

การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง
ที่มีรูปแบบคำถามแตกต่างกัน

ปริญญานิพนธ์
ของ
สุรี ใจกล้า

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2544
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

371. ๒๖๐13

๙ ๘๖๖ ๐

๙. 3

การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง
ที่มีรูปแบบคำถามแตกต่างกัน

บทคัดย่อ

ของ

สุริ ใจกล้า

27 ก.ย. 2544

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2544

155704 11/2/17

สุรี ใจกล้า. (2544). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง ที่มีรูปแบบคำถามแตกต่างกัน. ปริญญาโท กศ.ม. (การวัดผลการศึกษา).

กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ คณะกรรมการควบคุม : รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนันตพงษ์, อาจารย์ชวลิต รวยอาจิณ.

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อศึกษาความเชื่อมั่น และเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง ที่มีรูปแบบคำถามแตกต่างกันคือ แบบเลือกตอบ, แบบตัวเลือกคงที่ และแบบจับคู่

กลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2543 โรงเรียนสังกัดกรมสามัญศึกษา เขตกรุงเทพมหานคร จำนวน 1,200 คน ได้จากการสุ่มแบบ 2 ขั้นตอน เครื่องมือที่ใช้ในการศึกษาเป็นแบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 ฉบับ ประกอบด้วย ฉบับที่ 1 เป็นแบบทดสอบวัดความรู้ทางเครื่องมือช่างแบบเลือกตอบ ฉบับที่ 2 เป็นแบบทดสอบวัดความรู้ทางเครื่องมือช่างแบบตัวเลือกคงที่ ฉบับที่ 3 เป็นแบบทดสอบวัดความรู้ทางเครื่องมือช่างแบบจับคู่

ผลการวิจัยสรุปได้ดังนี้

แบบทดสอบวัดความรู้ทางเครื่องมือช่างที่มีรูปแบบคำถามแบบเลือกตอบมีค่าความเชื่อมั่น เท่ากับ 0.7533 แบบตัวเลือกคงที่ เท่ากับ 0.6994 และแบบจับคู่ เท่ากับ 0.7021 เมื่อทดสอบค่าความแตกต่างของค่าความเชื่อมั่นด้วยสูตร UX_1 พบว่า ค่าความเชื่อมั่นของแบบทดสอบทั้ง 3 รูปแบบแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

**THE COMPARISON OF RELIABILITY OF TOOL KNOWLEDGE TEST
WITH DIFFERENT-QUESTION FORMATS.**

**AN ABSTRACT
BY
SUREE JAIGLA**

**Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University
May 2001**

Suree Jaigla.(2001). The Comparison of Reliability of Tool Knowledge Test with Different-Question Formats. Master Thesis, M.Ed. (Educational Measurement). Bangkok, Graduated School, Srinakharinwirot University. Advisor Committee: Assoc. Prof. Dr. Boonchird Pinyoanuntapong, Mr. Chawalit Ruayajin.

The purpose of the research was to study and compare the reliability of Tool Knowledge Test with different-question formats by using Multiple Choice, Fixed Choice and Matching Achievement Tests.

The subject consisted of 1,200 Mathayomsuksa III students of the second semester in the academic year 2000 under the Department of General Education in Bangkok. The subject was selected by the two-stage random sampling. The instrument of the study was the Tool Knowledge Test with three formats as Multiple Choices, Fixed Choices and Matching.

The result was as follows:

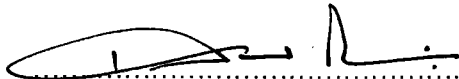
The reliability of Tool Knowledge Test by using Multiple-Choice Form was 0.7533, by using Fixed-Choice Form was 0.6994 and by using Matching Form was 0.7021. When compared with UX_1 , it was found that there was no significant difference among three formats.

ปริญญานิพนธ์
เรื่อง

การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างที่มีรูปแบบ
การตอบแตกต่างกัน

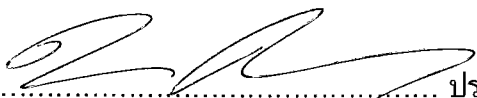
ของ
นางสาวสุรี ใจกล้า

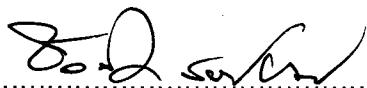
ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

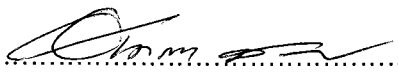

..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร.เสริมศักดิ์ วิศาลาภรณ์)

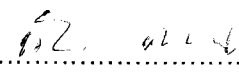
วันที่.....9.....เดือน พฤษภาคม พ.ศ. 2544

คณะกรรมการสอบปริญญานิพนธ์


..... ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญนนตพงษ์)


..... กรรมการ
(อาจารย์ ชวลิต รวยอาจัน)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ อังคณา สายยศ)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ นิภา ศรีไพโรจน์)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ได้รับความกรุณาจาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์ และ อาจารย์ชวลิต รวยอาจิณ ที่กรุณาให้คำปรึกษา ชี้แนะและแก้ไข ข้อบกพร่องต่างๆ จนกระทั่งปริญญานิพนธ์สำเร็จลงด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

กราบขอบพระคุณ รองศาสตราจารย์อังคณา สายยศ และรองศาสตราจารย์นิภา ศรีไพโรจน์, ที่ร่วมเป็นกรรมการสอบปริญญานิพนธ์ และได้ให้ข้อเสนอแนะที่เป็นประโยชน์ เพื่อให้ปริญญานิพนธ์มีความสมบูรณ์มากยิ่งขึ้น

กราบขอบพระคุณผู้ทรงคุณวุฒิทั้ง 5 ท่าน คือ รองศาสตราจารย์ นิภา ศรีไพโรจน์ อาจารย์ชวลิต รวยอาจิณ อาจารย์โอฬาร กาลวิบูลย์ อาจารย์ชัชชัย มณีโชติ และอาจารย์บุญร่วม เหล็กสูงเนิน ที่ให้คำแนะนำ แก้ไขข้อบกพร่องต่างๆ ในการสร้างเครื่องมือในครั้งนี้

กราบขอบพระคุณอธิการบดี อธิการ และคณบดี ของ มหาวิทยาลัย และ โรงเรียนต่างๆ ที่เป็นกลุ่มตัวอย่างในการศึกษา ที่ให้อนุเคราะห์ในการเก็บรวบรวมข้อมูล และ ขอขอบคุณนักเรียนที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

กราบขอบพระคุณบิดา มารดา ที่ได้อบรมสั่งสอน ให้ความเมตตาอุปการะ สนับสนุนการศึกษาของผู้วิจัยตลอดมา และกราบขอบพระคุณ คุณสมปอง ใจกล้า และผู้ร่วมงานทุกท่าน ที่ให้การช่วยเหลือ สนับสนุนเป็นกำลังใจ รวมทั้งเพื่อนทุกคนที่เป็นกำลังใจ ความดีและประโยชน์อันพึงมีจากปริญญานิพนธ์ฉบับนี้ขอมอบแด่ บิดา มารดา บุรพคณาจารย์ และผู้มีพระคุณทุกท่าน

สุรี ใจกล้า

สารบัญ

บทที่	หน้า
1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการศึกษาค้นคว้า	3
ความสำคัญของการศึกษาค้นคว้า	3
ขอบเขตของการศึกษาค้นคว้า	3
นิยามศัพท์เฉพาะ	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	
หลักสูตรมัธยมศึกษาตอนต้น	5
การวัดผล	8
การวัดด้านสติปัญญา	10
การทดสอบการปฏิบัติงาน	12
การวิเคราะห์ข้อสอบรายข้อ	31
ความเชื่อมั่นของแบบทดสอบ	32
งานวิจัยที่เกี่ยวข้อง	50
3 วิธีดำเนินการศึกษาค้นคว้า	
ประชากร	62
กลุ่มตัวอย่าง	62
วิธีดำเนินการสร้างเครื่องมือ	64
เครื่องมือที่ใช้ในการศึกษา	66
วิธีดำเนินการเก็บรวบรวมข้อมูล	69
สถิติในการวิเคราะห์ข้อมูล	69
4 ผลการวิเคราะห์ข้อมูล	
สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อมูล	72
ผลการวิเคราะห์ข้อมูล	72

