"I Will Study More . . . and Pray"

Metacognition About High-Stakes Test Preparation Among Culturally & Linguistically Diverse Students

Debra A. Giambo

Introduction

High-stakes testing continues in the U.S. public schools despite a plethora of concerns from the field (e.g., Council of the Great City Schools, 2015; Krashen, 2012a, 2012b; Lazarín, 2014; Ravitch, 2011, 2016a, 2016b) and among parents nationwide (e.g., Brown, 2015; Layton, 2015; Wallace, 2015). While there have been some modifications to testing requirements in some states (e.g., Layton, 2015), the majority of students and teachers in the U.S. continue to experience a rigorous emphasis on test preparation.

For culturally and linguistically diverse students, some of whom have not yet achieved academic English proficiency, high-stakes testing can bring additional issues and challenges (Coltrane, 2003; Hopkins et al., 2013; Huang, Han, & Schnapp, 2012; Luykx et al., 2007, Mahon, 2006; Murphy, 2007; Robinson-Cimpian, Thompson, & Umansky 2016; Sanchez et al., 2009; Solórzano, 2008). For English language learners (ELLs), high-stakes testing in English becomes a test of English proficiency, rather than a test of content knowledge (Tsang, Katz, & Stack, 2008), therefore raising questions regarding the reliability and validity of such assessments for students with limited English proficiency (Hopkins et al, 2013; Robinson-Cimpian et al., 2016).

California, Texas, and Florida had the highest number of language minority students among U.S. states as of 2012-2013 (National Center for Education Statistics, 2015a). In California, in the fall of 2014, 22.3% (1,392,263) of the students in public schools were ELLs. Additionally, 42.9% (2,672,128) speak a language other than English at home, including both ELLs and students fluent in English (California Department of Education, 2015c). In 2013-2014, 17% (900,476) of students enrolled in Texas' public school were ELLs (Texas Education Agency, 2014). In 2015-2016, 9.8% (273,570) of Florida's public school students were classified as ELLs (Florida Education Department, 2016a).

In these same states, scores from the 2015 National Assessment of Education Progress (National Center for Education Statistics, 2016) demonstrate a gap between the reading proficiency of 4th and 8th grade ELLs and non-ELLs. In California, 28% of 4th grade ELLs demonstrated proficiency in reading, compared to 72% of students who were not ELLs. Twenty-two percent of 4th grade ELLs in Texas demonstrated proficiency in reading, whereas 78% of non-ELLs demonstrated reading proficiency. In Florida, 9% of 4th grade ELLs' scores indicated proficiency in reading, compared to 91% of students who were not ELLs.

For 8th grade ELLs and non-ELLs in each of these states, the gap in reading proficiency scores was even wider: In California, the gap, respectively, was 14% and 86%; the scores in Texas, respectively, were 11% and 89%, and in Florida, they were 5% and 95% (National Center for Education Statistics, 2016).

As of 2011-2012, California, Texas, and Florida each required the passing of at least the reading/language arts and mathematics subtests for high school graduation (Center on Education Policy, 2012; National Center for Education Statistics, 2011-12), although such requirements have been in transition in recent years. The California High School Exit Examination (CAHSEE) was a requirement for earning a high school diploma beginning with the class of 2006 (California Department of Education, 2015a) and was preceded by a comprehensive competency exam (Center on Education Policy, 2009). The CAHSEE graduation requirement was temporarily suspended in January 2016 for graduating classes between 2015 and 2018 with a grandfather clause for students that went back to 2003.

From 1986 to 2011, the Texas Education Agency required the passing of a high-stakes test for high school graduation and, subsequently, end-of-course assessments in 2012-2014. Requirements have been revised for students who are 11th or 12th graders in 2014-2017 to stipulate that a student who has failed no more than two end-of-course assessments may earn a high school diploma by committee decision (Texas Education Agency, 2016a, 2016b).

Florida implemented the Florida Comprehensive Assessment Test (FCAT), a statewide assessment of students' academic content, in 1996. In 1999, the legislature added the use of the FCAT for graduation from high school and a grading system for schools based on FCAT scores. Subsequently, in 2001, the State Board of Education established passing scores on the FCAT to determine eligibility for a standard high school diploma. As of the 2015-2016 academic year, Florida continues to require students to pass the reading/English language arts subtest of high-stakes standardized test for graduation, now the Florida Standards Assessment (FSA), and also requires passing an end of course assessment in Algebra I (Florida Department of Education, 2015b).

In California, Texas, and Florida, the reported passing rates on the high-stakes test required for graduation indicate that White, non-Hispanic students showed the highest passing rate when compared to all students, Hispanic students, and ELLs. Significant numbers of Hispanic students and ELLs in Texas and Florida have not passed the assessments. Results reported here include selections of the most recently reported passing rates gleaned from each of the states' education agencies' websites.

In California, 85% of all students are

Debra A. Giambo is a professor in the Department of Curriculum, Instruction, and Culture in the College of Education at Florida Gulf Coast University, Fort Myers, Florida.

reported to have scored as passing the CAHSEE, while 93% of White, non-Hispanic students scored as passing, 80% of Hispanic students, and 43% of ELLs passed the exam (California Department of Education, 2015b). Passing rates in Texas on EOC exams for English II (as an example) in 2014-2015, indicate that 44% of White, non-Hispanic students scored as passing, while 35% of all students, 34% of Hispanic students, and 18% of ELLs scored as passing (Texas Education Agency, 2015).

In Florida, passing rates on the FCAT 2.0, required for a standard high school diploma in 2013 and 2014, were consistent across these two years for a variety of student groups in 10th grade (Florida Department of Education, 2016b). In 2014, statewide assessment results indicated that 55% of all students passed. When separating out White, non-Hispanic students, and Hispanic students, scores indicated 68% passing and 50% passing, respectively.

Results for ELLs in Florida showed a passing rate of 11%. Scores for 2013 saw similar patterns with a difference of 1-2% lower passing rates for each of the student groups mentioned above. While the passing rates for ELLs and Hispanic students are lower than that of other student groups discussed, overall, it is important to note that, depending on the reporting forum, high-stakes test passing rates of ELLs in Florida may have been inflated in the past (Giambo, 2009).

