 2005: 514).

An interview guideline was developed prior to the interviews. It contained a set of predetermined questions that needed to be explored during the interview (Nur Ainil 2013). In addition, a list of possible probes was also prepared in anticipation that further clarification regarding an interviewee’s responses may be necessary. As leading questions (such as “Do you use a dictionary when you search for the meaning of a word?” or “Do you use online translation?”) should be avoided in an interview, the probes were used only once a participant had mentioned that he or she used a certain strategy. Leading questions, which are questions that are phrased to suggest a
specific answer or to imply that one response is expected or more correct, need to be avoided (Herman & Bentley 1983, cited in Center for Refugee and Disaster Studies 2000) as they can influence the responses given by the participants (Legard et al. 2003).

**Findings**

**Vocabulary Size**

Nation (2012a) estimated that a written receptive vocabulary size of 8000 word families was the necessary goal for English language learners who had to deal with unsimplified spoken and written texts. Thus, this number (8000 word families) was considered as the desirable cut-off point in term of how big a vocabulary size the MDAB students should have. This is because, as English is the medium of instruction at UiTM, they had deal with both unsimplified spoken and written texts when they were in the programme and later when they continued their studies at a diploma level at the university. While some lecturers may code switch to their native Bahasa Malaysia to help students with poor lexical knowledge, their ability (or inability) to cope with unsimplified English reading texts (especially textbooks) may prove to be critical in their quest for a university diploma from UiTM.

Table 1 indicates the vocabulary size of the MDAB students. Of the 31 respondents, it was found that 27 respondents (87.1%) had a vocabulary size of 7999 word families or less. Meanwhile, three respondents (9.7%) were found to have a vocabulary size of between 8000-8999 word families, and one respondent (3.2%) had a vocabulary size of above 10000 word families. As such, only four respondents (12.9%) had a vocabulary size of 8000 word families and above. This means that the majority of the MDAB students did not reach the threshold of 8,000 word families that is the necessary vocabulary size needed for understanding unsimplified spoken and written texts. Therefore, it can be predicted that the majority of them may encounter some serious problems when it comes to understanding their lectures, and even more so when it comes to understanding textbooks which are written in English.

<table>
<thead>
<tr>
<th>Vocabulary Size (Word families)</th>
<th>Frequency</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 4000</td>
<td>1</td>
<td>3.2</td>
</tr>
<tr>
<td>4000 - 4999</td>
<td>4</td>
<td>12.9</td>
</tr>
<tr>
<td>5000 - 5999</td>
<td>7</td>
<td>22.6</td>
</tr>
<tr>
<td>6000 - 6999</td>
<td>10</td>
<td>32.3</td>
</tr>
<tr>
<td>7000 - 7999</td>
<td>5</td>
<td>16.1</td>
</tr>
<tr>
<td>8000 - 8999</td>
<td>3</td>
<td>9.7</td>
</tr>
<tr>
<td>9000 - 9999</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Above 10,000</td>
<td>1</td>
<td>3.2</td>
</tr>
</tbody>
</table>

**Vocabulary Learning Strategies**

Table 2 summarises the VLS used by the pre-diploma MDAB students based on their vocabulary size. The results show that, overall, the students with a vocabulary size of 8,000 word families and above, and the students with a vocabulary size of below 8,000 word families were both medium users of VLS. However, the students with a vocabulary size of 8,000 word families and above had higher overall mean score (M=2.85) as compared to the students with a vocabulary size of below 8,000 word families (M=2.79). This means that the students with a vocabulary size of 8,000
word families and above used more VLS in comparison to the students with a vocabulary size of below 8,000 word families.

