

การเปรียบเทียบคุณภาพแบบทดสอบโคลซใน การวัดความเข้าใจในการอ่านภาษาไทย
ชั้นประถมศึกษาปีที่ 6 โดยใช้วิธีในการตัดคำและเกณฑ์การให้คะแนน
ที่แตกต่างกัน

A comparison of Quality Testing of Cloze Test in Measuring Reading
Comprehension in Thai Subject for Prathomsuksa 6 Students through
Using Different Words Deletion Methods and Scoring Criteria

พิมรดา แสงภรา¹
สมนึก ภัททิยธนี²
โสภี อุ้นทะยา³

บทคัดย่อ

การวัดความเข้าใจในการอ่านเป็นเรื่องสำคัญในการเรียนรู้ด้านภาษาช่วยให้ทราบความสามารถในการอ่าน เพิ่มความเข้าใจได้เป็นอย่างดี แบบทดสอบโคลซเป็นแบบทดสอบประเภทหนึ่งที่เหมาะสมนำไปใช้วัดความเข้าใจในการอ่านภาษาไทยของนักเรียน การวิจัยครั้งนี้มีความมุ่งหมายเพื่อเปรียบเทียบคุณภาพของแบบทดสอบโคลซที่ใช้วัดความเข้าใจในการอ่านภาษาไทยที่มีวิธีการตัดคำและเกณฑ์การให้คะแนนที่แตกต่างกันโดยใช้แบบทดสอบโคลซ จำนวน 9 ฉบับ เปรียบเทียบคุณภาพของแบบทดสอบโคลซในด้านค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยงตรง ค่าความเชื่อมั่นและค่าความคลาดเคลื่อนในการวัดเกณฑ์ในการให้คะแนน มี 2 เกณฑ์ ได้แก่ เกณฑ์ที่ 1 การให้คะแนนกับคำตอบที่ตรงกับต้นฉบับหรือคำเฉลยนั้น และเกณฑ์ที่ 2 เป็นการให้คะแนนคำที่มีความเหมาะสมใกล้เคียงกับคำเฉลยด้วย กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นครั้งนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 จำนวน 270 คน ได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling) เครื่องมือที่ใช้ในการเก็บข้อมูล ได้แก่ แบบทดสอบชนิดเติมคำจำนวน 9 ฉบับ โดยฉบับที่ 1-3 ฉบับละ 67 ข้อ ฉบับที่ 4-6 ฉบับละ 47 ข้อ ฉบับที่ 7-9 ฉบับ 44 ข้อ สถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ การวิเคราะห์ความแปรปรวนทางเดียว (One - Way Analysis of Variance) ไคสแควร์ (Chi - Square Test) และวิธีการทดสอบค่าที (t-test) แบบกลุ่มตัวอย่างไม่อิสระ (Dependent Samples)

ผลการวิจัยปรากฏดังนี้

1. คุณภาพของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์การให้คะแนนทั้งสองเกณฑ์มีคุณภาพด้านต่างๆผลเป็นดังนี้

1.1 คุณภาพด้านค่าความยากและอำนาจจำแนกอยู่ในเกณฑ์ที่เหมาะสม โดยแบบทดสอบโคลซที่มีวิธีการตัดคำที่แตกต่างกันที่ให้คะแนนโดยเกณฑ์ที่ 1 มีค่าความยากตั้งแต่ .388 ถึง .603 และมีอำนาจจำแนกตั้งแต่ .221 ถึง .323 ส่วนแบบทดสอบโคลซที่ให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความยากตั้งแต่ .418 ถึง .621 มีอำนาจจำแนกตั้งแต่ .261 ถึง .352

¹ นิสิตปริญญาโท สาขาวิชาการวัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

² รองศาสตราจารย์ประจำภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

³ อาจารย์ ดร. ประจำภาควิชาภาษาไทย คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยมหาสารคาม

1.2 คุณภาพด้านค่าความเที่ยงตรงอยู่ในเกณฑ์ที่เหมาะสมโดยแบบทดสอบโคลซ ที่มีวิธีการตัดคำที่แตกต่างกันที่ให้คะแนนโดยเกณฑ์ที่ 1 ค่าความเที่ยงตรง มีค่าตั้งแต่ .331 ถึง .748 ส่วนแบบทดสอบโคลซที่ให้คะแนนโดยใช้เกณฑ์ที่ 2 ค่าความเที่ยงตรง มีค่าตั้งแต่ .516 ถึง .812

1.3 คุณภาพด้านความเชื่อมั่นอยู่ในเกณฑ์ที่เหมาะสมโดยแบบทดสอบโคลซ ที่มีวิธีการตัดคำที่แตกต่างกันที่ให้คะแนนโดยเกณฑ์ที่ 1 มีค่าความเชื่อมั่นตั้งแต่ .618 ถึง .952 ส่วนแบบทดสอบโคลซที่ให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความเชื่อมั่นตั้งแต่ .731 ถึง .972

1.4 คุณภาพด้านค่าความคลาดเคลื่อนมาตรฐานอยู่ในเกณฑ์ที่เหมาะสม โดยแบบทดสอบโคลซที่มีวิธีการตัดคำที่แตกต่างกันที่ให้คะแนนโดยเกณฑ์ที่ 1 มีค่าความคลาดเคลื่อนมาตรฐานในการวัดตั้งแต่ 2.169 ถึง 4.501 ส่วนแบบทดสอบโคลซที่ให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความคลาดเคลื่อนมาตรฐานในการวัดตั้งแต่ 2.216 ถึง 4.642

2 การทดสอบนัยสำคัญของแบบทดสอบโคลซที่ตรวจโดยใช้เกณฑ์การให้คะแนนที่แตกต่างกันผลเป็นดังนี้

2.1 ค่าความยากมาตรฐานของแบบทดสอบโคลซแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยแบบทดสอบโคลซที่ตรวจโดยใช้เกณฑ์ที่ 1 จะยากกว่าการตรวจโดยใช้เกณฑ์ที่ 2 ทั้ง 9 ฉบับ