Graduation rates in California, Texas, and Florida indicate patterns similar to the passing rates of the high-stakes tests in which we see discrepancies between White, non-Hispanic students, Hispanic students, and ELLs. Data reported by the National Center for Education Statistics (2015b) indicated that graduation rates for ELLs and Hispanics remained below those of White, non-Hispanic students, and all students combined in these states. Calculated using the Adjusted Cohort Graduation Rate (ACGR), the 2013-2014 graduation rates for all students, White, Hispanic, and ELLs, are displayed in Table 1.

In addition to considering the high-stakes-test passing rates, an understanding of the drop-out rates can provide a more complete picture of, for example, Florida's ELLs' academic achievements in high school. For example, the Florida Department of Education reports that, in 2006, the statewide drop-out rate in Florida was 3.5% (2008), and, in 2013-2014, the drop-out for all students fell to 1.9% and was 2.3% for Hispanic male students (Florida Department of Education, 2015a). However, it is possible that the actual drop-out rate was considerably higher due, in part, to the withdrawal codes used to indicate that a student has left school. Among several withdrawal codes, only one indicated that a student has dropped out of school. Other codes could be used when a student leaves without explanation or to enter a GED program, for example, and the student would not be counted as dropping out of school, even if the student drops out of the GED program (Giambo, 2009; Merrow, 2004).

A separate code indicated that a student's whereabouts was unknown (Florida Department of Education, 2015a). Considering the high-stakes testing requirement for high school graduation, the ELL pass rate, the use of withdrawal codes, and the pressure on school districts to show continual improvement, it is possible that the drop-out rate was under-reported and that the graduation rate, especially for ELLs, was over-reported (Giambo, 2009).

Investigating Perceptions

Since high stakes testing can be an inaccurate measure of educational achievement for culturally and linguistically diverse students (Coltrane, 2003; Hopkins et al., 2013; Huang et al, 2012; Luykx et al., 2007, Mahon, 2006; Murphy, 2007, Robinson-Cimpian et al., 2016; Sanchez et al., 2009; Solórzano, 2008; Tsang et al., 2008), an important step toward more promising opportunities for this group of students is to investigate their perceptions of the test and effective test preparation as well as the effects on their motivation to achieve academically.

Bandura's social cognitive theory includes the perspective that perceptions of self-efficacy influence motivation and action and are affected by social factors (Bandura, 1991). Central to this theory is the role of "cognitive, vicarious, self-reflective, and self-regulatory processes" in human adaptation and change (Bandura, 1989). The relevance of Bandura's theory to this study lies with the examination of culturally and linguistically diverse students' perceptions of a high-stakes test and their self-reflection on how they prepare for and adjust to the requirements of confronting the test.

In other words, self-perceptions of effectiveness in test preparation influences actions so that students who engage in self-regulatory processes are more likely to adapt and change. Thus, guided by social cognitive theory, it can be anticipated that students who are self-reflective in their preparation for a high-stakes test may be more likely to make changes in their preparation or subsequent academic behavior.

Student perceptions of assessments affect the manner in which they approach preparation as well as test results, and this has been found with middle and high school students as well as university students. When students perceive an assessment as inappropriate or inauthentic, they employ more superficial methods of preparation than if they perceive the assessment as valuable and appropriate (Alkharusi, 2013; Hong & Peng, 2008; Lizzo, Wilson, & Simons, 2002). Increased preparation efforts for a test perceived as valuable, however, do not necessarily mean that students who perceive themselves as using metacognitive strategies will do better on the test (Hong & Peng, 2008; Purpura, 1997).

Students who anticipate poor performance on a test are more likely to attribute their score to external factors, such as distractions in the testing situation as well as issues or situations outside of the test, than those who expect a good outcome (Gaier, 1962). Increased efficacy in approaching the goal of effectively helping culturally and linguistically diverse students improve performance on high-stakes assessments may involve better understanding of their approaches to testing via their perceptions of: the test, their preparation for the test, and the effect of the test on academic behaviors.

The purpose of this study was to investigate the following research questions:

Table 1

Adjusted Cohort Graduation Rate (2013-2014) for students in California, Texas, and Florida

States	*All students*	*White, non-Hispanic students*	*Hispanic students*	*ELLs*
California	81%	88%	77%	65%
Texas	88.3%	93%	85.5%	71.5%
Florida	76.1%	81.7%	75%	55.8%

1. What do culturally and linguistically diverse students think and feel about high-stakes testing (i.e., FCAT)?

2. What are culturally and linguistically diverse students' perceptions regarding effective preparation for high-stakes testing (in classes and individually)?

3. Do culturally and linguistically diverse students perceive that high-stakes testing affects their academic motivation, resulting in perceived differences in academic behavior?

Methodology

Measures

Measures employed in this study included a survey and a focus group interview protocol. Both the survey and the interview protocol were developed in conjunction with three high school ELL teachers, each with 15 to 26 years experience in education, and with an interdisciplinary panel of researchers.

The survey questions were examined by these stakeholders for balance in terms of allowing for both positive and negative responses, appropriateness, responsiveness to research questions, balance between Likert-type and free responses, and level of difficulty in terms of possible limitations in participants' English proficiency.

The survey included 18 items divided into three parts. Part one included four open-ended questions and one yes/no item focusing on students' future plans; part two was composed of six items requiring Likert-type scale responses that focused on students' general perceptions of the FCAT; and part three included five open-ended questions and two yes/no questions intended to elicit more specific information on students' perceptions.

Likert scales are often viewed as an effective way to measure attitudes (Alcázar-Olán, Deffenbacher, Escamilla-Tecalco, 2016; Geldhof, Gestsdottir, Stefansson, 2015; Likert, 1932; Page-Bucci, 2003) and seem to be similar in reliability and validity to traditional measures of self-efficacy (Maurer & Andrews, 2000). Additionally, the use of focus group interviews to elicit meaningful information on perspectives has long been recognized as effective (Butler, 2002; Center for Disease Control, 2008; De Groot, 2002; Israel & Galindo-Gonzalez, 2014; Patrick & Middleton, 2002; Zhu, 2005).