สารบัญ (ต่อ)

บทที่	หน้า
5 สรุป อภิปรายผล และ ข้อเสนอแนะ	
ความมุ่งหมายในการศึกษาค้นคว้า.....	75
สรุปผลการศึกษาค้นคว้า.....	76
อภิปรายผล.....	77
ข้อเสนอแนะ.....	77
 บรรณานุกรม.....	 80
 ภาคผนวก.....	 84
 ภาคผนวก ก.....	 85
รายชื่อผู้เชี่ยวชาญประเมินความสอดคล้องของข้อคำถาม กับนิยามศัพท์เฉพาะ.....	 86
 ภาคผนวก ข.....	 87
ค่าความยาก(p) อำนาจจำแนก(r) ของแบบทดสอบ วัดความรู้ทางเครื่องมือช่างแบบเลือกตอบจำนวน 50 ข้อ.....	 88
 ภาคผนวก ค.....	 90
แบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 รูปแบบ	91
 ประวัติย่อผู้วิจัย.....	 110

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 ระดับความเป็นจริงของการวัดภาคปฏิบัติ.....	15
2 แนวทางการวัดผลภาคปฏิบัติ.....	18
3 ประเภทต่างๆ ของแบบทดสอบวัดผลสัมฤทธิ์	26

บัญชีตาราง

ตาราง	หน้า
1 แสดงการแปลความหมาย.....	31
2 แสดงค่าอำนาจจำแนกจะพิจารณาระดับอำนาจจำแนก.....	32
3 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย.....	63
4 คะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบ วัดความรู้ทางเครื่องมือช่าง.....	73
5 ค่าความแปรปรวนของคะแนนในแต่ละรูปแบบ ผลรวมของ ค่าความแปรปรวนของคะแนนในแต่ละส่วน ผลรวมของ ค่าความแปรปรวนของความแปรปรวนรวม ค่าความเชื่อมั่น... 74	
6. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ ความรู้ทางเครื่องมือช่าง.....	74

บทที่ 1

บทนำ

ภูมิหลัง

การวัดผลการศึกษาเป็นส่วนสำคัญประการหนึ่งของกระบวนการเรียนการสอนในระดับการศึกษาต่างๆ เพราะเป็นเครื่องมือที่ช่วยพัฒนาคุณภาพการศึกษาโดยมีผลจากการวัดเป็นพื้นฐานในการตัดสินใจของครู และนักการศึกษาเพื่อประโยชน์ในการปรับปรุงวิธีการสอน การแนะแนว การประเมินหลักสูตร แบบเรียน การใช้อุปกรณ์การสอน ตลอดจนการจัดระบบบริหารทั่วไปของโรงเรียน และยังช่วยปรับปรุงการเรียนให้ถูกวิธียิ่งขึ้น (อนันต์ ศรีโสภณ, 2520 : 1) ซึ่งตรงกับปรัชญาการวัดผลที่ว่าทดสอบ เพื่อค้นและพัฒนาสมรรถภาพมนุษย์ วิธีวัดผลที่สำคัญ และใช้มากที่สุดในการศึกษาจนอาจนับได้ว่าเป็นงานหลักของการวัดผลสัมฤทธิ์ในโรงเรียน คือ การทดสอบ การทดสอบที่ดีจะช่วยให้ครูทราบสถานภาพของนักเรียนและของครูเองได้ง่ายกว่าวิธีอื่นๆ (ชวาล แพรัตกุล, 2516 : 88) การสอบมีคุณค่ามากน้อยเพียงใดนั้นย่อมขึ้นอยู่กับคุณภาพของแบบทดสอบที่ใช้ดังที่ ธอร์นไคด์ (Thorndike, 1971:27) ได้ให้ความเห็นว่แบบทดสอบที่สร้างอย่างดี และใช้ได้ผลนั้นต้องเป็นแบบทดสอบที่สามารถใช้เป็นแรงจูงใจในการสร้างนิสัยในการเรียนที่ดี ทำให้ผู้เรียนสามารถแก้ข้อผิดพลาดในการเรียนให้ดีขึ้น และเป็นแนวทางที่จะนำไปสู่ผลสัมฤทธิ์ในการเรียนรู้ได้

เครื่องมือที่ใช้ในการวัดผลการศึกษาในปัจจุบันจะมีคุณค่าหรือสามารถวัดได้ตรงกับความเป็นจริงเพียงใด ย่อมขึ้นอยู่กับความรู้จักเลือกเครื่องมือ การเลือกใช้รูปแบบของคำถามให้เหมาะสมกับเนื้อหาวิชานั้นว่าเป็นเรื่องสำคัญ สำหรับวิชางานช่างพื้นฐานเป็นวิชาที่ต้องใช้ความรู้ความเข้าใจ การจำแนกเครื่องมือ หรือชิ้นส่วนของเครื่องมือ และวิธีการใช้งานของเครื่องมือเพื่อเป็นพื้นฐานก่อนที่จะปฏิบัติต่อไป ดังนั้นครูจึงควรใช้แบบทดสอบเพื่อที่สามารถวัดความรู้ ความสามารถของผู้เรียน โดยเน้นในเรื่องความรู้ความเข้าใจ การจำแนก และวิธีการใช้งานเกี่ยวกับเครื่องมือเพราะเป็นหลักสำคัญในการเรียนวิชางานช่างพื้นฐาน ปัจจุบันนักวัดผลได้ให้ความสนใจเกี่ยวกับการพัฒนาคุณภาพแบบทดสอบเพื่อให้ได้แบบทดสอบที่มีประสิทธิภาพ เชื่อถือได้และวัดได้ตรงตามจุดประสงค์

แบบทดสอบแบบเลือกตอบ (ล้วน สายยศ และอังคณา สายยศ, 2539) เป็นข้อสอบที่นิยมใช้มากในปัจจุบันแบบทดสอบมาตรฐานใช้แบบเลือกตอบทั้งสิ้นทั้งนี้ก็เพราะข้อสอบแบบเลือกตอบสามารถวัดได้ครอบคลุมจุดประสงค์และตรวจให้คะแนนได้แน่นอน ข้อสอบแบบเลือกตอบใช้แทนข้อสอบรูปแบบอื่นๆ ได้ดีแม้แต่ข้อสอบความเรียงข้อสอบแบบเลือกตอบก็

สามารถใช้แทนได้ จากผลการวิจัยของคุณ (ล้วน สายยศ และอังคณา สายยศ. 2539; อ้างอิงจาก Cook; อ้างอิงจาก Ebel. 1979:137) ปรากฏผลว่าข้อสอบทั้งสองแบบที่วัดผลสัมฤทธิ์สิ่งเดียวกันมีสัมประสิทธิ์สหสัมพันธ์ .97 นั่นคือมีความสัมพันธ์กันสูงมากอาจใช้แทนกันได้ดีบางจุดประสงค์ นอกจากนี้ ล้วนและอังคณา สายยศ(ล้วน สายยศ และอังคณา สายยศ. 2539)ได้กล่าวว่า แบบทดสอบความเรียง แบบเติมคำ แบบจับคู่ และแบบถูกผิดวัดได้ดีตามจุดประสงค์เพียงบางประการ และมีข้อเสียของมันเอง แต่ละชนิดยากที่จะวัดได้ครอบคลุมจุดประสงค์ที่เราต้องการทั้งหมด และเมเยอร์(Mayer.1955:128) ได้ศึกษาการใช้แบบทดสอบแบบถูกผิดแบบเลือกตอบ และแบบเติมคำ พบว่าแบบทดสอบเลือกตอบและถูกผิดมีความเชื่อมั่นมากกว่าแบบทดสอบอัตนัยเติมคำ

แบบทดสอบแบบตัวเลือกคงที่ (ล้วน สายยศ และอังคณา สายยศ.2539) เป็นประเภทหนึ่งของแบบทดสอบเลือกตอบ เกิดขึ้นจากตัวเลือกจากตัวเลือกแต่ละข้อใดในคำถามโดดๆ ซ้ำกันอยู่บ่อยๆ ดังนั้นเพื่อให้คำถามและตัวเลือกมีประสิทธิภาพขึ้นจึงนำตัวเลือกที่ซ้ำมาเป็นตัวเลือกคงที่ และเพิ่มคำถามเป็นข้อๆ เท่านั้น