2.2 ค่าอำนาจจำแนก ค่าความเที่ยงตรงของแบบทดสอบโคลซที่มีวิธีการตัดคำ และเกณฑ์การให้คะแนนที่แตกต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนค่าความเชื่อมั่นของแบบทดสอบโคลซที่มีวิธีการตัดคำแตกต่างกัน มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเกณฑ์การให้คะแนนที่แตกต่างกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2.3 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 และตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และพบว่าการตรวจโดยใช้เกณฑ์ให้คะแนนที่แตกต่างกัน จะทำให้ค่าความแปรปรวนของแบบทดสอบโคลซมีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติทั้ง 9 ฉบับ

โดยสรุปในการวิจัยครั้งนี้ พบว่า การตัดคำแบบกำหนดอัตราส่วนตายตัวและมีการชี้แนะ โดยกำหนดจำนวนคำของคำตอบ เป็นแบบทดสอบที่มีความเหมาะสม ส่วนการเลือกใช้เกณฑ์การตรวจให้คะแนนควรเลือกใช้ตามวัตถุประสงค์ของการวัด ถ้าเป็นการสอบคัดเลือกหรือสอบแข่งขันเพื่อจัดอันดับ ควรใช้เกณฑ์ที่ 1 ส่วนการสอบวัดผลประสงค์การเรียนรู้ สอบกลางปี หรือปลายปีการศึกษา ควรตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 จะมีความเหมาะสม

คำสำคัญ : แบบทดสอบโคลซ

ABSTRACT

The measurement of reading comprehension is important in learning literacy because it can help both reading ability and enhancing at good reading comprehension. Cloze test is one of the tests that are appropriate to measure students in Thai reading comprehension. The purposes of the study were to compare the quality of cloze tests in measuring reading comprehension in Thai subject which has different words deletion method and scoring criteria by using nine types of comparing quality of the cloze test in difficulty, discrimination, validity, reliability and deviation of measurement. There are two types of scoring criteria namely, the first criteria is giving scores related to the correct answers and the second is giving scores that is the appropriate words or nearly with the answer keys. The samples were 270 of Prathomsuksa 6 students in the second semester of the 2012 academic year from Roi-et primary educational service area office 1 which were selected by stratified random sampling technique. The tools of data collection namely nine types of cloze tests by 1st -3rd papers have 67 items, 4th -6th papers have 47 items, 7th -9th have 44 items. The statistics of data analysis are One-Way Analysis of Variance, Chi-Square Test, testing method is t-test of dependent samples

The results of the study were as follows:

1. The quality of all nine tests which checked and given scores are shown the quality as follows :

1.1 The quality of difficulty and discrimination are in appropriate criteria by using cloze test of different words deletion given scores and used first criteria. The difficulty value starts at .388 to .603 and discrimination value starts at .221 to .323 cloze test of words deletion given scores by using the second criteria which difficulty value starts at .418 to .621. and the discrimination value starts at .261 to .352.

1.2 The quality of validity value is in appropriate criteria by using cloze test of different words deletion given scores and used first criteria. The validity value starts at .331 to .748. The cloze test of different words deletion given scores by using second criteria of validity value starts at .516 to .812

1.3 The quality of reliability value is in appropriate criteria by using cloze test of different words deletion and used first criteria. The reliability value starts at .618 to .952 and cloze test given scores by using second criteria has reliability value at .731 to .972.

1.4 The quality of standard deviation is in appropriate criteria using cloze test of different words deletion given scores and used first criteria has standard deviation in measuring started at 2.169 to 4.501. The cloze test of different words deletion given scores by using second criteria of standard deviation in measuring starts at 2.216 to 4.642

2. The significance test of cloze test of different words deletion by giving scores criteria as follows:

2.1 The standard difficulty value of cloze tests with different scoring methods were significantly different .01 level by using the first cloze test criteria which is more difficult than using the second criteria.

2.2 The discrimination value. The validity of cloze test of words deletion was not statistical significantly different difference of giving score criteria and reliability value was significantly different .01 levels and was significantly different of giving score criteria.

2.3 The standard deviation, there are overall nine types of testing measurement using both first and second scoring method criteria were statistical significantly different .01 level and found that the checking of different criteria makes variance value of cloze test different which was not statistical significantly different of nine types method criteria.

In this research conclusion found that cloze test of words deletion fixed proportion and suggestion by fixing word number answers is appropriate test. The selection of giving scores criteria should select on purpose of measurement.

Keyword : Cloze Test

บทนำ

ภาษาไทยเป็นเอกลักษณ์ของชาติเป็นสมบัติทางวัฒนธรรมอันก่อให้เกิดความเป็นเอกภาพ และเสริมสร้างบุคลิกภาพของคนในชาติให้มีความเป็นไทย เป็นเครื่องมือในการติดต่อสื่อสารเพื่อสร้างความเข้าใจและความสัมพันธ์ที่ดีต่อกัน ทำให้สามารถประกอบกิจกรรม การงาน และดำรงชีวิตร่วมกันในสังคมประชาธิปไตยได้อย่างสันติสุข และเป็นเครื่องมือในการแสวงหาความรู้ (สำนักวิชาการและมาตรฐานการศึกษา. 2552 : 1-4) การอ่านเป็นสิ่งสำคัญอย่างหนึ่งในชีวิต ของคนเราเป็นเครื่องมือในการแสวงหาความรู้ ผู้ที่มีนิสัยรักการอ่านและมีทักษะในการอ่าน ย่อมแสวงหาความรู้และศึกษาเล่าเรียนได้อย่างมีประสิทธิภาพ สามารถนำความรู้ที่ได้จากการอ่านไปใช้ในการพูดและเขียนได้อย่างดีการอ่านจึงจัดเป็นทักษะสำคัญในการเรียนทุกระดับ การอ่านนับว่ามีคุณค่าและเป็นสิ่งที่ควรได้รับการฝึกฝนให้มีทักษะเพิ่มขึ้นคือ การอ่านเพื่อความเข้าใจเพราะถ้านักเรียนเข้าใจเรื่องที่อ่านได้ถูกต้องก็จะได้รับความรู้จากเรื่องที่อ่านได้เต็มที่ การประเมินความสามารถการอ่านของนักเรียนอยู่เสมอ จะช่วยให้เข้าใจนักเรียนและรู้ถึงปัญหาที่เกี่ยวข้องกับการอ่าน วิธีการประเมินการอ่านอาจทำได้หลายวิธี ในต่างประเทศจะมีแบบทดสอบมาตรฐาน ครูสามารถนำแบบทดสอบมาทดสอบเพื่อประเมินระดับการอ่านของนักเรียนแต่ในประเทศไทยยังขาดแบบทดสอบมาตรฐานที่ใช้วัดความเข้าใจการอ่าน โดยเฉพาะ (อรนุช วรอำศวปติ. 2530 : 2) ในการวัดความสามารถในการอ่านของนักเรียนต้องใช้แบบทดสอบที่ครูสร้างขึ้นเอง ซึ่งครูส่วนใหญ่ยังขาดความรู้ ความชำนาญในการวัดผล มักเลือกสร้างเฉพาะข้อสอบแบบเลือกตอบ ที่มีความสะดวกในการใช้ การตรวจให้คะแนนและมีความเป็นปรนัยสูงแต่ต้องใช้เวลาในการสร้าง มีขั้นตอนการสร้างที่ละเอียดจึงเป็นปัญหาสำหรับครูที่จะสร้างแบบทดสอบแบบเลือกตอบที่มีคุณภาพ (ทรงศักดิ์ ภูสีอ่อน. 2542 : 2) ได้มีผู้นำเอาวิธีการประเมินความสามารถในการอ่านที่สร้างได้ง่าย รวดเร็ว ประหยัดค่าใช้จ่าย (สุพร พันธุ์สุภา. 2539 : 3) มาใช้ในการประเมินความสามารถในการอ่าน เรียกว่าแบบทดสอบโคลซ (Cloze Procedure) หรือแบบทดสอบโคลซ (Cloze Test) เป็นแบบทดสอบแบบเติมคำ (Completion Test) ชนิดหนึ่ง ผู้ออกข้อสอบจะเลือกบทความที่มีความสมบูรณ์ในตัวและมีความยาวเพียงพอ