The focus group interview protocol consisted of five open-ended questions with supporting prompts. The content was intentionally similar to the categories of topics on the survey in order to elicit more in-depth information. Consideration was given to the possibility that students completing the survey may have some limitations in their English proficiency as well as the possibility of limited motivation to complete the survey, so the focus group interview was intended to be an opportunity for students to provide additional information.

Participants

Participants in both the survey and the focus group interviews were culturally and/or linguistically diverse students in South Florida. Survey participants included elementary (4th to 6th graders) and middle and high school students, and focus group participants included middle and high school students. All were active in programs that support underrepresented students in their preparation for post-secondary education, offered through local colleges and universities in cooperation with the students' home schools. Participation was limited to those for whom permission to participate had been obtained.

All participants had diverse backgrounds based on ethnicity (i.e., their families had come to the U.S. from another country) and/or language (i.e., they were experiencing or had experienced limited English proficiency). Many participants had come to the United States at a variety of ages while some had been born in the U.S. At the time of the interviews and surveys, all participants demonstrated facility with English sufficient for completion of the survey and/or focus group interview based on consultations between and recommendations from the college program coordinator and the cooperating teacher at the participants' schools.

The 62 survey participants included 25 elementary students, 12 middle school students, and 25 high school students. According to self-report, nine were currently receiving ELL services, nine had received ELL services in the past, and 44 had never received ELL services. Additionally, survey participants self-reported information about their first language, home language, and language used with friends. Thirty-seven reported speaking English as their first language, seven spoke Spanish as their first language, and 17 spoke Haitian-Creole as their first language, with one unreported. At home, 40 spoke English, nine spoke Spanish, and 13 spoke Haitian Creole, and this includes one participant who reported speaking both English and Haitian Creole equally frequently at home. With friends, 55 used English, two spoke Spanish, three spoke Haitian Creole, and one spoke both English and Haitian Creole equally frequently (with one non-response).

Of the 38 participants in the focus group interviews, 23 were middle school students and 15 were high school students. All participants were from schools in two adjacent county school districts. Of the 23 middle school focus group participants, 14 reported that they had received ELL services in the past, and nine reported never having received ELL services. Nine of the 15 high school focus group interview participants reported that they were currently receiving ELL services, one had received services in the past, and five reported never having received ELL services.

Focus group participants also reported on their use of languages other than English. Among the middle school participants, six spoke Spanish, 14 spoke Haitian Creole (seven of these also reported speaking French), and the remaining three spoke Arabic, Tagalog, or Hungarian. The high school participants reported a similar distribution, with three who spoke Spanish, 11 who spoke Haitian Creole (four of whom also reported speaking French), and one who spoke French and Dutch.

Procedures

Survey

To maximize participants' comfort level, the surveys were administered by staff members of the college-bound programs, and assistance was available, if needed, during the completion of the survey. Participants completed the surveys during one of their college or university on-campus visits.

Focus Group Interviews

Participants were interviewed in the college-bound program offices during one of their on-campus visits. Most participants had been to these locations multiple times, so the location was familiar to them. The interviews were conducted by a researcher and a graduate student, lasted approximately 45 minutes, and included an average of six participants, with a range from four to 10, depending on logistics.

An interview protocol was followed in each interview which allowed for introductions followed by an emphasis on the importance of each person's opinion. Participants were told the purpose of the interviews (i.e., to understand their thoughts about the FCAT) and logistical rules (i.e., allowing each person to respond without others talking, not needing to agree with anyone else, respecting others

and their responses, and respecting the confidentiality of responses following the close of the interview).

Handwritten notes were taken during the interviews, but interviews were not audio-taped to encourage participation and reduce anxiety. Numbers were used for identification purposes both on the table during the interview and in the notes to maintain anonymity in the written record of the interviews. At the conclusion of the interviews, students were thanked and reminded to respect the confidentiality of everyone's responses.

Data Analysis

Data were entered into three databases: one for Likert-type survey responses, one for survey free-response items, and one for focus group interview responses. Free responses were themed, and both numerical and free response items were analyzed for frequency of responses. Direct quotes were entered into the database to enhance and clarify responses and to maintain and illustrate student voices.

Results

Research Question 1: What do culturally and linguistically diverse students think and feel about high-stakes testing (i.e., FCAT)?

While elementary participants in this study tended overall to offer more positive responses about how the FCAT makes them feel, the majority of middle and high school participant responses indicated a neutral view of the assessment. Among the elementary school participants, 68% (17 out of 25) responded on the survey that the FCAT makes them feel good, while 58% (21 out of 37) middle and high school students indicated that the FCAT made them feel neutral (i.e., neither good nor bad) (See Table 2 for specifics).

Focus group interviews with middle and high school students indicated that 39% (15 out of 38) of participants agreed with a negative view of the FCAT. Students indicated concern about the use of the FCAT to determine future educational opportunities and graduation from high school as well as having to take the FCAT over and over beginning in 3rd grade.

A similar pattern of responses occurred in the focus group interviews, with younger students who were, in this case, middle school participants, demonstrating a more neutral view and high school participants exhibiting stronger, negative reactions and stating that the FCAT involves unrealistic standards, wastes time, and provides pressure for students newly arrived to the U.S. One view, expressed by a middle school participant, was reflective of a frequently expressed concern: "It can be a bad day and can affect the rest of your life."

The majority of the middle school and high school participants indicated that they do not feel that the FCAT allows them to show a lot of what they know and so may consider the test to be an unfair assessment of their knowledge and skills, while the elementary participants again responded more positively. Specifically, 68% (25 out of 37) of middle and high school participant responses indicated that the FCAT either lets them show "some" to "not much" of what they learned in school. However, the majority of elementary participants indicated that the FCAT gave them an opportunity to show "a lot" or "more than some" of what they know: 84% (21 out of 25) indicated showing a lot of what they know, 12% (three out of 25) responded they can show some of what they know on the FCAT.

Research Question 2: What are culturally and linguistically diverse students' perceptions regarding effective preparation for high-stakes testing (in classes and individually)?

Responses to relevant survey questions were themed to examine trends in responses. In some instances, participants indicated more than one idea in response to a question, and, in those cases, the total number of responses exceeds the number of participants. Responses were tallied among themes, and tallies are provided without percentages due to the occurrence of multiple ideas from some participants.