แบบทดสอบแบบจับคู่(ล้วน สายยศ และอังคณา สายยศ.2539)แบบทดสอบชนิดนี้เหมาะสำหรับการวัดความรู้เกี่ยวกับคำจำกัดความ เหตุการณ์ และตัวอย่างของความคิดรวบยอดที่มีความเกี่ยวข้องกัน หรือสัมพันธ์กัน ซึ่งจะเหมาะที่จะวัดเรื่องเกี่ยวกับตัวบุคคล (ใคร) การกระทำของบุคคล (ทำอะไร) เหตุการณ์สำคัญๆ (เมื่อไร) และสถานที่ เช่นจับคู่สถานที่กับเหตุการณ์, จับคู่ชื่อผู้เขียนหนังสือ, จับคู่ชื่อเครื่องมือและประโยชน์ของเครื่องมือ ในแต่ละกรณีอาจจับคู่บุคคลกับผลงานหรือเวลา จากการวิจัยของโดนัลด์ และคนอื่นๆ (Donald and ot her N.d. 179 - 182) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบแบบเติมคำ, แบบจับคู่, แบบเลือกตอบ โดยนำไปทดสอบกับนักศึกษาปริญญาตรีที่เรียนวิชาจิตวิทยาเรื่องสั้น พบว่าแบบทดสอบเติมคำมีค่าความเชื่อมั่นสูงกว่าแบบทดสอบจับคู่ และแบบทดสอบจับคู่มีค่าความเชื่อมั่นสูงกว่า แบบทดสอบเลือกตอบ

จากหลักฐานและผลการศึกษางานวิจัยดังกล่าวมาแล้วนั้น จะเห็นว่าคุณภาพของแบบทดสอบขึ้นอยู่กับรูปแบบของแบบทดสอบเป็นสำคัญ ด้วยเหตุนี้ทำให้ผู้วิจัยสนใจที่จะศึกษารูปแบบคำถาม ของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง ว่าแบบทดสอบเลือกตอบแบบตัวเลือกคงที่ แบบจับคู่ มีผลต่อค่าความเชื่อมั่นมากน้อยเพียงใด ซึ่งผลจากการวิจัยจะทำให้เกิดประโยชน์สำหรับครูผู้สอนและผู้ที่เกี่ยวข้องกับการวัดประเมินผล ในการเลือกใช้แบบทดสอบให้เหมาะสม ซึ่งจะเป็นแนวทางในการสร้างเครื่องมือวัดความรู้ทางเครื่องมือช่าง ที่มีคุณภาพยิ่งขึ้น

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายสำคัญเพื่อศึกษารูปแบบคำถามของแบบทดสอบวัดความรู้ทางเครื่องมือช่างมีผลต่อค่าความเชื่อมั่นมากน้อยเพียงใด โดยกำหนดเป็นจุดมุ่งหมายเฉพาะของการวิจัยครั้งนี้

1. เพื่อหาค่าความเชื่อมั่น ของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง แบบเลือกตอบ, แบบตัวเลือกคงที่ และแบบจับคู่
2. เพื่อเปรียบเทียบความเชื่อมั่น ของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง แบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่

ความสำคัญของการศึกษาค้นคว้า

ผลการศึกษาค้นคว้าในครั้งนี้ ทำให้ได้เครื่องมือวัดความรู้ทางเครื่องมือช่างที่เหมาะสมสำหรับนักเรียนมัธยมศึกษาปีที่ 3 ที่เลือกเรียนวิชางานช่างพื้นฐาน และทำให้ทราบว่ามีรูปแบบคำถามของแบบทดสอบวัดความรู้ทางเครื่องมือช่างมีผลต่อค่าความเชื่อมั่นหรือไม่อีกด้วยอันจะเป็นแนวทางในการสร้างเครื่องมือวัดความรู้ทางเครื่องมือช่างที่มีคุณภาพยิ่งขึ้น

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่กำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2543 ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 117 โรงเรียน 1,290 ห้องเรียน จำนวนนักเรียนทั้งหมด 53,774 คน
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่เลือกเรียนวิชางานช่างพื้นฐาน ที่กำลังศึกษาภาคเรียนที่ 2 ปีการศึกษา 2543 ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวนนักเรียน 1,200 คน โดยสุ่มตัวอย่างแบบ 2 ขั้นตอน
3. แบบทดสอบวัดความรู้ทางด้านช่าง เป็นการทดสอบความรู้เกี่ยวกับเครื่องมือช่างในระดับที่ผู้วิจัยสร้างขึ้น โดยสร้างขึ้นเป็นสามรูปแบบคือ แบบเลือกตอบ , แบบตัวเลือกคงที่ และแบบจับคู่
4. ตัวแปรที่ใช้ในการศึกษาค้นคว้า
 - ก. ตัวแปรอิสระ ได้แก่ รูปแบบของคำถาม จำแนกดังนี้
 1. แบบเลือกตอบ
 2. แบบตัวเลือกคงที่
 3. แบบจับคู่

3. แบบจับคู่

ข. ตัวแปรตาม ได้แก่ ความเชื่อมั่นของแบบทดสอบ

นิยามศัพท์เฉพาะ

1. ความเชื่อมั่น (Reliability) ของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง หมายถึง สัดส่วนของความแปรปรวนของคะแนนจริงต่อความแปรปรวนของคะแนนที่สอบได้ สำหรับงานวิจัยในครั้งนี้หาค่าความเชื่อมั่นของแบบทดสอบแบบคะแนนจริงสัมพันธ์ โดยใช้สูตรของ เฟลด์ต์ – ราซู (Feldt-Raju)

2. แบบทดสอบวัดความรู้ทางเครื่องมือช่าง หมายถึง ชุดของคำถามที่วัดความจำการรู้จักเครื่องมือต่างๆ ที่ใช้ในงานช่างไม้ ช่างปูน ช่างก่อสร้าง ช่างไฟฟ้า เป็นต้น ที่ใช้ในการทำงานแต่ละอย่างโดยดูว่าเครื่องมือหรือลักษณะของการนำไปใช้งานมีส่วนสัมพันธ์หรือเกี่ยวข้องกัน

3. รูปแบบของข้อสอบวัดความรู้ทางเครื่องมือช่าง ที่ผู้วิจัยสร้างขึ้น แบ่งออกเป็น 3 รูปแบบ คือ

3.1 แบบเลือกตอบ หมายถึง ชุดของคำถามที่ประกอบด้วย 2 ส่วนคือส่วนที่เป็นรูปภาพที่กำหนดให้เป็นคำถามและส่วนรูปภาพที่เป็นคำตอบให้เลือก จำนวน 5 ตัวเลือกโดยกำหนดให้รูปภาพทั้ง 2 ส่วนมีความสัมพันธ์หรือเกี่ยวข้องกัน

3.2 แบบตัวเลือกคงที่ หมายถึง ชุดของคำถามที่ประกอบด้วย 2 ส่วนคือส่วนที่เป็นชุดของข้อความที่เป็นคำถามและส่วนที่เป็นชุดของรูปภาพเป็นตัวเลือก โดยกำหนดให้ชุดของตัวเลือกมีความสัมพันธ์หรือเกี่ยวข้องกัน

3.3 แบบจับคู่ หมายถึง ชุดของคำถามที่ประกอบด้วยชุดของรูปภาพ 2 ชุด คือชุดหนึ่งเป็นชุดที่เป็นคำถาม อีกชุดหนึ่งเป็นตัวเลือก ให้พิจารณาจากชุดที่เป็นคำถามว่ามีความสัมพันธ์หรือเกี่ยวข้องกันชุดที่เป็นตัวเลือกในข้อใด

4. ผู้เชี่ยวชาญ หมายถึง อาจารย์ผู้สอนวิชาการวัดผลการศึกษาหรือผู้ที่มีประสบการณ์และความรู้ทางช่าง