ทำการตัดคำในข้อความออกจาดตัดโดยใช้อัตราส่วนตายตัวหรือไม่ตายตัวก็ได้ นำมาให้ผู้สอบเติมหรือเลือกคำที่กำหนดให้ (บุญชม ศรีสะอาด. 2540 : 78) ซึ่งบอร์มัท (Bormuth. 1967 : 291 - 299) ได้กล่าวถึงแบบทดสอบโคลซว่าเป็นแบบทดสอบที่เชื่อถือได้ มีความเที่ยงตรงสูง สามารถให้คะแนนได้ง่ายและมีความเป็นปรนัยสูง โดยเพลสเตอร์ (Plaister. 1973 : 33) ได้แนะนำครูสามารถแก้ปัญหาในการอ่านเพื่อให้ นักเรียนเกิดความเข้าใจโดยวิธีโคลซ แบบทดสอบโคลซมาตรฐานเดิมจะเป็นการเติมคำลงในช่องว่างที่เว้นจากเนื้อเรื่องทุกคำที่ 5 อย่างเป็นระบบ ต่อมามีการเปลี่ยนแปลงระยะของคำที่เว้นช่องทุกคำที่ 6,7 และ 8 โดยใช้เส้นตรงที่มีความยาวเท่ากันประมาณ 14 ตัวอักษรแทนช่องว่างนั้น ซึ่งแบบทดสอบโคลซแบบนี้มีจุดอ่อนในเรื่องของการให้คะแนน และการควบคุมคุณลักษณะที่ต้องการวัด และที่สำคัญมีงานวิจัยพบว่าตัวแปรบางตัวส่งผลต่อคุณภาพของแบบทดสอบ เช่นจุดเริ่มต้นเว้นคำ (Blank Length) รูปแบบการเว้นคำ (Passage Difficilty) ระยะเวลาเว้นคำ (Deletion Rate) และวิธีการให้คะแนน ซึ่งองค์ประกอบเหล่านี้ส่งผลต่อคุณภาพแบบทดสอบ (กัลยาณี นันทะศรี. 2542 : 4) งานวิจัยที่เกี่ยวข้องกับเกณฑ์การให้คะแนนนั้น ออลเลอร์ (Oller. 1979 : 367- 375) ได้ศึกษาเกี่ยวกับการให้คะแนนในแบบทดสอบโคลซ สรุปว่ากลุ่มที่เป็นเจ้าของภาษาวิธีการให้คะแนนที่ดีที่สุดคือการให้คะแนนเฉพาะคำที่ถูกต้องตามต้นฉบับ ส่วนกลุ่มคนที่ไม่เจ้าของภาษาคอรอนูโลมให้คะแนนแก่คำที่มีความหมายใกล้เคียงกันกับคำที่ถูกต้องนั้นได้ การเลือกใช้เกณฑ์ในการให้คะแนนของแบบทดสอบโคลซ ถ้าเป็นการสอบแข่งขันเพื่อจัดอันดับ ควรใช้เกณฑ์ที่ 1 ซึ่งจะมีค่าอำนาจจำแนกไม่แตกต่างจากเกณฑ์ที่ 2 และมีความสะดวกในการตรวจให้คะแนนมากกว่า แต่ถ้าเป็นการสอบวัดจุดประสงค์ หรือสอบกลางภาค ปลายภาคควรที่จะใช้เกณฑ์การตรวจคะแนนเกณฑ์ที่ 2 เพื่อเป็นการตรวจสอบว่านักเรียนผ่านจุดประสงค์นั้นแล้วหรือยังมีข้อบกพร่องตรงไหนจะมีเหมาะสมมากกว่า (ทรงศักดิ์ ภูสีอ่อน. 2542 : 115)จากงานวิจัยที่เกี่ยวข้องกับแบบทดสอบโคลซที่ใช้วัดความเข้าใจในการอ่านภาษาไทย มีการพัฒนารูปแบบแบบทดสอบโคลซมากมายซึ่งมีทั้งแบบทดสอบโคลซมาตรฐานและแบบทดสอบโมดิฟายด์ โคลซเช่น แบบเลือกตอบ แบบเลือกกลุ่มคำ แบบบอกความหมาย แบบพหุขนะตัวแรกชี้แนะ และแบบมีรูปสระ รูปวรรณยุกต์และไม้ทัณฑฆาตชี้แนะการสร้างแบบทดสอบโคลซไม่มีความยุ่งยาก ประหยัดเวลา ง่ายในการนำไปใช้ ผู้วิจัยจึงสนใจในการสร้างแบบทดสอบที่ใช้การตัดคำแบบแบ่งอัตราส่วนตามลักษณะของการชี้แนะและอัตราส่วนในการตัดคำรวมถึงเกณฑ์การให้คะแนนที่เหมาะสม สามารถนำไปใช้วัดความสามารถในการอ่านภาษาไทยของนักเรียนให้เหมาะสมและมีประสิทธิภาพมากยิ่งขึ้น