Regarding student perceptions of what is helpful in doing well on the assessment, total responses indicated the following were helpful, listed in descending frequencies: studying (15), learning in class (12), teachers (6), faith/confidence/positive thinking/perseverance (5), and five did not provide a response to the question (See Table 3.)

While some responses indicate an awareness of the benefits of actively participating in preparing to do well on the assessment (e.g., studying, learning in

Table 2
Survey Responses Regarding Feelings about the FCAT

Responses	*Elementary n = 25*	*Middle n = 12*	*High n = 25*	*Totals n = 62*
1 - Good	17 (68%)	0	1 (4%)	18 (29%)
2	3 (12%)	1 (8%)	4 (16%)	7 (11%)
3 - Neither good nor bad	5 (20%)	9 (75%)	12 (48%)	26 (42%)
4	0	1 (8%)	2 (8%)	3 (5%)
5 - Bad	0	1 (8%)	6 (24%)	7 (11%)

Table 3
Participant Perceptions of What Helps Them Do Well on the High-Stakes Test (number of survey responses in order of frequency)

What helps me do well	*Elementary*	*Middle*	*High*	*Total number of responses*
Studying	5	5	5	15
Learning in class	5	2	5	12
The teacher	5	0	1	6
Faith, confidence, positive thinking, perseverance	3	0	2	5
No response	1	0	4	5
Reading	2	0	2	4
Eating breakfast	0	3	1	4
Concentrating/ taking time	1	1	1	3
Re-reading passages and questions	1	1	0	2
Preparation	0	1	1	2
Music	0	1	1	2
Practicing FCAT-type questions	0	0	2	2
Sleeping enough the night before	1	1	0	2
Relaxing	2	0	0	2
Using a dictionary	0	0	1	1
Common sense	0	0	1	1
Knowing what to do	0	0	1	1
Nothing	0	0	1	1

class, perseverance, reading, concentrating, taking time, re-reading questions and passages, practicing FCAT-type questions, using a dictionary), these were unspecific in nature, and many of the responses were further removed from active participation in learning and improving preparation (e.g., faith, confidence, positive thinking, eating breakfast, preparation, music, sleeping the night before, relaxing, common sense, knowing what to do, doing nothing).

When asked what participants could do to help themselves do well on the high-stakes assessment, responses indicated the following: studying (25), getting adequate sleep (8), paying attention in class and on the test (7), having breakfast (7), and seven either did not respond or responded that they did not know. There were between one and four responses indicating the following would be helpful: staying calm (4), practicing (3), studying the night before (2), nothing (2), working hard (2), reading more (2), understanding and trying to remember (2), listening to music (1), reading the book (1), trying one's best (1), staying confident (1), not talking about the test (1), having good attendance (1), doing nothing different because of success on the test in the past (1), preparing (without being specific; 1), and praying (1). (See Table 4). As with the previous question, many responses indicated the benefits of actively participating in preparation but were unspecific or not directly connected to preparation for a high-stakes test.

When asked on the survey how classes helped prepare participants for the FCAT, 31% (19 out of 62) participants did not respond to the question. Fifteen responses (24%) indicated that classes help them prepare without providing a response as to how that occurs, 23% (14 out of 62) indicated that they help because what they learn is what they have to do on the FCAT. Additional responses were low in frequency (i.e., teacher provides helpful tips for the test [2 responses] and teacher enthusiasm [1 response]). Other responses were not related to the question (i.e., enjoying classes, because classes help, and some questions on the FCAT are difficult). These responses regarding classroom help in preparation for the FCAT were, therefore, were predominantly either nonexistent or nonspecific in nature.

In the focus group interviews, participants varied on their responses about how their classes helped to prepare them for the assessment. Some middle school participants expressed that teachers helped prepare them as well (26%) while some middle and high school participants felt that teachers did not seem to care about their preparation or prepare them well (42%). There was some acknowledgment among the middle school participants that their language arts and mathematics classes were the only classes that were helpful in preparing for the assessment, as these were (at the time) the only subjects covered by the assessment.

Middle school participants focused more on the dull colors used in the test as well as the irrelevance of the reading passages to their lives (30%), while the high school participants concentrated more on the pressure they felt regarding the high-stakes of the assessment (33%). While participants brought up some issues and ideas that may help students score somewhat higher on the assessment, comparison of responses in the focus groups and on the surveys shows consistency, overall, regarding the lack of substance when considering effective preparation for the assessment.

Research Question 3: Do culturally and linguistically diverse students perceive that high-stakes testing affects their academic motivation, resulting in perceived differences in academic behavior?

When asked if they thought they could do better in their classes after taking the FCAT, and why or why not, more than half of survey participants and nearly half of focus group participants responded positively. Survey data indicate that 56% (35 out of 62 participants) thought the FCAT would help them in their classes, while 37% (23 out of 62) responded negatively. Six percent provided no response.

Some examples of responses from the focus group interviews include: "I see I don't know stuff on the FCAT, [and I want to] do better next year"; "It makes me want to think harder"; "I want to learn what they be talking about." One student explained that a friend described the FCAT as a video game: If you beat a level, you go on to the next level; your skills improve as the levels get harder. Some indicated this perception was due to the lower pressure after the test is over.

When asked how the FCAT helps or would not help them do better in their classes, more than a third of participants (34% or 21 out of 62) provided no response (See Table 5.) Fifteen percent provided responses within the theme that the FCAT shows them what they need to learn, while 18% felt that the FCAT would not help in their classes due to their view that they did not see the relevance of the test to their classwork. Each of the additional responses were provided by 6% (four out of 62) or fewer of participants. One motivated participant pointed out that, although he wanted to learn from his mistakes on the test, the results were not provided in itemized format, so he was unable to do so.