สมมุติฐานการวิจัย

ความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง แบบเลือกตอบแบบตัวเลือกคงที่ และแบบจับคู่ มีค่าแตกต่างกัน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาครั้งนี้ ผู้วิจัยได้ทำการศึกษาเนื้อหาต่าง ๆ ตามลำดับต่อไปนี้

1. หลักสูตรมัธยมศึกษาตอนต้น
2. การวัดผล
 - 2.1 การวัดด้านสติปัญญา
 - 2.2 การวัดภาคปฏิบัติ
3. ชนิดของแบบทดสอบ
4. ความเชื่อมั่นของแบบทดสอบ
5. งานวิจัยที่เกี่ยวข้อง

1. หลักสูตรมัธยมศึกษาตอนต้น

ผู้ที่เรียนจบหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) จะต้องเรียนหมวดวิชาต่าง ๆ ดังนี้ (กรมวิชาการ กระทรวงศึกษาธิการ 2534 : 2, 92-93)

1. วิชาบังคับ จำนวน 57 หน่วยการเรียนรู้ ได้แก่รายวิชาดังต่อไปนี้

1.1 วิชาบังคับแกน จำนวน 39 หน่วยการเรียนรู้

ภาษาไทย	12 หน่วยการเรียนรู้
วิทยาศาสตร์	9 หน่วยการเรียนรู้
คณิตศาสตร์	6 หน่วยการเรียนรู้
สังคมศึกษา	6 หน่วยการเรียนรู้
พลานามัย	3 หน่วยการเรียนรู้
ศิลปศึกษา	3 หน่วยการเรียนรู้

1.2 วิชาบังคับเลือก จำนวน 18 หน่วยการเรียนรู้

สังคมศึกษา	6 หน่วยการเรียนรู้
พลานามัย	6 หน่วยการเรียนรู้
การทำงาน	6 หน่วยการเรียนรู้

2. วิชาเลือกเสรี จำนวน 33 หน่วยการเรียนรู้ ให้เลือกจากรายวิชาในกลุ่มวิชาต่าง ๆ ต่อไปนี้

2.1 กลุ่มวิชาภาษา

ภาษาไทย

ภาษาต่างประเทศ

2.2 กลุ่มวิชาวิทยาศาสตร์-คณิตศาสตร์

วิทยาศาสตร์

คณิตศาสตร์

2.3 กลุ่มวิชาสังคมศึกษา

2.4 กลุ่มวิชาพัฒนาบุคลิกภาพ

พลานามัย

ศิลปศึกษา

2.5 กลุ่มวิชาการงานและอาชีพ

อาชีพ

วิชาบังคับเลือก ให้เลือกเรียนจากรายวิชาต่อไปนี้ ภาคเรียนละ 1 รายวิชา โดยจะเลือกเรียนรายวิชาใดก่อนหลัง ก็ได้

ง 011 งานบ้าน

ง 012 การจัดการในบ้าน

ง 013 งานช่างพื้นฐาน

ง 014 งานเกษตรพื้นฐาน

ง 015 งานผลิตภัณฑ์จากวัสดุท้องถิ่น

งานช่างพื้นฐาน คือ การศึกษางานช่างพื้นฐานที่จำเป็นในบ้าน วิเคราะห์ วางแผนและลงมือปฏิบัติในเรื่องการใช้ การเก็บ การบำรุงรักษาเครื่องมือ และอุปกรณ์งานช่าง การอ่านแบบเบื้องต้น การซ่อมแซมและดัดแปลงเครื่องใช้ในบ้าน และการรักษาความปลอดภัยในการทำงาน เพื่อให้มีความรู้ ความเข้าใจในงานช่างพื้นฐาน สามารถซ่อมแซมและดัดแปลงเครื่องใช้ในบ้านได้อย่างถูกต้องตามกระบวนการทำงาน และปลอดภัย มีนิสัยขยัน อดทน ประณีต รอบคอบ มีระเบียบและประหยัด

งานช่างพื้นฐาน (กระทรวงศึกษาธิการ 2534) มีเนื้อหาวิชาดังนี้

บทที่ 1 งานช่างพื้นฐาน

1. งานไม้

2. งานไฟฟ้า

3. งานโลหะ

4. งานเครื่องยนต์

5. งานแบบ

บทที่ 2 เครื่องมือ วัสดุและอุปกรณ์ที่ใช้ในงานช่างพื้นฐาน

1. เครื่องมือที่ใช้ในงานช่างพื้นฐาน

2. เครื่องมือสำหรับวัด
3. เครื่องมือสำหรับตัดและผ่า
4. เครื่องมือประเภทตอก
5. เครื่องมือสำหรับขันและไข
6. เครื่องมือประเภทจับยึด
7. เครื่องมือประเภทไสและตกแต่ง
8. เครื่องมือประเภทเจาะ

วัสดุและอุปกรณ์ที่ใช้ในงานช่างพื้นฐาน

บทที่ 3 ความปลอดภัยในการทำงานช่างพื้นฐาน

สาเหตุของการเกิดอุบัติเหตุในงานช่าง
 ความถี่ของอุบัติเหตุที่เกิดขึ้นในงานช่าง
 กฎโรงงานและการป้องกันอุบัติเหตุในโรงงาน
 กฎแห่งความปลอดภัยในการปฏิบัติงานช่างพื้นฐานในโรงงาน
 หลักปฏิบัติทั่วไปเกี่ยวกับการปฏิบัติงานช่างพื้นฐาน
 การป้องกันอุบัติเหตุในงานช่างพื้นฐาน
 การแก้ไขปัญหาเมื่อเกิดอุบัติเหตุ

บทที่ 4 การอ่านแบบเบื้องต้น

ลักษณะงานเขียนแบบ
 เครื่องมือและอุปกรณ์ในการเขียนแบบ
 การเขียนและการอ่านภาพสามมิติ
 มาตรฐานในงานเขียนแบบ
 การกำหนดขนาดและเส้นต่าง ๆ
 การเขียนและการอ่านภาพไอโซเมตริก

บทที่ 5 งานไม้ประเภทเครื่องเรือน

ความหมายและความสำคัญของเครื่องเรือน
 ประเภทของเครื่องเรือนที่ใช้ในบ้านพักอาศัยแยกตามวัสดุที่ใช้
 ทักษะพื้นฐานในการซ่อมแซมดัดแปลงเครื่องเรือนภายในบ้าน
 การซ่อมแซมและดัดแปลงเครื่องเรือน
 การเคลือบผิวและการตกแต่งผิวงาน

บทที่ 6 งานไฟฟ้า

เครื่องใช้ไฟฟ้าภายในบ้าน
 อุปกรณ์ไฟฟ้าภายในบ้าน
 การต่อสายไฟฟ้า

การต่อสายไฟฟ้าเข้ากับหลักต่อสาย
วงจรไฟฟ้าและการเขียนแบบไฟฟ้า
การบำรุงรักษาเครื่องใช้ไฟฟ้าในบ้าน

2. การวัดผล (Measurement)

ลัวินและอังคณา สายยศ (ลัวิน สายยศ และอังคณา สายยศ. 2539 : 10) กล่าวว่า การวัดเป็นกิจกรรมประจำวันของมนุษย์ตั้งแต่เกิดจนกระทั่งตาย ตอนเกิดมาแพทย์และพยาบาลจะวัดเรวว่าหนักเท่าไร สูงเท่าไร เพื่อดูว่าน้อยกว่า หรือมากกว่าเกณฑ์มาตรฐานของคนไทยทั่วไป พอโตได้เวลาเข้าโรงเรียนครูก็จะวัดว่ามีความรู้ความสามารถเพียงใดควรสอบผ่านหรือไม่ผ่านต้องเรียนเสริมหรือไม่ ตลอดชีวิตความเป็นนักเรียนไม่ว่าระดับใดจะได้รับการสอบวัดมากที่สุด