ความมุ่งหมายของการวิจัย

1. เพื่อเปรียบเทียบวิธีการตัดคำของแบบทดสอบโคลซทั้ง 9 ฉบับ คือ แบบทดสอบโคลซที่ตัดคำแบบกำหนดอัตราส่วนตายตัวและไม่มีการชี้แนะ แบบทดสอบโคลซที่ตัดคำแบบกำหนดอัตราส่วนตายตัวและกำหนดคำแรก แบบทดสอบที่ตัดคำแบบกำหนดอัตราส่วนตายตัวและกำหนดจำนวนตัวอักษร แบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดอัตราส่วนตายตัวและไม่มีการชี้แนะ แบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดอัตราส่วนตายตัวและกำหนดอักษรตัวแรก แบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดอัตราส่วนตายตัวและกำหนดจำนวนตัวอักษร แบบทดสอบโคลซตัดคำแบบผสมและไม่มีการชี้แนะ แบบทดสอบโคลซตัดคำแบบผสมและกำหนดคำอักษรคำแรก แบบทดสอบโคลซตัดคำแบบผสมและกำหนดจำนวนตัวอักษร
2. เพื่อเปรียบเทียบคุณภาพของแบบทดสอบโคลซทั้ง 9 ฉบับ ในด้าน ค่าความยาก ค่าอำนาจจำแนก ค่าความเชื่อมั่น ค่าความเที่ยงตรงและค่าความคลาดเคลื่อนมาตรฐาน

3. เพื่อเปรียบเทียบเกณฑ์การตรวจให้คะแนนที่ต่างกัน ได้แก่การให้คะแนนเฉพาะคำตอบที่ถูกต้องตรงกับต้นฉบับ (The Exact Word Method) และการให้คะแนนคำที่มีความเหมาะสมหรือใกล้เคียงด้วย (Contextual Appropriateness Method)

ระเบียบวิธีวิจัย

ขอบเขตของการศึกษา

1. ประชากรที่ใช้ในการศึกษาครั้งนี้ เป็นชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในอำเภอจตุรพักตรพิมาน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 จำนวน 826 คน จาก 51 โรงเรียน

2. กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นชั้นประถมศึกษาปีที่ 6 ของโรงเรียนในอำเภอจตุรพักตรพิมาน สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 1 จำนวน 270 คน จาก 51 โรงเรียน ได้มาโดยการสุ่มแบบแบ่งชั้น (Stratified Random Sampling)

3. ตัวแปรที่ใช้ในการวิจัย

3.1 ตัวแปรต้น คือ ลักษณะแบบทดสอบโคลซในการตัดคำและเกณฑ์การตรวจให้คะแนน ได้แก่ ลักษณะแบบทดสอบโคลซในการตัดคำโดยการกำหนดอัตราส่วนและการชี้แนะ และเกณฑ์การให้คะแนน

3.2 ตัวแปรตาม คือ คุณภาพของแบบทดสอบ ได้แก่ ค่าความยากของแบบทดสอบค่าอำนาจจำแนกของแบบทดสอบ ค่าความเชื่อมั่นของแบบทดสอบ ค่าความเที่ยงตรงของแบบทดสอบและค่าความคลาดเคลื่อนมาตรฐานในการวัด

4. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลมี 2 ประเภท คือแบบทดสอบโคลซวัดความเข้าใจในการอ่านจำนวน 9 ฉบับ และแบบทดสอบแบบเลือกตอบจำนวน 1 ฉบับ

5. การวิเคราะห์ข้อมูล

5.1 การวิเคราะห์แบบทดสอบโดยการหา ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าความแปรปรวน

5.2 การหาคุณภาพของแบบทดสอบโดยการหา ค่าความยาก (Difficulty) ค่าความยากมาตรฐาน (Standard Difficulty) ค่าอำนาจจำแนก (Discrimination) ค่าอำนาจจำแนกเฉลี่ยของแบบทดสอบ โดยเปลี่ยนค่าอำนาจจำแนกของข้อสอบเป็นค่าพิชเชอร์ซี หาค่าความเชื่อมั่นของแบบทดสอบ (Reliability) โดยคำนวณจากสูตร KR_{20} ค่าความเที่ยงตรงของแบบทดสอบ (Validity)หาค่าความเที่ยงตรงของแบบทดสอบ โดยการหาสัมประสิทธิ์สหสัมพันธ์อย่างง่ายของเพียร์สัน (Pearson) ระหว่างคะแนนที่ได้จากแบบทดสอบแต่ละฉบับกับคะแนนที่ได้จากแบบทดสอบแบบเลือกตอบและค่าความคลาดเคลื่อนมาตรฐานในการวัด (Standard Error of measurement)

5.3 การวิเคราะห์ความแตกต่างของแบบทดสอบโดยทดสอบความแตกต่างของค่าความยากของแบบทดสอบโคลซทั้ง 9 ฉบับ ด้วยการวิเคราะห์ความแปรปรวนทางเดียว (One - Way Analysis of Variance) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติ จะทดสอบความแตกต่างรายคู่ โดยวิธีของเชฟเฟ (Scheffe) ทดสอบความแตกต่างของค่าอำนาจจำแนก ความเชื่อมั่น และความเที่ยงตรงทั้ง 9 ฉบับ โดยวิธีการทดสอบค่าไคสแควร์ (Chi-Square Test) ถ้าพบความแตกต่างอย่างมีนัยสำคัญทางสถิติ จะทดสอบความแตกต่างรายคู่ ด้วยวิธีทดสอบค่าอัตราส่วนซี (Z) เปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซทั้ง 9 ฉบับโดยการทดสอบความเป็นเอกพันธ์ของความแปรปรวนตามวิธีของฮาร์ดเลย์ (Hartley) ถ้าความแตกต่างมีนัยสำคัญทางสถิติ จะทดสอบความแตกต่างเป็นรายคู่ด้วยวิธีทดสอบค่าอัตราส่วนเอฟ