Table 4

Perceptions on What Participants Can Do to Help Themselves Do Better on the Assessment (number of survey responses in order of frequency)

What I can do to help me do better	*Elementary*	*Middle*	*High*	*Total responses*
Study	12	4	9	25
Sleep	4	3	1	8
Pay attention in class/on test	4	1	2	7
Eat breakfast	5	1	1	7
No response/ don't know	1	1	5	7
Stay calm	3	0	1	4
Practice	1	1	1	3
Study night before	0	0	2	2
Nothing	0	2	0	2
Work hard	1	1	0	2
Read more	0	0	2	2
Understand/try to remember	0	0	1	2
Listen to music	0	0	1	1
Read the book	0	1	0	1
Try my best	1	0	0	1
Stay confident	1	0	0	1
Not talk about the test	1	0	0	1
Have good attendance	1	0	0	1
Do nothing different (past success on test)	1	0	0	1
Prepare	0	1	0	1
Pray	0	0	1	1

Approximately one-third of focus group participants responded that the FCAT would not result in their doing better in their classes, and that there seemed to be a disconnect between the FCAT and class work. For example, "[The] FCAT did not apply to class work"; "Because it's not a true test of what I learned in class." Additional remarks indicated that instructors often do not want to discuss the FCAT after it is over: "If we ask a question in reading [class] after the FCAT, they won't explain it and say to ask next year." One student responded both positively and negatively and said that, although s/he gets nervous on test day and forgets everything, s/he recognized the benefit of having to focus to take the test.

Discussion

The culturally and linguistically diverse students in this study demonstrated a difference in their view of the FCAT depending on their grade level. More elementary participants had a positive view of the assessment, while middle and high school participants' views demonstrated neutral or negative views of the assessment. Some middle and high school participants connected this with the pressure to pass the test and concern over the use of the test scores. Thus, the results of this study indicated a difference in viewpoint between younger and older students.

While the high-stakes test in Florida has recently changed to the Florida Standards Assessment, the use of high-stakes testing has largely remained the same. The FCAT was originally intended to measure student achievement of the Sunshine State Standards, but the use of the assessment has changed over time. For the years that the participants in this study had been in school, the test was, in part, used as a gatekeeper, and students had to pass the test to receive a standard high school diploma, regardless of their grades in school.

High-stakes tests in Florida have also been and continue to be used to assign grades to schools, which can affect school funding from the state. Many parents, students, and teachers in FL and nationwide have protested an overuse of instructional time for the purpose of test preparation (e.g., Brown, 2015; Jacobowitz, 2015; Shammas, 2016; WESH, 2015). In 2015, a bill was passed in Florida to limit time spent on standardized tests, and it received mixed reviews from lawmakers, parents, and teacher organizations, with some seeing it was a step in the right direction and others lamenting its insufficiencies (Cotterell, 2015; WESH, 2015), possibly due to the increased pressures resulting from the use of the test.

The use of such tests is particularly significant for culturally and linguistically diverse students. For these students, high-stakes tests can pose significant and specific challenges, over and above those for many other students (Coltrane, 2003; Hopkins et al., 2013; Huang et al, 2012; Luykx et al., 2007, Mahon, 2006; Murphy, 2007, Robinson-Cimpian, 2016; Sanchez et al., 2009; Solórzano, 2008). Students from diverse backgrounds, who are also ELLs, are particularly susceptible to these pressures, as high-stakes testing in English can become a test of English proficiency, rather than a test of content knowledge (Tsang et al., 2008), thus rendering it impossible to meet the objectives of the test (Hopkins et al, 2013; Robinson-Cimpian et al., 2016). Culturally and linguistically diverse students, including ELLs, are even more vulnerable to testing pressures than are many other students.

While it is not surprising that middle and high school students might view a required, time-consuming, high-stakes test negatively, the reasons for such a view should be considered as significant. Many participants were concerned about the use of the test to determine graduation from high school and future educational opportunities. Many participants were also concerned that the test did not allow them to show what they knew and could do and that much of what was on the test had not been taught to them. One student said, "We should have a better way of testing." Another stated, "It's a useless test which does not properly evaluate students."

Many participants were clearly affected by the pressure placed on them to pass the test. One suggestion from participants was to reduce this pressure, possibly by removing it as a graduation requirement. The issue of pressure was a recurring one in both the surveys and the focus group interviews and one that seemed to color students' views of the test. This pressure is seen by some to be the result of changes in the use of the results of the test, which has also resulted in high pressure for administrators and teachers in addition to the students (ESOL teacher, name excluded upon request, personal communication, March 20, 2015).

When asked to indicate what might be helpful in preparing effectively for the high-stakes assessment, including what participants themselves can do to prepare, responses to these items provided strikingly limited information, and the nonspecific, superficial nature of many responses (e.g., "study more," "sleep," "classes help me learn") as well as the lack of responses was meaningful. There was a plethora of nonspecific responses and a lack of responses to these types of questions across questions and measures, and when responses were provided, they tended to be superficially considered and non-specific.

The relevance of high-stakes tests to students' learning and lives, academically and otherwise, was not clear for many participants. It is important to note that according to previous research the relevance a student assigns to a testing task can predict test performance (Roberts &

Table 5
Themed Perceptions of How the FCAT Helps/Does Not Help in Academic Classes

Participant survey responses	*n*	*%*
No responses/don't know	21	34%
Reasons why the FCAT helps		
The FCAT shows me what I need to learn.	9	15%
I have to work harder and pay attention.	4	6%
There is less stress when the test is over.	3	5%
The FCAT is a review of class content.	3	5%
The FCAT is preparation for the following year.	2	3%
The FCAT shows what I know.	2	3%
The FCAT gives teachers information on class placement.	1	2%
The FCAT gives me confidence.	1	2%
Reasons why the FCAT does not help in class		
The FCAT does not apply to classwork.	11	18%
The FCAT makes me feel like I have to know everything.	3	5%
I already know the answers.	2	3%

Dansereau, 2008), so these disconnects for students may affect their performance on the test as well as in their classes.

The lack of responses to items asking students to indicate the best way to prepare for a high-stakes test as well as how to effectively increase their achievement on the test may be an indication of limitations in metacognitive skills and a lack of awareness of diverse learning strategies. More specifically, the superficiality as well as lack of responses on these items cannot be seen as indicative of interrupted completion of the survey, as the same participants provided responses to subsequent items. It is possible that participants skipped these items due to limitations in their ability to know, process, and convey information about what they know. In other words, it is possible that participants lack metacognitive awareness to know and to be able to explain what would be most helpful to them both in preparation for a high-stakes test as well as in processing information about their learning and abilities in relation to the test.