การวัดผล พจนานุกรมฉบับราชบัณฑิตยสถาน(ลัวิน สายยศและอังคณา สายยศ.2539 : 10; อ้างอิงจาก ราชบัณฑิตยสถาน 2525 : 735) ให้ความหมายของการวัดว่า เป็นการตรวจสอบขนาด หรือปริมาณของสิ่งต่าง เช่น ส่วนยาว ส่วนกว้าง ส่วนสูง หรือความรู้ เป็นต้น วัดผลการศึกษาหมายถึง ประเมินผลการเรียนการสอน เพื่อค้นหาและพัฒนาสมรรถภาพให้เด็กเรียนดีขึ้น ครูสอนเก่งขึ้น และใช้การตัดสินเที่ยงตรงแน่นอน และยุติธรรมมากขึ้น นักวัดผลส่วนใหญ่(ลัวิน สายยศ และอังคณา สายยศ. 2539 : 10; อ้างอิงจาก Martuza. 1977 ; Ebel. 1979 ; Hopkins and Stanley. 1981) ให้ความหมายของการวัดผลว่า เป็นกระบวนการกำหนดตัวเลขให้แก่สิ่งใดสิ่งหนึ่งอาจจะเป็นวัตถุ สิ่งของหรือบุคคลก็ได้ เพื่อจุดประสงค์ที่จะชี้ให้เห็นความแตกต่างของ คุณลักษณะที่ต้องการจะวัด

การวัดผลเป็นการวัดคุณลักษณะ เช่น ความสูง ความยาว การเรียนรู้ ความกลัว เชาว์ปัญญา เป็นต้น ไม่ใช่เป็นการวัดคน สัตว์ สิ่งของแต่อย่างใดเลย ดังนั้นการวัดผลแต่ละครั้งจึงขึ้นอยู่กับจุดมุ่งหมายว่าจะวัดคุณลักษณะใด จากสิ่งใด ใช้เครื่องมืออะไร หน่วยเป็นอะไร ควรมองให้คลุมดังนี้

- คุณลักษณะที่จะวัด
- วัดจากสิ่งใด
- เครื่องมืออะไร
- หน่วยเป็นอะไร

สี่ประการที่กล่าวถึงจะเรียกว่า ครบคุณลักษณะของการวัดผล

ในการวัดผลมีวิธีการวัดใหญ่ๆ อยู่ 2 อย่างคือ

1. การวัดผลทางตรง หมายความว่า เป็นความสามารถในการวัดสิ่งนั้นๆ ได้โดยตรงจริงๆ สิ่งที่ต้องการวัดมีรูปธรรม เช่น วัดความยาวของโต๊ะ เราสามารถเอาไม้เมตรไปทาบวัดความยาวของโต๊ะออกมาเป็นหน่วยของเมตรได้ทันที การวัดแบบนี้มีโอกาสได้คุณลักษณะที่

เป็นจริงอยู่มาก เราเรียกว่าเป็นการวัดด้านกายภาพ วิทยาการใดที่วัดได้โดยวิธีการนี้จะเจริญได้รวดเร็วเพราะมีเครื่องมือวัดที่เป็นมาตรฐานสากลมีความแม่นยำเที่ยงตรง วิทยาศาสตร์กายภาพ (Physical science) จึงพัฒนาไปได้ไกลมาก

2. การวัดผลทางอ้อม หมายถึง การวัดคุณลักษณะหนึ่งโดยอาศัยวัดจากอีกสิ่งหนึ่ง เช่น การวัดผลการเรียน การวัดเชาวน์ปัญญา การวัดเจตคติ การวัดความกังวลใจ เป็นต้น คุณลักษณะเหล่านี้เอาเครื่องมือไปหาวัดโดยตรงไม่ได้ ต้องผ่านกระบวนการทางสมอง (Mental process) ก่อนเสมอ ผลการวัดที่ได้มาจึงเป็นผลจากการผ่านกระบวนการทางสมองขึ้นหนึ่งก่อน ซึ่งอาจจะเชื่อมั่นได้จริงหรือเปล่านั้นก็ไม่แน่อนนัก แต่ที่แน่ที่สุดก็คือ เราไม่มีโอกาสวัดคุณลักษณะนั้นๆ ตรงไปตรงมา ความคลาดเคลื่อนย่อมมีเป็นธรรมดา

การประเมินผล (Evaluation)

คำนี้มักใช้คู่กับการวัดผลอยู่เสมอแต่จริงๆ แล้ว มีความหมายแตกต่างกัน โดยเฉพาะในการเรียนการสอนหรือการทำกิจกรรมต่างๆ ใช้คำว่าประเมินผลการเรียนกันอยู่เป็นประจำ หลังจากการเรียนการสอนเสร็จสิ้นแล้ว เพื่อจะได้รู้ว่าผลการเรียนของนักเรียนก้าวหน้าเป็นอย่างไร การประเมินผลจึงนิยามว่า เป็นกระบวนการพิจารณาตัดสินที่เป็นระบบครอบคลุมถึงจุดมุ่งหมายที่ตั้งไว้ นั่นคือ ประเมินว่ากิจกรรมที่ทำทั้งหลายเป็นไปตามจุดมุ่งหมายที่ตั้งไว้เพียงใด บางกรณีจึงต้องใช้ปริมาณจากการวัดมาพิจารณาตัดสินด้วยคุณธรรมแล้วลงสรุป บางกรณีไม่ต้องใช้ตัวเลขจากการวัด เป็นแต่เพียงการหาข้อมูลจากด้านอื่นมาประกอบการพิจารณาตัดสิน เช่น ประวัติ ระเบียบสะสม เป็นต้น การประเมินจึงมีความหมายกว้างและคลุมกว่าการวัดผล การประเมินผลจึงมักใช้ข้อมูลของการวัดผลมาพิจารณาตัดสินดำเนินการเป็นขั้นตอนต่อเนื่องจากการวัดผลที่ดีจึงเป็นฐานของการประเมินผลที่ดีด้วย บางทีการนิยามการประเมินผลจึงมองในแง่การอธิบายปริมาณหรือตัวเลขจากการวัดและ/หรือ การอธิบายข้อมูลเชิงคุณภาพจากสิ่งไม่ต้องวัดรวมกันเข้ากับการพิจารณาตัดสินอย่างมีคุณธรรม การประเมินผลจึงเป็นเรื่องของการใช้เหตุผลเป็นฐานในการพิจารณาด้วยว่า อะไรเหมาะ อะไรดี อะไรควร เป็นต้น (ลัวัน สายยศ และอังคณา สายยศ. 2539 : 12; อ้างอิงจาก Gronlund. 1981)

การวัดผลที่ดีต้องวัดได้ครอบคลุมในสิ่งที่ต้องการวัด และมั่นใจว่าสามารถวัดสิ่งนั้นได้แน่นอนด้วย การวัดให้ครอบคลุมมีผู้เสนอไว้หลายกลุ่มแต่นำเชื่อถือคือแนวคิดของ บลูมและคณะ เสนอการวัดไว้ 3 ด้าน คือ

1. ด้านสติปัญญา (Cognitive Domain) เป็นด้านของการรู้คิด มีผิด มีถูก พฤติกรรมสำคัญที่ใช้วัดด้านนี้คือ ความจำ ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ และการประเมินค่า แต่ละพฤติกรรมมีลักษณะเฉพาะของมันแต่บางส่วนต้องอาศัยซึ่งกันและกัน

2. ด้านความรู้สึก (Affective Domain) เป็นการวัดสภาพการเปลี่ยนแปลงของจิตใจ

เมื่อมีสิ่งหนึ่งสิ่งใดมากระทบพบแล้วเกิดการรับรู้ เกิดความสนใจ อยากจะเกี่ยวข้องกับ จิตใจ ชี้นำในสิ่งนั้นจนรู้คุณค่าทั้งทางดีและไม่ดี เมื่อรู้คุณค่าหลายๆ อย่างแล้วจะเกิดความรู้สึกซาบซึ้งในสิ่งที่ประพจน์ปฏิบัติตาม นั่นคือ เกิดเจตคติขึ้น คุณค่า หรือคุณธรรมรวมทั้งเจตคติจะเป็นตัวการให้เกิดลักษณะนิสัยประจำตัวคน การวัดด้านนี้จะเริ่มจากการรับ การสนองตอบ การรู้คุณค่า การจัดระบบคุณค่า และการสร้างลักษณะนิสัย ตามลำดับ