(F) ทดสอบความแตกต่างของค่าความยากของแบบทดสอบโคลซแต่ละฉบับที่ได้จากเกณฑ์การตรวจคะแนนที่ต่างกัน โดยวิธีการทดสอบค่าที (t-test) แบบกลุ่มตัวอย่างไม่อิสระ (Dependent Sample) การทดสอบความแตกต่างของค่าอำนาจจำแนก ค่าความเชื่อมั่นและค่าความเที่ยงตรง ของแบบทดสอบโคลซในแต่ละฉบับที่ได้จากการตรวจให้คะแนนโดยใช้เกณฑ์ที่ต่างกัน ด้วยวิธีทดสอบค่าอัตราส่วนซี (Z) และทดสอบความแตกต่างของค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซ ในแต่ละฉบับที่ได้จากการตรวจให้คะแนนโดยใช้เกณฑ์ที่ต่างกันด้วยการทดสอบค่าอัตราส่วนเอฟ (F)

สรุปผลการวิจัย

1. แบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 มีค่าตั้งแต่ 22.64 ถึง 40.36 และที่ได้จากการตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ 26.01 ถึง 41.18 เนื่องจากจำนวนข้อสอบของแบบทดสอบโคลซแต่ละฉบับมีจำนวนไม่เท่ากัน จึงพิจารณาเป็นเปอร์เซ็นต์ของคะแนนเฉลี่ยพบว่า ค่าเปอร์เซ็นต์ของคะแนนเฉลี่ยของแบบทดสอบโคลซ ทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 มีค่าตั้งแต่ 51.45 ถึง 60.34 และเปอร์เซ็นต์เฉลี่ยของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ 59.11 ถึง 68.46 ส่วนค่าความเบี่ยงเบนมาตรฐานของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 มีค่าตั้งแต่ 6.4 ถึง 16.22 ส่วนค่าความเบี่ยงเบนมาตรฐานของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ 9.48 ถึง 19.62

2. คุณภาพของแบบทดสอบโคลซทั้ง 9 ค่าความยากเฉลี่ยของแบบทดสอบโคลซ ทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่า มีค่าตั้งแต่ .388 ถึง .603 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ .418 ถึง .621

ค่าอำนาจจำแนกของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่า มีค่าตั้งแต่ .221 ถึง .323 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่มีค่าตั้งแต่ .261 ถึง .352

ค่าความเที่ยงตรงของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่าแบบทดสอบโคลซทั้ง 9 ฉบับ มีค่าความเที่ยงตรงเชิงโครงสร้างตั้งแต่ .331 ถึง .748 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความเที่ยงตรงเชิงโครงสร้างตั้งแต่ .516 ถึง .812

ค่าความเชื่อมั่นของแบบทดสอบโคลซทั้ง 9 ฉบับ แบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่า มีค่าตั้งแต่ .618 ถึง .952 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ .731 ถึง .972

ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซทั้ง 9 ฉบับ แบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่า มีค่าตั้งแต่ 2.169 ถึง 4.501 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ 2.216 ถึง 4.642

3. การเปรียบเทียบคุณภาพของแบบทดสอบโคลซ

การศึกษาเปรียบเทียบค่าความยากมาตรฐานของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 แบบทดสอบโคลซที่ตัดค่าแบบไม่กำหนดสัดส่วนตายตัว กำหนดค่าแรก มีค่าความยากมาตรฐานเฉลี่ยสูงสุด จากการทดสอบความแตกต่างของค่าความยากมาตรฐานเฉลี่ย พบว่า แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อทดสอบความแตกต่างรายคู่ พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 4 คู่ และมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 3 คู่ แบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 แบบทดสอบโคลซที่ตัดค่าแบบไม่กำหนดสัดส่วนตายตัว กำหนดค่าแรก มีค่าความยากมาตรฐานเฉลี่ยสูงสุด แบบทดสอบโคลซที่ตัดค่าแบบผสมและ

อภิปรายผล

1. คุณภาพของแบบทดสอบ

1.1 ค่าความยากเฉลี่ยของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่ามีค่าตั้งแต่ .388 ถึง .603 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ .418 ถึง .621 ซึ่งอยู่ในเกณฑ์ .20 ถึง .08 เป็นแบบทดสอบที่มีคุณภาพ (สมนึก ภัททิยธนี. 2551 : 204)

1.2 ค่าอำนาจจำแนกของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่ามีค่าตั้งแต่ .221 ถึง .323 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ .261 ถึง .352 ซึ่งอยู่ในเกณฑ์ .02 ขึ้นไป (สมนึก ภัททิยธนี. 2551 : 204) ซึ่งแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับสามารถแยกนักเรียนกลุ่มสูงและกลุ่มต่ำได้

1.3 ค่าความเที่ยงตรงของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่าแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ มีค่าความเที่ยงตรงเชิงโครงสร้างตั้งแต่ .331 ถึง .748 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความเที่ยงตรงเชิงโครงสร้างตั้งแต่ .516 ถึง .812 ซึ่งอยู่ในเกณฑ์ .50 ถึง 1.00 (สมนึก ภัททิยธนี. 2551 : 204) แต่มีแบบทดสอบบางฉบับในเกณฑ์การตรวจให้คะแนนเกณฑ์ที่ 1 มีค่าไม่อยู่ในเกณฑ์ อาจเนื่องมาจากวิธีการตรวจให้คะแนนส่งผลต่อความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ

1.4 ค่าความเชื่อมั่นของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ แบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่ามีค่าตั้งแต่ .618 ถึง .952 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ .731 ถึง .972 ซึ่งแบบทดสอบทั้งหมดทั้ง 9 ฉบับที่ตรวจให้คะแนนทั้ง 2 เกณฑ์ มีค่าความเชื่อมั่นสูง มีความเหมาะสมที่จะนำมาวัดความเข้าใจในการอ่านภาษาไทย