Many participants reported spending time in class on test preparation involving drill and practice and on test-based workbooks. Although the research on effects of drill and practice test preparation for elementary, middle, and high school students is limited, the indications are that such types of test preparation may not improve student test results. In one study, comparisons among test preparation procedures for 3rd and 5th grade students found that teaching-to-the-test-style preparation did not improve students test performance over teaching content guided by state standards (Welsh, Eastwood, & D'Agostino, 2014).

For older learners, completion of practice tests and tutoring for college-level admission tests have been shown to have little to no affect on test scores (Briggs, 2007, 2009; Scholes & Lain, 1997). Drill and practice methods for university students may be more effective when students are engaged with the material (Kamarulzaman & Shaari, 2015). However, "deliberate practice" has been shown to be insufficient alone and affected by individual factors in improving certain skills (Campitelli & Gobet, 2011). While these drill and practice types of activities may help students to become more familiar with the format of the test, they may also take time away from instruction that enhances critical thinking and problem-solving skills while supporting student learning regarding the standards.

Metacognitive skills can be increased through classroom instruction (Tok, 2013; Vrieling, Bastiaens, & Stijnen, 2012) as can critical thinking skills (Bangert-Drowns, & Bankert, 1990; Davoudi & Sadeghi, 2015; Wismath, Orr, & Good, 2014). Metacognitive instruction has been shown to significantly improve the ability of low-achieving students', including many of Hispanic origin, to solve math problems (Cardelle-Elawar, 1992). There is some indication that metacognitive skills develop along with intellectual ability, but do not completely depend on intellectual ability, and are a stronger predictor of learning performance than is intelligence (Veenman, Wilhelm, & Beishuizen, 2004; Veenman & Spaans, 2005). It is quite possible that spending academic time on teaching students to think about their own learning, rather than conducting a drill and practice type of test preparation may be a more effective way to help students raise test achievement.

Implications and Recommendations

Responses from study participants, combined with the superficiality and lack of specific responses in many cases, seem to indicate that students may lack the metacognitive skills to assess both their preparation prior to testing as well as that which would help improve their preparation. It is possible that, with the pressure associated with the use of high-stakes test scores for school administrators, teacher, and students, the focus on test preparation drill may be taking time away from instruction that would enhance critical thinking skills, even while such instruction can also help improve academic performance (Bangert-Drowns, & Bankert, 1990; Cardelle-Elawar, 1992; Davoudi & Sadeghi, 2015; Tok, 2013; Veenman & Spaans, 2005; Vrieling et al., 2012; Wismath et al., 2014) and thus could, more directly, improve high-stakes testing scores.

This potential is particularly important for students who are more challenged by a high-stakes test, such as culturally and linguistically diverse students. While this study addresses culturally and linguistically diverse students, the results may not be exclusive to them. More research in this area is needed, especially specific to diverse and ELL students.

Assisting students to make sense of the relevance of high-stakes testing to their lives may also help increase test performance. When there is such a disconnect for students, their test performance may be affected (Roberts & Dansereau, 2008). The initial, intended use of the FCAT as an assessment of student achievement of the Sunshine State Standards, was obviously unclear to participants in this study. Helping students to see connections between the content area standards, instruction and learning in class, and high-stakes testing may be a valuable use of time with an eye to increasing test performance.

Re-examination of the effects and the effectiveness of employing high-stakes accountability measures in ways that decrease instructional time, that place more of a burden on specific groups of students, and that fail to encourage and reinforce critical thinking and connection-building skills must be undertaken. Research evidence supporting such a re-examination is readily available for the consideration of policymakers. While accountability of students and teachers is politically popular and can sometimes be educationally appropriate, accountability that is effective and fair, in which the objectives of the assessment can be accomplished, and that enhances and improves instruction and learning is both preferred and imperative.

References

Alcázar-Olán, R. J., Deffenbacher, J. L., Escamilla-Tecalco, H. (2016). Developing a valid version of an inventory to measure anger in Mexican adolescents of middle school level: The ML-STAXI-MS. *Youth & Society, 48*(1), 126-143.

Alkharusi, H. (2013). Canonical correlational models of students' perceptions of assessment tasks, motivational orientations, and learning strategies. *International Journal of Instruction, 6*(1), 1308-1470.

Bangert-Drowns, R. L., & Bankert, E. (1990, April). *Meta-analysis of effects of explicit instruction for critical thinking*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA. ERIC ID# ED328614.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist, 44*, 1175-1184.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248-287.

Briggs, D. C. (2007). *The effects of admissions test preparation: Evidence from NELS: 88*. Retrieved from http://epicpolicy.org/files/Briggs_Theeffectofadmissionstestpreparation.pdf

Briggs, D. (2009). *Preparation for college admission exams*. 2009 NACAC Discussion Paper. National Association for College Admission Counseling. Retrieved from http://www.nacacnet.org/PublicationsResources/Research/Documents/TestPrepDiscussionPaper.pdf

Brown, E. (2015, March 7). Some parents across

the country are revolting against standardized testing. *The Washington Post*. Retrieved from https://www.washingtonpost.com/local/education/some-parents-across-the-country-are-revolting-against-standardized-testing/2015/03/05/e2abd062-c1e1-11e4-9ec2-b418f57a4a99_story.html

Butler, D. L. (2002). Qualitative approaches to investigating self-regulated learning: contributions and challenges. *Educational Psychologist, 37*, 1, 59-63.

California Department of Education. (2015). *California high school exit examination* (CAHSEE). Retrieved from http://www.cde.ca.gov/ta/tg/hs/index.asp

California Department of Education. (2015b). *California high school exit exam* (CAHSEE) state demographic summary report. Retrieved from http://cahsee.cde.ca.gov/ExitProf1.asp?cLevel=State&cYear=2014-15&cChoice=ExitProf1&cAdmin=C&tDate=000000&TestType=E&cGrade=10&Pageno=1

California Department of Education. (2015). *Facts about English learners in California.* Retrieved from http://www.cde.ca.gov/ds/sd/cb/cefelfacts.asp

Campitelli, G., & Gobet, F. (2011). Deliberate practice: Necessary but not sufficient. *Current Directions in Psychological Science, 20*(5), 280-285.