3. ด้านทักษะกลไก (Psychomotor Domain) เป็นการวัดด้านการกระทำหรือปฏิบัติ นั่นเอง การปฏิบัตินั้นจำเป็นจะต้องใช้มือ เท้า ซึ่งเป็นอวัยวะของร่างกายให้สัมพันธ์กับความคิด ยิ่งสัมพันธ์กันมากเท่าไรก็ยิ่งมีทักษะมากขึ้นเท่านั้น นั่นคือ การปฏิบัติยิ่งเพิ่มความคล่องแคล่ว ว่องไวแม่นยำและรวดเร็ว ยิ่งเป็นสิ่งที่พึงปรารถนาในการฝึกด้านนี้

2.1 การวัดด้านสติปัญญา

การวัดด้านสติปัญญา (ถ้วน สายยศ และอังคณา สายยศ. 2539 : 41-46) บลูมได้เป็นบรรณาธิการ Taxonomy of Educational Objectives พิมพ์เผยแพร่ครั้งแรก ค.ศ. 1956 จำแนกจุดประสงค์ออกเป็นสามย่อยพร้อมกับมีรายละเอียดไว้อย่างน่าศึกษา รายละเอียดด้านใหญ่ๆ มีอยู่ 6 ข้อเรียงตามลำดับความสามารถที่น่าจะเป็นและยังไม่ได้ทดลองก่อน ดังนั้นจึงถือว่าเป็นปรัชญาหรือแนวคิดก็แล้วกัน ส่วนย่อย ๆ ด้านสติปัญญามีดังนี้

1.1 ความรู้ (Knowledge) เป็นความสามารถในการระลึกนึกออกสิ่งใดที่ได้เรียนรู้มาแล้ว คือความจำนั่นเอง ยังแบ่งส่วนย่อยอีกตามลำดับดังนี้

1.1.1 ความรู้ด้านเนื้อหา (Knowledge of Specifics) เป็นความสามารถในการจดจำเนื้อหาของสิ่งที่เรียนหรือประสบพบมา แบ่งออกเป็น 2 อย่างคือ

1.1.1.1 ความรู้เกี่ยวกับศัพท์และนิยาม (Knowledge of Terminology) ความรู้ความจำด้านนี้เป็นสัญลักษณ์ ศัพท์ นิยาม ที่ตกลงกันไว้เพื่อใช้ในเรื่องใดเรื่องหนึ่งเพื่อให้เป็นความหมายที่สะดวก

1.1.1.2 ความรู้เกี่ยวกับความจริง (Knowledge of Specific Facts) เป็นความสามารถในการจดจำสิ่งที่เป็นความจริงที่เรียนรู้มา ความจริงในที่นี้เป็นลักษณะ วันเดือนปี สถานที่ บุคคล เหตุการณ์ เป็นต้น ที่เกิดขึ้นเป็นจริงมาแล้ว

1.1.2 ความรู้เกี่ยวกับวิธีดำเนินการในเนื้อหา (Knowledge of Ways and Means of Dealing with Specifics) ความสามารถด้านนี้เป็นความจำเป็นในด้านวิธีการจัดระบบจัดการศึกษา พิจารณาวิพากษ์วิจารณ์รวมทั้งวิธีแสวงหาความรู้และลำดับขั้นของเวลา แบ่งย่อยออกเป็นดังนี้

1.1.2.1 ความรู้เกี่ยวกับระเบียบประเพณี (Knowledge of Conventions) เป็นความสามารถในการจดจำประเพณี วัฒนธรรม ธรรมเนียม หรือการกระทำที่เป็นนิสัยยึดถือกันในกลุ่ม หรือในเนื้อหาวิชานั้นๆ

1.1.2.2 ความรู้เกี่ยวกับแนวโน้มและลำดับขั้น (Knowledge of Trends and Sequences) เป็นความสามารถในการจำเพื่อหาส่วนที่เกี่ยวข้องกับการแสดงออกทางแนวโน้ม และลำดับขั้นตอนในเรื่องใดเรื่องหนึ่ง

1.1.2.3 ความรู้เกี่ยวกับการจำแนกประเภท (Knowledge of Classifications and Categories) เป็นความจำในเรื่องจัดประเภท กลุ่มชุดของความรู้ในเนื้อหาวิชาที่เรียนรู้อยู่แล้ว

1.1.2.4 ความรู้เกี่ยวกับเกณฑ์ (Knowledge of Criteria) หมายถึง ความจำเกณฑ์ต่างๆ ในการเกิดหลักการ มโนภาพ ความคิดเห็น และอื่นๆ

1.1.2.5 ความรู้เกี่ยวกับวิธีการ (Knowledge of Methodology) เป็นลักษณะการจำวิธีการในการค้นหาความรู้ จำเทคนิค และกระบวนการต่างๆ ที่เคยเรียนมาแล้ว

1.1.3 ความรู้เกี่ยวกับความคิดรวบยอด (Knowledge of the Universals and Abstractions in a Field) ความจำแบบนี้เป็นความจำขั้นสูงขึ้นไป แบ่งออกเป็น 2 อย่าง คือ

1.1.3.1 ความรู้เกี่ยวกับหลักวิชาและขยายนัยทั่วไป (Knowledge of Principles and Generalizations) เมื่อสอนหลักวิชาและการขยายนัยทั่วไปในหลักวิชานั้นๆ แล้ว จุดประสงค์ที่ต้องการจะให้จำสิ่งนั้นๆ ให้ได้

1.1.3.2 ความรู้เกี่ยวกับทฤษฎีและโครงสร้าง (Knowledge of Theories and Structures) ระดับนี้จุดประสงค์เพื่อให้ผู้เรียนสามารถจำทฤษฎีและโครงสร้างของสิ่งที่เรียนมาแล้วให้ได้

1.2 ความเข้าใจ (Comprehension) หมายถึง ความสามารถในการแปลความ ตีความ และขยายความจากสื่อความหมายต่าง ๆ ที่ได้พบเห็น แบ่งส่วนย่อยออก 3 อย่างคือ

1.2.1 การแปลความ (Translation) เป็นความสามารถในการถ่ายทอด ความหมายจากภาษาระดับหนึ่งมาเป็นอีกระดับหนึ่งให้เข้าใจกันง่ายขึ้น

1.2.2 การตีความ (Interpretation) หมายถึงความสามารถในการสรุป การแปลความ มองภาพส่วนรวมมาเป็นใจความสั้น ๆ อย่างได้ความ

1.2.3 การขยายความ (Extrapolation) เป็นความสามารถในการทำนายหรือ คาดคะเนข้อเท็จจริงล่วงหน้าโดยอาศัยแนวโน้มที่มีมาแล้ว

1.3 การนำไปใช้ (Application) เป็นความสามารถในการนำหลักวิชาไปใช้แก้ปัญหา ในสถานการณ์ใหม่

1.4 การวิเคราะห์ (Analysis) เป็นความสามารถในการแยกแยะเพื่อหาส่วนย่อย ๆ ของเหตุการณ์ เรื่องราวหรือเนื้อหาต่างๆ ว่าประกอบด้วย อะไรมีความสำคัญอย่างไร อะไรเป็นเหตุอะไรเป็นผล และที่เป็นไปอย่างนั้นอาศัยหลักการอะไร การวิเคราะห์แบ่งแยกย่อยออกเป็น 3 อย่างคือ

1.4.1 วิเคราะห์ความสำคัญ (Analysis of Elements) เป็นการวิเคราะห์ว่าสิ่งที่ย่อย

นั้นอะไรสำคัญ หรือจำเป็น หรือมีบทบาทที่สุด ตัวไหนเป็นเหตุ ตัวไหนเป็นผล

1.4.2 วิเคราะห์ความสัมพันธ์ (Analysis of Relationships) เป็นการหาความสัมพันธ์ หรือความเกี่ยวข้องส่วนย่อยในปรากฏการณ์หรือเนื้อหานั้น เพื่อนำมาอุปมาอุปไมย

1.4.3 วิเคราะห์หลักการ (Analysis of Organizational Principles) เป็นความสามารถที่จะจับเค้าเงื่อนของเรื่องราวนั้นว่ายึดหลักการใด มีเทคนิค หรือยึดคติใด