1.5 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ แบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ผลปรากฏว่ามีค่าตั้งแต่ 2.169 ถึง 4.501 และแบบทดสอบที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าตั้งแต่ 2.216 ถึง 4.642 ถึงแม้ค่าความเชื่อมั่นของแบบทดสอบมีค่าความเชื่อมั่นสูงแต่ลักษณะการตัดคำของแบบทดสอบอาจส่งผลต่อความคลาดเคลื่อนของแบบทดสอบด้วย ซึ่งผู้สอบเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 มีความสนใจในเรื่องต่างๆ สันนิษฐานการให้เด็กทำข้อสอบที่หาคำมาเติมเองก่อให้เกิดความเบื่อหน่ายจึงส่งผลต่อความคลาดเคลื่อนในการวัด

2. การทดสอบความแตกต่างของแบบทดสอบโคลซ

จากการศึกษาเปรียบเทียบค่าความยากมาตรฐานของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 แบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดสัดส่วนตายตัว กำหนดค่าแรก มีค่าความยากมาตรฐานเฉลี่ยสูงที่สุด จากการทดสอบความแตกต่างของค่าความยากมาตรฐานเฉลี่ย พบว่า แตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 เมื่อทดสอบความแตกต่างรายคู่ พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 4 คู่ และมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 3 คู่ แบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 แบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดสัดส่วนตายตัว กำหนดค่าแรก มีค่าความยากมาตรฐานเฉลี่ยสูงที่สุด แบบทดสอบโคลซที่ตัดคำแบบผสมและไม่มีกรขึ้นแคะคำตอบ มีค่าความยากมาตรฐานเฉลี่ยต่ำที่สุด เมื่อทดสอบความแตกต่างระหว่างค่าความยากมาตรฐานเฉลี่ยของแบบทดสอบโคลซทั้งหมดทั้ง 9 ฉบับ พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 4 คู่ ค่าความยากมาตรฐานเฉลี่ยของแบบทดสอบโคลซแต่ละฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์การตรวจที่แตกต่างกัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เมื่อทดสอบความแตกต่างระหว่างค่าความยากมาตรฐานเฉลี่ยของแบบทดสอบโคลซที่ตรวจให้คะแนนโดยใช้เกณฑ์การตรวจที่แตกต่าง

กัน พบว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ จากผลการวิจัยเป็นเช่นนี้ อาจเนื่องมาจากแบบทดสอบโคลซที่มี การตัดทำแตกต่างกัน มีวิธีการชี้แนะเพื่อช่วยในการตอบของนักเรียนต่างกัน จึงทำให้นักเรียนใช้ความรู้ ความสามารถในการตอบแตกต่างกันส่งผลให้ค่าความยากของแบบทดสอบแตกต่างกัน ซึ่งผลการวิจัยสรุปได้ว่า ลักษณะของแบบทดสอบต่างกันมีผลต่อค่าความยากของแบบทดสอบ

จากการศึกษาเปรียบเทียบค่าอำนาจจำแนกของแบบทดสอบโคลซทั้ง 9 ฉบับ พบว่า ค่าอำนาจจำแนกของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนทั้งสองเกณฑ์ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ อาจเนื่องมาจาก ค่าความยากง่ายเฉลี่ยของแบบทดสอบโคลซแต่ละฉบับที่มีค่าต่ำ นั่นคือแบบทดสอบโคลซแต่ละฉบับนั้นยากมากซึ่งสอดคล้องกับค่ากล่าวของอนันต์ ศรีโสภา (2525 : 205 -222) ที่ว่าองค์ประกอบอย่างหนึ่งที่มีต่ออิทธิพลต่อค่าอำนาจจำแนกของแบบทดสอบคือ ค่าความยากง่ายของแบบทดสอบข้อใดข้อหนึ่งที่มีความยากเกินไปหรือง่ายเกินไปก็มีแนวโน้มที่จะทำให้ข้อสอบนั้นมีค่าอำนาจจำแนกต่ำหรือจะทำให้ไม่สามารถแยกผู้สอบออกได้ว่ากลุ่มใดเก่งหรือกลุ่มใดอ่อนและอาจมีผลเนื่องมาจากว่าแบบทดสอบโคลซทั้ง 9 ฉบับเป็นแบบทดสอบให้หาคำมาเติม และให้ตัวชี้แนะเพื่อช่วยในการตอบของนักเรียนส่งผลให้แบบทดสอบมีค่าอำนาจจำแนกแตกต่างกันน้อย จึงสรุปได้ว่าลักษณะของแบบทดสอบต่างกันและเกณฑ์การให้คะแนนที่ต่างกันไม่มีผลต่อค่าอำนาจจำแนกของแบบทดสอบ

จากการศึกษาเปรียบเทียบค่าความเที่ยงตรงของแบบทดสอบโคลซทั้ง 9 ฉบับ พบว่า ค่าความเที่ยงตรงของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนทั้งสองเกณฑ์ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐานที่ตั้งไว้ การชี้แนะของคำตอบส่งผลให้ค่าความเที่ยงตรงเพิ่มขึ้นและจะเห็นว่าแบบทดสอบโคลซทั้ง 9 ฉบับมีค่าความเที่ยงตรงใกล้เคียงกันมากอาจเนื่องมาจากแบบทดสอบโคลซเป็นแบบทดสอบที่ให้นักเรียนหาคำมาเติมเองนักเรียนเก่งจึงสามารถเข้าใจเรื่องราวได้ดีและเติมคำในช่องว่างได้ถูกต้อง ส่วนนักเรียนที่อ่อน อ่านหนังสือไม่ออกก็จะไม่เข้าใจเนื้อเรื่องที่อ่าน ตอบไม่ถูกและยากในการที่จะเดาคำตอบได้ถูกต้อง จึงทำให้ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบโคลซแต่ละฉบับกับคะแนนที่ได้จากแบบทดสอบวัดความเข้าใจในการอ่านแบบเลือกตอบมีความสัมพันธ์กันค่อนข้างสูงซึ่งสอดคล้องกับผลการวิจัยทรงศักดิ์ ภูสีอ่อน (2542 : 112) ได้เปรียบเทียบคะแนนของแบบทดสอบโคลซปรากฏว่า ความสัมพันธ์ระหว่างคะแนนของแบบทดสอบโคลซแต่ละฉบับกับคะแนนที่ได้จากแบบทดสอบวัดความเข้าใจในการอ่านแบบเลือกตอบมีความสัมพันธ์กันค่อนข้างสูง จึงสรุปได้ว่าลักษณะของแบบทดสอบต่างกันและเกณฑ์การให้คะแนนที่ต่างกันไม่มีผลต่อค่าความเที่ยงตรงของแบบทดสอบ

จากการศึกษาเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 จากการทดสอบความแตกต่างพบว่ามีความแตกต่างกัน อย่างน้อย 1 คู่เมื่อทดสอบความแตกต่างเป็นรายคู่พบว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 4 คู่ และแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 1 คู่ และค่าความเชื่อมั่นของแบบทดสอบโคลซทั้ง 9 ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 จากการทดสอบความแตกต่างพบว่าอย่างน้อย 1 คู่ที่มีความแตกต่าง เมื่อทดสอบรายคู่พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 8 คู่ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จำนวน 2 คู่ ส่วนคู่อื่นๆ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ค่าความเชื่อมั่นของแบบทดสอบโคลซแต่ละฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์การตรวจที่ต่างกัน จากการทดสอบความแตกต่างพบว่าแตกต่างกันอย่างไม่มีนัยสำคัญเมื่อพิจารณาโดยรวมแล้ว แบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 และเกณฑ์ที่ 2 มีค่าความเชื่อมั่นอยู่ในเกณฑ์สูงมากและมีค่าใกล้เคียงกันด้วย ทั้งนี้อาจเป็นผลเนื่องจาก แบบทดสอบโคลซนั้น ผู้สอบจะต้องใช้ความสามารถของตนเองโดยอาศัยการระลึกได้ และใช้

ข้อความที่เหลืออยู่มาประมวลความรู้ด้านต่าง ๆ โดยอาศัยประสบการณ์และความสามารถเพื่อหาคำมาเติมในช่องว่าง ผู้สอบต้องมีความเข้าใจเนื้อเรื่องที่อ่านอย่างแท้จริงจึงจะสามารถทำแบบทดสอบโคลซได้ จากคำกล่าวของกรลอนสันด์ (Gronlund. 1976 : 119) และแมกนุสสัน (สุพัฒน์ สุกมลสันต์. 2530 : 15 ; อ้างอิงมาจาก Magusson. 1966 : 66) ที่ว่าปัจจัยมีผลต่อค่าความเชื่อมั่นของแบบทดสอบคือการกระจายของคะแนนนั้นคือ ค่าความเชื่อมั่นย่อมขึ้นอยู่กับความแปรปรวนระหว่างคะแนนผลการสอบ ถ้าความแปรปรวนระหว่างผลการสอบมีน้อย ย่อมได้ค่าความเชื่อมั่นต่ำว่าคะแนนผลการสอบที่มีความแปรปรวนมาก ซึ่งในการวิจัยครั้งนี้ใช้แบบทดสอบโคลซซึ่งในแต่ละช่องว่างของแบบทดสอบโคลซแต่ละฉบับที่สร้างจากเรื่องเดียวกัน มีทั้งช่องที่อาจเติมคำที่ยากและง่ายได้จึงทำให้การกระจายของคะแนนสอบมีมาก และมีผลทำให้แบบทดสอบโคลซแต่ละฉบับมีค่าความเชื่อมั่นในระดับที่สูงมากและใกล้เคียงกันอีกด้วยส่วนค่าความเชื่อมั่นของแบบทดสอบโคลซแต่ละฉบับที่มีเกณฑ์การตรวจให้คะแนนที่ต่างกันคือ เกณฑ์ที่ 1 และเกณฑ์ที่ 2 นั้น แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติซึ่งมีแนวโน้มว่า แบบทดสอบโคลซที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1

จากการเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซทั้ง 9 ฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 จากการทดสอบความแตกต่างของความแปรปรวนพบว่ามีอย่างน้อย 1 คู่ที่แตกต่างกันเมื่อทดสอบเป็นรายคู่พบว่ามีความแตกต่างอย่างมีนัยสำคัญที่ระดับ .01 จำนวน 8 คู่ และแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 3 คู่ ส่วนคู่อื่นๆ มีความแปรปรวนแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 จากการทดสอบความแตกต่างของความแปรปรวนพบว่ามีอย่างน้อย 1 คู่ที่แตกต่างกันเมื่อทดสอบเป็นรายคู่พบว่าคู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 มีจำนวน 3 คู่ คู่ที่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 มีจำนวน 5 คู่ ส่วนคู่อื่นๆ แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติส่วนค่าความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบโคลซแต่ละฉบับที่ตรวจให้คะแนนโดยใช้เกณฑ์การตรวจที่แตกต่างกัน(ฉบับเดียวกันแต่ตรวจให้คะแนนโดยใช้เกณฑ์การตรวจที่ต่างกัน คือ เกณฑ์ที่ 1 และเกณฑ์ที่ 2 จากการทดสอบความแตกต่างพบว่าความคลาดเคลื่อนของแบบทดสอบโคลซทั้ง 9 ฉบับ ที่ตรวจให้คะแนนโดยเกณฑ์ที่ 1 และเกณฑ์ที่ 2 พบว่า ทุกฉบับแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ โดยแบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดอัตราส่วนตายตัวและกำหนดคำแรกของคำตอบ มีค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงที่สุด ส่วนแบบทดสอบโคลซที่ตัดคำแบบไม่กำหนดอัตราส่วนตายตัวและกำหนดจำนวนตัวอักษรของคำตอบ มีค่าความคลาดเคลื่อนมาตรฐานในการวัดต่ำที่สุด ถึงแม้ว่าแบบทดสอบโคลซทั้ง 9 ฉบับจะมีค่าความเชื่อมั่นใกล้เคียงกันแต่ลักษณะการตัดคำที่ต่างกันและการที่มีตัวชี้แนะในการตอบ ของแบบทดสอบโคลซอาจส่งผลต่อความคลาดเคลื่อนมาตรฐานในการวัดของแบบทดสอบได้