Table 2

<table>
<thead>
<tr>
<th>Strategy Category</th>
<th>8000 word families and above (N=4)</th>
<th>Below 8000 word families (N=27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determination</td>
<td>Mean 2.93 Frequency Medium 4 Rank 3</td>
<td>Mean 2.74 Frequency Medium 3 Rank 3</td>
</tr>
<tr>
<td>Social (Discovery)</td>
<td>Mean 3.20 Frequency Medium 1 Rank 1</td>
<td>Mean 3.38 Frequency Medium 1 Rank 1</td>
</tr>
<tr>
<td>Social (Consolidation)</td>
<td>Mean 2.56 Frequency Medium 5 Rank 5</td>
<td>Mean 2.54 Frequency Medium 5 Rank 5</td>
</tr>
<tr>
<td>Memory</td>
<td>Mean 3.03 Frequency Medium 3 Rank 3</td>
<td>Mean 2.72 Frequency Medium 4 Rank 4</td>
</tr>
<tr>
<td>Cognitive</td>
<td>Mean 2.05 Frequency Low 6 Rank 6</td>
<td>Mean 2.41 Frequency Low 6 Rank 6</td>
</tr>
<tr>
<td>Metacognitive</td>
<td>Mean 3.07 Frequency Medium 2 Rank 2</td>
<td>Mean 3.05 Frequency Medium 2 Rank 2</td>
</tr>
<tr>
<td>Overall</td>
<td>Mean 2.85 Frequency Medium 2 Rank 2</td>
<td>Mean 2.79 Frequency Medium 2 Rank 2</td>
</tr>
</tbody>
</table>

Other than that, the data also reveal that the strategy categories that were most used and least used by both groups of students were social (discovery) and cognitive strategies. Both groups of students were medium users of social (discovery) strategies and low users of cognitive strategies. Nevertheless, for these two strategy categories, the mean scores for the group of students with a vocabulary size of below 8,000 word families were higher than that of the former group. While the respective mean scores for the two strategy categories were M=3.38 and M=2.41 for the students with a vocabulary size of below 8,000 word families, they stood at M=3.20 and M=2.05 for the students with a vocabulary size of above 8,000 word families. This means the students with a vocabulary size of below 8,000 word families used more social (discovery) and cognitive strategies than the students with a vocabulary size of 8,000 word families and above.

Most Frequently Used Vocabulary Learning Strategies

Table 3 summarises the ten VLS most frequently used by the pre-diploma MDAB students based on their vocabulary size. The findings reveal that all the ten VLS most frequently used by the students with a vocabulary size of 8,000 word families and above were of high frequency use. On the other hand, nine of the ten VLS most frequently used by the students with a vocabulary size of below 8,000 word families were also of high frequency use. The tenth most frequently used strategy by the latter group fall under medium frequency use.
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I watch English movies.</td>
<td>META</td>
<td>4.30</td>
<td>High</td>
<td>I study the sound of a word.</td>
<td>MEM</td>
<td>4.75</td>
<td>High</td>
</tr>
<tr>
<td>I listen to English songs.</td>
<td>META</td>
<td>4.30</td>
<td>High</td>
<td>I listen to English songs.</td>
<td>META</td>
<td>4.75</td>
<td>High</td>
</tr>
<tr>
<td>I use a bilingual dictionary to find the translation of a new English word.</td>
<td>DET</td>
<td>4.19</td>
<td>High</td>
<td>I watch English movies.</td>
<td>META</td>
<td>4.75</td>
<td>High</td>
</tr>
<tr>
<td>I watch English TV programmes.</td>
<td>META</td>
<td>3.89</td>
<td>High</td>
<td>I visualise the spelling of a word.</td>
<td>MEM</td>
<td>4.50</td>
<td>High</td>
</tr>
<tr>
<td>I use an online translation to translate a new English word.</td>
<td>DET</td>
<td>3.74</td>
<td>High</td>
<td>I watch English TV programmes.</td>
<td>META</td>
<td>4.50</td>
<td>High</td>
</tr>
<tr>
<td>I ask my friends for the meaning of a new English word.</td>
<td>SOC (DIS)</td>
<td>3.63</td>
<td>High</td>
<td>I study the spelling of a word.</td>
<td>MEM</td>
<td>4.25</td>
<td>High</td>
</tr>
<tr>
<td>I ask my teacher for Malay translation of a new English word.</td>
<td>SOC (DIS)</td>
<td>3.56</td>
<td>High</td>
<td>I ask my teacher to check my word lists for accuracy.</td>
<td>SOC (CON)</td>
<td>4.00</td>
<td>High</td>
</tr>
<tr>
<td>I study the spelling of a word.</td>
<td>MEM</td>
<td>3.56</td>
<td>High</td>
<td>I guess for the meaning of a new English word by looking at</td>
<td>DET</td>
<td>4.00</td>
<td>High</td>
</tr>
</tbody>
</table>
The strategy that was most frequently used by the students with a vocabulary size of 8,000 word families and above was “I study the sound of a word”. This was followed by the strategy “I listen to English songs”, “I watch English movies”, “I visualise the spelling of a word” and “I watch English TV programmes”. In comparison, the strategy that was most frequently used by the students with a vocabulary size of below 8,000 word families was “I watch English movies”, which was followed by “I listen to English songs”, “I use a bilingual dictionary to find the translation of a new English word”, “I watch English TV programmes” and “I use an online translation to translate a new English word”. The findings further reveal that five strategies (“I watch English movies”, “I listen to English songs”, “I use a bilingual dictionary to find the translation of a new English word”, “I watch English TV programmes” and “I visualise the spelling of words”) appeared to be among the top ten most used VLS by both groups of students.