1.5 การสังเคราะห์ (Synthesis) เป็นความสามารถในการผสมผสานส่วนย่อยต่างๆ เข้าด้วยกัน เพื่อให้เป็นสิ่งใหม่อีกรูปหนึ่ง มีคุณลักษณะโครงสร้าง หรือหน้าที่ใหม่ แปลกแตกต่างไปจากของเดิม แบ่งย่อยออกเป็น 3 อย่าง

1.5.2 สังเคราะห์ข้อความ (Production of Unique Communication) เป็นความสามารถในการสังเคราะห์ข้อความโดยสื่อหรือโดยการพูด การเขียน การวิพากษ์วิจารณ์หาข้อยุติบางประการ

1.5.3 สังเคราะห์แผนงาน (Production of Plans and Proposed Set of Operations) เป็นความสามารถด้านการวางแผนหรือการจัดกระทำการออกแบบหรือการกระทำให้ลักษณะเดียวกันนี้

1.5.4 การสังเคราะห์ความสัมพันธ์ (Derivation of a Set of Abstract Relations) เป็นความสามารถในการนำนามธรรมย่อยๆ มาสัมพันธ์กันเกิดเป็นทฤษฎี สมมุติฐาน สูตร หรือกฎขึ้น

1.6 การประเมินค่า (Evaluation) เป็นความสามารถในการพิจารณาตัดสินสรุปเกี่ยวกับคุณค่าของความคิดทุกชนิด เพื่อเปรียบเทียบเกณฑ์ หรือมาตรฐานที่กำหนดให้แบ่งย่อยออกเป็น 2 อย่างคือ

1.6.2 ประเมินค่าโดยอาศัยข้อเท็จจริงภายใน (Judgment in terms of Internal Evidence) การประเมินแบบนี้พิจารณาความถูกต้อง สมเหตุสมผล ความสอดคล้อง โดยอาศัยเกณฑ์ภายในของสิ่งนั้นเป็นสำคัญ

1.6.3 ประเมินค่าโดยอาศัยเกณฑ์ภายนอก (Judgment in terms of External Criteria) การประเมินแบบนี้อาศัยเกณฑ์หรือมาตรฐานจากภายนอกเอาไว้มากำหนด เปรียบเทียบ เกณฑ์เหล่านี้อาจเป็นเกณฑ์ที่สังคม หรือระเบียบประเพณีกำหนดไว้ก็ได้

2.2 การวัดภาคปฏิบัติ (Measurement of Performance)

ความหมายของการวัดภาคปฏิบัติ

ได้มีผู้ให้ความหมายของการวัดภาคปฏิบัติไว้หลายท่านดังนี้

ทักแมน (Tuckman.1981) ได้ให้ความหมายของ การวัดภาคปฏิบัติว่า(Performance) เป็นการแสดงความรู้ ความเข้าใจ ความคิด มโนทัศน์ และอื่นๆ คือ พุทธิพิสัย และทักษะพิสัย (Cognitive and Psychomotor Domain)

ไพศาล หวังพานิช (ไพศาล หวังพานิช. 2526) ได้ให้ความหมาย การวัดภาคปฏิบัติ หมายถึงการวัดผลงานภาคปฏิบัติ หรือความสามารถในการปฏิบัติ เป็นการวัดที่ให้ผู้เรียนได้แสดงพฤติกรรมตรงออกมาด้วยการกระทำ โดยถือว่าการปฏิบัติเป็นความสามารถในการผสมผสาน หลักวิธีการต่างๆ ที่ได้รับการฝึกฝนมา ให้ปรากฏออกมาเป็นทักษะของผู้เรียน

ชูศักดิ์ เปลียนงู (ชูศักดิ์ เปลียนงู. 2539) กล่าวว่าเป็นการประเมินที่มีทั้ง กระบวนการ(Process) และผลผลิต(Product) ดังนั้นในการวัดผลควรมีการวัดผลเรื่อง ความสัมพันธ์ (Co-ordination) ระหว่างสายตา มือ จังหวะของการทำงาน เวลาของการแสดงทักษะ ทำทาง ความถูกต้องตลอดจนความเร็วในการปฏิบัติงาน

สมบูรณ์ ชิตพงศ์ (สมบูรณ์ ชิตพงศ์. 2522) ให้ความหมายของการวัดภาคปฏิบัติว่าเป็นแบบทดสอบที่มีจุดประสงค์ที่ต้องการให้ผู้สอบได้ปฏิบัติ การทดสอบแบบนี้ต้องการวัดวิธีการ(Process) หรือดูผลงาน (Product) ในการปฏิบัติ เช่น การทดสอบภาคปฏิบัติในวิชา ศิลปะ งานช่างอุตสาหกรรม สุขศึกษา พลศึกษา

พวงแก้ว ปุณยนกและสุวิมล ว่องวาณิช (พวงแก้ว ปุณยนก และ สุวิมล ว่องวาณิช. 2524) ได้กล่าวว่า การวัดภาคปฏิบัติ มีความหมายครอบคลุมทั้ง Cognitive Skills และ Non cognitive Skills โดยใช้แทน ความรู้ ความเข้าใจ ความคิด มโนทัศน์ และอื่นๆ ซึ่งรวมอยู่ในด้านพุทธิพิสัย ทักษะพิสัย และจิตพิสัย โดยสิ่งที่วัดอาจเป็นความสามารถด้านใดด้านหนึ่ง แม้กระทั่งความสามารถด้านภาษา จุดสำคัญอยู่ที่ว่าพฤติกรรม ที่แสดงออกให้เห็นเป็น การตอบสนองต่อสิ่งเร้าในรูปของการปฏิบัติ ทั้งนี้สิ่งเร้าที่นำมาเสนออาจเป็น Verbal หรือ Nonverbal ก็ได้

จากความหมายของ การวัดภาคปฏิบัติ ที่กล่าวมาแล้วข้างต้น จะเห็นว่านักการศึกษา ได้ให้ความหมายของ การวัดภาคปฏิบัติ แตกต่างกันไปตามความคิดเห็นของแต่ละบุคคล แต่จากความหมายที่แตกต่างกันนั้น มีสิ่งที่นักการศึกษาได้กล่าวไว้ตรงกัน คือการวัดภาคปฏิบัติ มีจุดมุ่งหมายให้ผู้สอบ ได้แสดงพฤติกรรมออกมาในขณะที่กำลังทดสอบ ซึ่งประกอบด้วยวิธีการ (Process) และผลงาน (Product) ที่เกิดขึ้นจากการปฏิบัติภายใต้สถานการณ์ที่ผู้ประเมินกำหนดขึ้น

การวัดภาคปฏิบัติสามารถแบ่งตามระดับความเป็นจริงได้ดังนี้ (สุมาลี จันทร ชลธ. 2542 : 168-169)