Cardelle-Elawar, M. (1992). *Effects of metacognitive instruction on low-achievers in mathematics problem solving*. Paper presented at the Annual Convention of the American Association for Educational Research, San Francisco, CA.

Center for Disease Control. (2008 July). Data collection methods for program evaluation: Focus groups. *ETA Evaluation Briefs*. Retrieved from http://www.cdc.gov/healthyyouth/evaluation/pdf/brief13.pdf

Center on Education Policy. (2009). *State high school exit exams: Trends in test programs, alternate pathways, and pass rates*. Retrieved from www.cep-dc.org/cfcontent_file.cfm?Attachment=CEP_HSEE09Report_up011310.pdf

Center on Education Policy. (2012). *State high school exit exams: A policy in transition.* Retrieved from http://www.cep-dc.org/displayDocument.cfm?DocumentID=408

Coltrane, B. (2003). English language learners and high-stakes tests. An overview of the issues. *ERIC Digest*. Retrieved from http://www.ericdigests.org/2003-4/high-stakes.html

Cotterell, B. (2015 April 14). New law reduces testing at Florida schools. *Reuters News Service*. Retrieved from http://www.reuters.com/article/us-usa-florida-education-idUSKBN0N52F720150414

Council of the Great City Schools. (2015). *Student testing in American's great city schools: An inventory and preliminary analysis*. Retrieved from www.cgcs.org/cms/lib/DC00001581/.../87/Testing%20Report.pdf

Davoudi, M., & Sadeghi, N. A. (2015). A systematic review of research on questioning as a high-level cognitive strategy. *English Language Teaching, 8*(10), 76-90.

De Groot, E. V. (2002). Learning through interviewing: Students and teachers talk about learning and schooling. *Educational Psychologist, 37*(1), 41-52.

Florida Department of Education. (2008). Graduation and dropout rates by district, 2005-06 school year. Education Information & Accountability Services, Data Publications & Reports. Retrieved from http://www.fldoe.org/eias/eiaspubs/2006grad.asp

Florida Department of Education. (2009). *Student performance results: Demographic report*. Retrieved from https://app1.fldoe.org/FCATDemographics/

Florida Department of Education. (2015a). *Florida's single-year dropout rates*. Retrieved from http://www.fldoe.org/core/fileparse.php/7584/urlt/Dropout-Rate-Report-1314-jtk.pdf

Florida Department of Education. (2015b). *Graduation requirements for Florida's statewide assessments*. Retrieved from http://www.fldoe.org/core/fileparse.php/3/urlt/fcatpass.pdf

Florida Department of Education. (2016a). *Student enrollment*. Retrieved from https://edstats.fldoe.org/SASWebReportStudio/gotoReportSection.do?sectionNumber=1

Florida Department of Education. (2016b). *Students performance results: State reading demographic report* (interactive report). Retrieved from https://app1.fldoe.org/FCATDemographics/Selections.aspx?reportTypeID=6&level=State&subj=Reading

Gaier, E. L. (1962). Students' perceptions of factors affecting test performance. *Journal of Educational Research, 55*(10), 561-566.

Geldhof, G. J., Gestsdottir, S., & Stefansson, K. (2015). Selection, optimization, and compensation: The structure, reliability, and validity of forced-choice versus Likert-type measures in a sample of late adolescents. International *Journal of Behavioral Development, 39*(2), 171-185.

Giambo, D. (2010). High stakes testing, high school graduation, and limited English proficient students: A case study. *American Secondary Education Journal, 38*(1), 44-56.

Hong, E., & Peng, Y. (2008). Do Chinese students' perceptions of test value affect test performance? Mediating role of motivational and metacognitive regulation in test preparation. *Learning and Instruction, 18*, 499-512.

Hopkins, M., Thompson, K. D., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance: A key issue in ESEA reauthorization. *Educational Researcher, 42*(2), 101-108.

Huang, J., Han, T., & Schnapp, K. (2012). Do high-stakes test really address English language learners' learning needs? A discussion of issues, concerns, and implications. *International Journal of Learning & Development, 2*(1). Retrieved from http://www.macrothink.org/journal/index.php/ijld/article/view/1472/1188

Israel, G. D., & Galindo-Gonzalez, S. (2014). Using focus group interviews for planning or evaluating extension programs. *UF/IFAS Extension Document* AEC387. Retrieved from http://edis.ifas.ufl.edu/pdffiles/PD/PD03600.pdf

Jacobowitz, R., & Tobin, K. T. (2015). *Time on test: The fixed costs of 3-8 standardized testing in New York State*. New Paltz, NY: State University of New York at New Paltz: The Benjamin Center for Public Policy Initiatives. Retrieved from http://www.newpaltz.edu/media/the-benjamin-center/Time%20on%20Test%20-%20The%20Fixed%20Costs%20of%203-8%20Standardized%20Testing%20in%20New%20York%20State.pdf

Kamarulzaman, M. S., & Shaari, A. J. (2015). Comparing the effect of EDPA and FDPA on university students' examination results. *Malaysian Online Journal of Educational Technology, 3*(1), 28-34.

Krashen, S. (2012a, July 25). How Common Core will change testing in schools. *The Washington Post*. Retrieved from https://www.washingtonpost.com/blogs/answer-sheet/post/how-common-core-will-change-testing-in-schools/2012/07/25/gJQAzIJB9W_blog.html

Krashen, S. (2012b, July 21). Letter to the editor. *New York Times* Sunday dialogue: Improving our schools. Retrieved from http://www.nytimes.com/2012/07/22/opinion/sunday/sunday-dialogue-improving-our-schools.html?_r=0

Layton, L. (2015, October 24). Study says standardized testing is overwhelming nation's public schools. *The Washington Post*. Retrieved from https://www.washingtonpost.com/local/education/study-says-standardized-testing-is-overwhelming-nations-public-schools/2015/10/24/8a22092c-79ae-11e5-a958-d889faf561dc_story.html

Lazarín, M. (2014). *Testing overload in America's schools*. The Center for American Progress. Retrieved from https://www.americanprogress.org/issues/education/report/2014/10/16/99073/testing-overload-in-americas-schools/

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*, 140, 55.

Lizzo, A., Wilson, K., & Simons, R. (2002). University student perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27*(1), 27-52.

Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record, 109*(4), 897-926.