These quantitative data from the VLSQ were corroborated by the students’ responses in the series of semi-structured interview. Firstly, six out of eight participants of the interview conceded that they used bilingual dictionaries to check for the meanings of unfamiliar English words. A participant (M2) remarked that “I'm trying to understand ...trying to understand the word, and if I cannot understand I must use a dictionary to translate the word until I understand the word”. Explaining why he used a bilingual dictionary more frequently than a monolingual dictionary, M2 explained that “To understand...because it’s easy”. Another participant (M3) reported that “If I very very not understand what the word...very difficult word to understand...I used dictionary. Malay-English dictionary”. On the other hand, M1 clarified that “If I search on English-English dictionary [and] I don’t understand I search on English-Malay dictionary”.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I visualise the spelling of a word.</td>
<td>MEM</td>
<td>3.50</td>
<td>High</td>
<td>I use a bilingual dictionary to find the translation of a new English word.</td>
<td>DET</td>
<td>3.75</td>
<td>High</td>
</tr>
<tr>
<td>I study the sound of a word.</td>
<td>MEM</td>
<td>3.41</td>
<td>Medium</td>
<td>I guess for the meaning of a new English word by looking at its part of speech.</td>
<td>DET</td>
<td>3.75</td>
<td>High</td>
</tr>
</tbody>
</table>
The results also show that three metacognitive strategies (“I listen to English songs”, “I watch English movies” and “I watch English TV programmes” made into the top five VLS most used by both groups of students. Data from the semi-structured interviews corroborated this finding. First, two interview participants (F5 and F7) acknowledged that the strategy “I watch English movies” was a strategy that they frequently used. F5 explained that “Watch Malay story...subtitle English at below. Okay I know, understand the meaning...or English story I look at [Malay] subtitle...Many times lah. Always watched drama”. F7, on the other hand, mentioned that she watched English movies quite frequently: “Three or two [movies every month]”. She further stated that “When I see movie, the subtitle... at the bottom of the...so I see it and I don’t know, okay I made a list and I search the meaning”. Second, three interview participants (M2, M3 and F7) stated that they used the strategy “I listen to English songs” with two of them (M2 and M3) saying they always used this strategy. M2 maintained that “Hearing... English music, to know the new word...about unfamiliar word that I don’t know. Then if I don’t know the word I keep hearing many time”. He further explained “Yes, a song. Every day when I want to sleep. Before I want to sleep I sing...I hearing to the song”. Meanwhile, M3 said that “I really like to hear songs... in English. Even I don’t know what...what the song in Malay, I try to understand and then I list the song [the words] I try to find in dictionary”. This strategy is a strategy that M3 used “every day”. On the other hand, F7 declared that “Sometimes I heard songs and what the meaning I don’t know I quickly list it”. Last, F7 was the only interview participant who revealed that she used the strategy “I watch English TV programmes”. She said that “I’m...hear radio, English radio. On the radio and TV...at TV, National Geographic. At the same time...National Geographic, and one more...ntv7...[the TV programme] Bella”.