จากผลการทดสอบความเป็นเอกพันธ์ของความแปรปรวนของคะแนนความคลาดเคลื่อนของแบบทดสอบโคลซแต่ละฉบับ ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ต่างกัน ที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 และเกณฑ์ที่ 2 แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งไม่สอดคล้องกับสมมติฐาน หากมองโดยภาพรวมแล้วแบบทดสอบโคลซที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 2 มีค่าความคลาดเคลื่อนมาตรฐานในการวัดสูงกว่าแบบทดสอบโคลซที่ตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 ทั้งนี้อาจเป็นผลเนื่องมาจาก วิธีการตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1 คือ การตรวจที่เชื่อถือได้ มีคำตอบที่แน่นอนเพราะมีเพียงคำตอบเดียวที่เป็นคำตอบ จึงเป็นปรนัยสามารถตรวจให้คะแนนได้ตรงกันส่วนการตรวจได้คะแนนโดยใช้เกณฑ์ที่ 2 คือการตรวจให้คะแนนที่มีความเหมาะสมหรือใกล้เคียงด้วย แม้จะไม่ตรงกับคำที่ผู้ออกข้อสอบระบุไว้ นั้น ซึ่งอาจจะไม่สามารถตรวจให้คะแนนได้ตรงกัน หรือบางที่ผู้ตรวจก็ลังเลว่าสมควรจะให้คะแนนกับคำตอบคำนั้นหรือไม่ซึ่งจะส่งผลให้เกิดความคลาดเคลื่อนในการวัดได้มากกว่า การตรวจให้คะแนนโดยใช้เกณฑ์ที่ 1

ข้อเสนอแนะ

1. ด้านการนำไปใช้

1.1 ในการเลือกใช้เกณฑ์การให้คะแนนของแบบทดสอบโคลซนั้น ถ้าเป็นการสอบคัดเลือกหรือสอบแข่งขันเพื่อจัดอันดับ ควรใช้เกณฑ์ควรใช้เกณฑ์การตรวจให้คะแนนเกณฑ์ที่ 1 ซึ่งจะมีค่าอำนาจจำแนกไม่แตกต่างกันกับเกณฑ์ที่ 2 และมีความสะดวกในการให้คะแนนมากกว่าอีกด้วย แต่ถ้าเป็นการสอบวัดจุดประสงค์ หรือสอบกลางภาค ปลายภาค ควรที่จะใช้เกณฑ์การตรวจให้คะแนนที่ 2 เพื่อเป็นการตรวจสอบว่านักเรียนผ่านจุดประสงค์นั้นแล้วหรือยังและมีข้อบกพร่องตรงไหนจะเหมาะสมมากกว่า

1.2 การนำแบบทดสอบโคลซไปทดสอบกับผู้เรียน ควรตั้งเกณฑ์ให้เหมาะสม คือ ควรตั้งเกณฑ์ให้ต่ำกว่าการทดสอบโดยใช้แบบทดสอบแบบเลือกตอบ

2. ด้านการวิจัย

2.1 ควรศึกษาเจตคติและแรงจูงใจของครูและนักเรียนที่มีต่อแบบทดสอบโคลซ เพื่อเป็นแนวทางในการปรับปรุงให้มีประสิทธิภาพมากยิ่งขึ้น

2.2 ควรทำการศึกษาเกี่ยวกับ การสร้างข้อความ โดยนำคำที่เราต้องการสอนมาสร้างเป็นเรื่องราวและตัดคำเหล่านั้นออก แล้วจึงให้นักเรียนเติมคำเหล่านั้น ซึ่งเป็นลักษณะของแบบทดสอบโคลซ เช่น การสอนคำใหม่ในบทเรียน เป็นต้นเพื่อตรวจสอบว่า จะสามารถนำมาประยุกต์ใช้ในการเรียนการสอนให้เด็กเข้าใจและมีประสิทธิภาพมากน้อยเพียงใด

บรรณานุกรม

กระทรวงศึกษาธิการ. **ตัวชี้วัดและสาระการเรียนรู้แกนกลาง กลุ่มสาระการเรียนรู้ภาษาไทย**

ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐานพุทธศักราช 2551. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว, 2551.

กัลยาณี นันทะศรี. การเปรียบเทียบคุณภาพของแบบทดสอบโมดิฟายด์โคลซ 5 รูปแบบในการวัดความเข้าใจในการอ่านภาษาอังกฤษของนักเรียนชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ กศ.ม. มหาสารคาม : มหาวิทยาลัยมหาสารคาม, 2542.

ทรงศักดิ์ ภูศรีอ่อน. การเปรียบเทียบคุณภาพของแบบทดสอบโคลซ ในการวัดความเข้าใจในการอ่านภาษาไทย โดยใช้วิธีการตัดคำและเกณฑ์การให้คะแนนที่แตกต่างกัน. วิทยานิพนธ์ กศ.ม. มหาสารคาม : มหาวิทยาลัยมหาสารคาม, 2542.

บุญชม ศรีสะอาด. การวิจัยทางการวัดผลและประเมินผล. กรุงเทพฯ : สุวีริยาสาส์น, 2540.

สมนึก ภัททิยธนี. การวัดผลการศึกษา. พิมพ์ครั้งที่ 6. กทม. : ประสานการพิมพ์, 2551.

สุพัฒน์ สุกมลสันต์. แบบทดสอบโคลซและการประยุกต์ใช้. วิทยการวิจัย. 2(2) : พฤษภาคม - สิงหาคม 2530

สุพรรณ พันธุ์สุภา. ความสัมพันธ์ระหว่างทักษะทางการเขียนภาษาไทยกับประสิทธิภาพของการสื่อสารที่วัดได้จากแบบทดสอบโคลซของนักเรียนชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ กศ.ม. มหาสารคาม : มหาวิทยาลัยศรีนครินทรวิโรฒ, 2539

อรนุช วรอำศวปติ. การศึกษาคุณภาพแบบทดสอบโคลซ ชนิดเว้นช่องว่างให้เติมคำชนิดกำหนดตัวอักษรตัวแรกของคำ ชนิดกำหนดจำนวนตัวอักษร และชนิดบอกความหมายของคำ ในการวัดสมรรถภาพทางการอ่านภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 4. วิทยานิพนธ์ ศษ.ม. เชียงใหม่ : มหาวิทยาลัยเชียงใหม่, 2530.