Mahon, E. A. (2006). High-stakes testing and English language learners: Questions of validity. *Bilingual Research Journal, 30*(2), 479-497.

Maurer, T. J., & Andrews, K. D. (2000). Traditional, Likert, and simplified measures of self-efficacy. *Educational and Psychological Measurement, 60*(6), 965-73.

Merrow, J. (2004). Disappearing dropouts. *The Jim Leher News Hour.* Retrieved from www.merrow.org/podcast/01.mp3

Murphy, S. (2007). Culture and consequences: The canaries in the coal mine. *Research in the Teaching of English, 42*(2), 228-244.

National Center for Education Statistics. (2011-12). *Table 5.5: State high school exit exams, by*

exam characteristics and state, 2011-12. U.S. Department of Education, Institute of Education Sciences. Retrieved from http://nces.ed.gov/programs/statereform/tab5_5.asp

National Center for Education Statistics. (2015a). *Table 204.20: Number and percentage of public school students participating in programs for English language learners, by state: Selected years, 2002-03 through 2012-2013*. Retrieved from https://nces.ed.gov/programs/digest/d14/tables/dt14_204.20.asp

National Center for Education Statistics. (2015b). *Table 1. Public high school 4-year adjusted cohort graduation rate (ACGR), by race/ethnicity and selected demographics for the United States, the 50 states, and the District of Columbia: School year 2013–2014.* Retrieved from https://nces.ed.gov/ccd/tables/ACGR_RE_and_characteristics_2013-14.asp

National Center for Education Statistics. (2016). *The nation's report card* (Interactive report). Retrieved from http://www.nationsreportcard.gov/dashboards/report_card.aspx#

Page-Bucci, H. (2003). *The value of Likert scales in measuring attitudes of on-line learners*. Retrieved from http://www.hkadesigns.co.uk/websites/msc/reme/likert.htm

Patrick, H. & Middleton, M. J. (2002). Turning the kaleidoscope: What we see when self-regulated learning is view with a qualitative lens. *Educational Psychologist, 37*(1), 27-40.

Purpura, J. E. (1997). An analysis of the relationships between test-takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning, 47*(2), 289-325.

Ravich, D. (2011). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Ravitch, D. (2016a, January 29). *District administration poll: Significant resistance to testing*. Retrieved from http://dianeravitch.net/2016/01/29/district-administration-poll-significant-resistance-to-testing/

Ravitch, D. (2016b, January 26). *Fairtest: The fight against misuse of testing continues*. Retrieved from http://dianeravitch.net/2016/01/26/fairtest-the-fight-against-misuse-of-testing-continues/

Roberts, F. W., & Dansereau, D. F. (2008). Studying strategy effects on memory, attitudes, and intentions. *Reading Psychology, 29*(6), 552-580.

Robinson-Cimpian, J. P., Thompson, K. D., & Umansky, I. L. (2016). Research and policy considerations for English learner equity. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 129-137.

Sanchez, M. T., Ehrlich, S., Midouhas, E., & O'Dwyer, L. (2009). *Analyzing performance by grade 10 Hispanic high school students on the Massachusetts State Assessment: Issues and answers*. REL 2009-No. 071. Retrieved at http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2009071.pdf

Scholes, R., & Lain, M. M. (1997, March 24-28). *The effects of test preparation activities on ACT assessment scores.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. ERIC # ED409341.

Shammas, B. (2016, March 4). Frustrated south Florida families opting out of state exams. *Sun Sentinel*. Retrieved from http://www.sun-sentinel.com/local/palm-beach/fl-opt-out-movement-20160304-story.html

Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*(2), 20-329.

Texas Education Agency. (2014). *Enrollment in Texas public schools 2013-2014*. Retrieved from http://tea.texas.gov/acctres/enroll_index.html

Texas Education Agency. (2015). *Fall 2015 EOC: Combined summary reports*. Retrieved from http://tea.texas.gov/Student_Testing_and_Accountability/Testing/State_of_Texas_Assessments_of_Academic_Readiness_(STAAR)/STAAR_Statewide_Summary_Reports_2015-2016/

Texas Education Agency. (2016a). *State assessments: Past, present, and future*. Retrieved from tea.texas.gov/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=25769825328&libID=25769825424

Texas Education Agency. (2016b). *State graduation requirements*. Retrieved from http://tea.texas.gov/graduation.aspx

Tok, S. (2013). Effects of the know-want-learn strategy on students' mathematics achievement, anxiety and metacognitive skills. *Metacognition and Learning, 8*(2), 193-212.

Tsang, S. L., Katz, A., & Stack, J. (2008). Achieving testing for English language learners, ready or not? *Education Policy Analysis Archives, 16*(1). Retrieved from http://epaa.asu.edu/epaa/v16n1/

Veenman, M.V.J. & Spaans, M.A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning & Individual Differences, 15*(2), 159-176.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*(1), 89-109.

Vrieling, E., Bastiaens, T., & Stijnen, S. (2012). Consequences in increased self-regulated learning opportunities on student teachers' motivation and use of metacognitive skills. Australian *Journal of Teacher Education, 37*(8). Retrieved from http://files.eric.ed.gov/fulltext/EJ995229.pdf

Wallace, K. (2015, April 24). Parents all over U.S. 'opting out' of standardized student testing. *CNN*. Retrieved from http://www.cnn.com/2015/04/17/living/parents-movement-opt-out-of-testing-feat/

Welsh, M. E., Eastwood, M., & D'Agostino, J. V. (2014). Conceptualizing teaching to the test under standards-based reform. *Applied Measurement in Education, 27*(2), 98-114.

WESH Orlando. (2015, April 14). *Florida governor signs bill to limit standardized tests*. Retrieved from http://www.wesh.com/news/florida-gov-signs-bill-to-limit-standardized-tests/32360186

Wismath, S., Orr, D., & Good, B. (2014). Metacognition: Student reflections on problem solving. *Journal of Excellence in College Teaching, 25*(2), 69-90.

Zhu, W., & Flaitz, J. (2005). Using focus group methodology to understand international students' academic language needs: A comparison of perspectives. *Teaching English as a Second or Foreign Language, 8*(4). Retrieved from http://tesl-ej.org/ej32/a3.html.