Least Frequently Used Vocabulary Learning Strategies

Table 4 lists the five VLS least used by the pre-diploma MDAB students based on their vocabulary size. The two strategies that were among the least used strategies by both groups of students with a vocabulary size of 8,000 word families and above, and students with a vocabulary size of below 8,000 word families were “I keep flash cards” and “I put English labels on physical objects”. The other three strategies that were least used by the former group were “I skip or pass new words”, “I keep word lists” and “I used spaced word practice”. Meanwhile, for the latter group, the other three strategies that were least used by them were “I use flash cards to learn new English words”, “I use an online thesaurus to find the meaning of new English words” and “I interact with native speakers online”. These findings were verified through the semi-structured interview. None of the interview participants mentioned any of these VLS except one participant who mentioned that she kept word list. She claimed that “I see the list of the vocabulary and I search the meaning but I get more meaning from the words. So I know the words is not only used in one meaning but many meanings”.
Table 4
Least frequently used vocabulary learning strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
<th>Strategy</th>
<th>Category</th>
<th>Mean</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I use flash cards to learn new English words.</td>
<td>DET</td>
<td>1.48</td>
<td>Low</td>
<td>I skip or pass new words.</td>
<td>META</td>
<td>1.25</td>
<td>Low</td>
</tr>
<tr>
<td>I use an online thesaurus to find the meaning of a new English word.</td>
<td>DET</td>
<td>1.69</td>
<td>Low</td>
<td>I keep flash cards.</td>
<td>COG</td>
<td>1.25</td>
<td>Low</td>
</tr>
<tr>
<td>I keep flash cards.</td>
<td>COG</td>
<td>1.74</td>
<td>Low</td>
<td>I keep word lists.</td>
<td>COG</td>
<td>1.25</td>
<td>Low</td>
</tr>
<tr>
<td>I interact with native speakers online.</td>
<td>SOC (CON)</td>
<td>1.85</td>
<td>Low</td>
<td>I put English labels on physical objects.</td>
<td>COG</td>
<td>1.50</td>
<td>Low</td>
</tr>
<tr>
<td>I put English labels on physical objects.</td>
<td>COG</td>
<td>1.96</td>
<td>Low</td>
<td>I use spaced word practice.</td>
<td>META</td>
<td>1.50</td>
<td>Low</td>
</tr>
</tbody>
</table>

Discussion
One major finding of the study is that the majority of the MDAB pre-diploma students had a written receptive vocabulary size of below 8,000 word families. This means that they did not reach the desired threshold of 8,000 word families that is the necessary vocabulary size needed for understanding unsimplified spoken and written texts (Nation & Beglar 2007; Nation 2012). This finding is consistent with the results in Ahmad Azman’s (2009) and Engku Haliza et al.’s (2013) studies. The respondents in Ahmad Azman’s (2009) study were also found to have limited receptive vocabulary size, while the subjects in Engku Haliza et al.’s (2013) study also had not achieved “the recommended threshold of vocabulary” (p.44). This finding on the MDAB pre-diploma students’ vocabulary size suggests that they may encounter some serious problems when they continue their studies in the more advanced levels at UiTM. This is because English is the medium of instruction at the university. In fact, lack of English language proficiency has been identified as a major factor in the poor academic achievement of the MDAB students at UiTMJS (Rosnani & Nurul Hasni 2012). As such, a systematic VLS training could be one measure to be taken to
address this issue as lack of exposure to a systematic VLS training has been identified as a major factor that contributes to poor vocabulary size (Engku Haliza et al. 2013). Besides, it has been reported that systematic VLS trainings have positive effects on vocabulary learning (Zhao 2009, cited in Mazlin 2013) which, in turn, could help students to improve their English vocabulary size.

Another major finding of the study is that, overall, the students with a vocabulary size of 8,000 word families and above, and the students with a vocabulary size of below 8,000 word families were both medium users of VLS. These results were very similar to the findings in the studies by Mohd Sahandri et al. (2009), and Kafipour et al. (2011). In these two studies, their respondents were also found to be medium VLS users for overall strategy and for each VLS sub-strategy. The only difference was that in this present study while the MDAB students proved to be medium VLS users for overall VLS and five VLS sub-categories, they were low strategy users for one category. The category that showed low frequency use was the cognitive strategy. However, these findings of the study were different from that of the study by Lou (2014) whose respondents were found to be high strategy users for metacognitive strategies, and medium frequency users of cognitive and social/affective strategies.

The study also found that the strategy that was most frequently used by the students with a vocabulary size of 8,000 word families and above was “I study the sound of a word”. In comparison, the strategy that was most frequently used by the students with a vocabulary size of below 8,000 word families was “I watch English movies”. The findings further reveal that five strategies (“I watch English movies”, “I listen to English songs”, “I use a bilingual dictionary to find the translation of a new English word”, “I watch English TV programmes” and “I visualise the spelling of words”) appeared to be among the top ten most used VLS by both groups of students. Similarly, using a bilingual dictionary was found to be one of the most frequently used VLS in Scmitt’s study (1997), while using English language media (which includes watching movies and TV programmes, and listening to songs) in the study by Mohd Sahandri et al. (2011). Finally, findings reveal that two strategies (“I keep flash cards” and “I put English labels on physical objects”) were least used by both groups of the students.

**Limitations**

In terms of limitation, firstly, as this is a case study which involves only pre-diploma MDAB students at UiTMJS during Session 2 2013/2014 semester, the findings of the study cannot be generalised to other MDAB students at different UiTM campuses throughout Malaysia.

Secondly, as the English vocabulary size explored in this study is limited only to the students’ written receptive vocabulary size which refers to their vocabulary knowledge for reading, the findings on their vocabulary size are limited only to the students’ reading vocabulary size and do not include other kinds of vocabulary size such as listening, speaking, writing and focal vocabulary size. However, the selection of written receptive vocabulary size as the type of vocabulary size investigated in this study is appropriate as it is the written receptive vocabulary size of around 8,000 word families that is estimated to be the critical vocabulary size for English language learners who have to deal with unsimplified spoken and written words (Nation & Beglar 2007; Nation 2012).

Other than that, the use of the Vocabulary Learning Strategy Questionnaire as one of the instruments for data collection also interposes another limitation. This is
because in giving their responses to the questionnaire, some of them may not be absolutely honest while some others may have responded to strategies they think they should have used instead of focusing on strategies they have used. However, in order to compensate for the anticipated issues connected to a self-report survey questionnaire as stated above, a series of semi-structured interviews were conducted with eight of the MDAB students so that they could provide supporting information and further insights about the VLS that they use to learn English. Nevertheless, as the number of participants for the interviews was small, this could be considered as another shortcoming. Some students’ unwillingness to be interviewed for the study had made it impossible for this researcher to interview more students. Furthermore, as this researcher herself conducted the interviews, her lack of experience and interviewing skills may have resulted in some underexplored findings from a potentially rich source of information.

**Recommendation**

In replicating this study, it is recommended that a larger number of intermediate and low proficiency ESL students should be involved. For example, in the case of MDAB students, respondents can be from several UiTM campuses instead of from only one campus. If this is undertaken, comparisons can be made on the vocabulary size and vocabulary learning strategies used by the students from different campuses and/or by pre-commerce and pre-science MDAB students. Other than that, when a larger number of students is involved, the relationship between the students’ vocabulary size and VLS, and the contribution of VLS to vocabulary size could be explored.

Another recommendation would be for future research to utilise the online version of the VST. The availability of this online version would allow a researcher to include a large sample as test scores would be automatically calculated.

In addition, it is also suggested that experimental research on some vocabulary learning strategies could be undertaken to learn more about their effects on language learners’ vocabulary knowledge. For example, it would be interesting to explore the effects of technologically-driven vocabulary learning strategies such as using online translation and mobile dictionary on the learners’ vocabulary size.

**References**


ENGLISH VOCABULARY SIZE AND VOCABULARY LEARNING


Ling, S. 2015. Poor language proficiency is harming our youth’s future. The Star Online, 14 November.