1. ระดับการรับรู้ (Recognition) การวัดในระดับนี้เป็นการทดสอบว่า

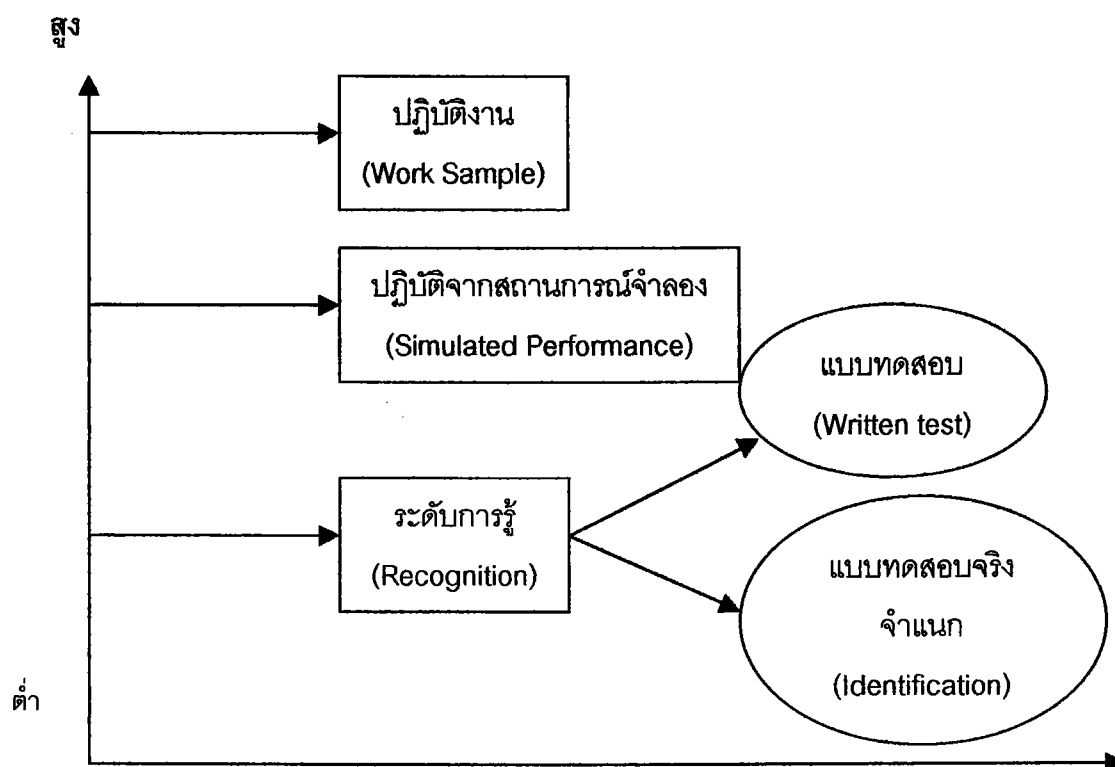
ผู้เรียนได้เรียนรู้ด้านการปฏิบัติมากน้อยเพียงใด ผู้เรียนยังไม่ได้ลงมือปฏิบัติจริง แต่เป็นการทดสอบความรู้ ที่เกี่ยวข้องกับการปฏิบัติ หรือทดสอบพื้นฐานที่จำเป็นในการปฏิบัติงานแบบทดสอบที่ใช้ในการวัดการปฏิบัติในระดับนี้ได้แก่

2. การทดสอบเชิงจำแนก (Identification test) เป็นการทดสอบเพื่อวัดความรู้ความสามารถในการจำแนกเครื่องมือหรือชิ้นส่วนของเครื่องมือ รวมทั้งความรู้พื้นฐานเกี่ยวกับหน้าที่หรือความรู้พื้นฐาน ในการที่จะปฏิบัติงานต่อไป การทดสอบประเภทนี้อาจเป็นการให้เขียนตอบการสอบถามให้ตอบหรือให้แสดงวิธีการใช้งานของเครื่องมือพื้นฐานก่อนที่จะปฏิบัติงานขั้นต่อไป

3. การทดสอบภาคปฏิบัติด้วยแบบทดสอบข้อเขียน (Written test) เป็นการวัดความรู้ด้านทฤษฎีหรือความรู้เกี่ยวกับขั้นตอนการปฏิบัติของผู้เรียนซึ่ง อาจสร้างเครื่องมือในลักษณะของแบบทดสอบวัดผลสัมฤทธิ์ (Achievement test)

4. ระดับการปฏิบัติจากสถานการณ์จำลอง (Simulated Performance) เป็นการทดสอบโดยกำหนดสถานการณ์ที่ใกล้เคียงกับสภาพความเป็นจริง ทั้งนี้เนื่องจากข้อจำกัด บางประการเกี่ยวกับด้านความปลอดภัย ข้อจำกัดด้านเวลา ด้านการจัดการและด้านการลงทุน เป็นต้น การทดสอบจึงจัดสถานการณ์จำลองที่ใกล้เคียงความเป็นจริงเพื่อวัดความสามารถในการปฏิบัติของผู้เรียน เช่น การจัดสภาพสถานการณ์จำลองเพื่อฝึกนักบิน การสอบขับรถในสภาพที่จัดให้ การให้แสดงบทบาทสมมุติตามสถานการณ์ การจัดสถานการณ์จำลองในการลงคะแนนเสียงเลือกตั้ง

5. ระดับการปฏิบัติงานจริงโดยใช้ตัวอย่างงาน (Work Sample) เป็นการทดสอบโดยให้ผู้เรียนปฏิบัติงาน ซึ่งนำมาเป็นตัวอย่างในการสอบวัดความสามารถในการปฏิบัติงาน เป็นการวัดที่มีสภาพความเป็นจริงมากที่สุด ดังนั้น จึงมีลักษณะที่เชื่อถือได้ และมีความตรงมากที่สุด การทดสอบด้วยตัวอย่างงานจึงมักถูกใช้เป็นวิธีการหลักในการวัดภาคปฏิบัติ โดยเฉพาะงานอาชีพ ซึ่งบางอาชีพมีความเกี่ยวข้องกับความปลอดภัย ชีวิตและทรัพย์สิน ดังนั้น จึงต้องฝึกและทดสอบให้ใกล้เคียงกับความจริงมากที่สุด เช่น ในการวัด ความสามารถในการขับรถ ผู้สอบต้องถูกทดสอบภายใต้สถานการณ์จริง หรือถ้าเป็น สถานการณ์จำลองก็มีสภาพเหมือนสถานการณ์จริงทุกอย่างมีสภาพปัญหาที่ผู้สอบต้อง แก้ปัญหาในสถานการณ์จริงได้ในทางอุตสาหกรรม การกำหนดตัวอย่างงานให้ทำ จะต้องเลือก ตัวอย่างงานที่มีขั้นตอนการปฏิบัติ ซึ่งเป็นตัวแทนของการปฏิบัติทั้งหมดที่ได้เรียนมา



ภาพประกอบ 1 ระดับความเป็นจริงของการวัดภาคปฏิบัติ

จุดประสงค์การเรียนรู้และการทดสอบภาคปฏิบัติ

จุดประสงค์การเรียนรู้และการทดสอบภาคปฏิบัติ (สุมาลี จันทร์ชลอ. 2542 : 170-172) กล่าวว่า การสอบวัดสิ่งใดก็ตามต้องคำนึงถึงจุดประสงค์การเรียนรู้เป็นสำคัญ จุดประสงค์ (Objectives) เป็นเป้าหมายซึ่งกำหนดไว้เพื่อเป็นแนวทางในการจัดการเรียนและเป็นเป้าหมายที่คาดว่าผู้เรียนจะบรรลุ ภายหลังจากผ่านกระบวนการเรียนมาแล้ว ส่วนการทดสอบเป็นการวัดความสามารถของผู้เรียน ซึ่งส่วนมากจะกระทำหลังจากที่ผู้เรียนผ่านขั้นตอนการเรียนรู้มาแล้ว ลักษณะของจุดประสงค์มีหลายประเภท เช่นบางจุดประสงค์เป็นเป้าหมายด้านความรู้สึกรู้สึก หรือ จิตใจ การเขียนวัตถุประสงค์ภาคปฏิบัติที่ดี จะต้องเขียนด้วยภาษาที่ชัดเจน และให้แนวทางที่เป็นประโยชน์ต่อการวัด วัตถุประสงค์ที่กำหนดขึ้นนี้หมายถึงเป้าหมายสำหรับผู้เรียน หรือ ผู้เข้าร่วมกิจกรรมทุกคน การเขียนวัตถุประสงค์ควรมีองค์ประกอบต่อไปนี้

1. กำหนดพฤติกรรม (Behavior) เฉพาะที่สามารถวัดหรือสังเกตเห็นได้ เป็นการกำหนดว่านักเรียนจะต้องทำอะไรที่แสดงให้ผู้สอนรู้ว่าเขาได้เรียนรู้แล้ว เช่น ผู้เรียนสามารถแยกคาร์บูเรเตอร์ได้ หมายความว่าสามารถสังเกตได้ว่าเป็นการแยกคาร์บูเรเตอร์ที่ถูกต้องวิธี หรือ พฤติกรรมบางอย่างสามารถวัดหรือสังเกตได้จากทางอ้อม เช่นการให้คำนวณค่าบางค่าผลจากการคำนวณโดยข้อสอบจะใช้เป็นข้อสรุปความสามารถบางประการเกี่ยวกับ

ความสามารถในการคำนวณ จะเป็นองค์ประกอบหนึ่งที่บอกว่าผู้เรียนปฏิบัติงานได้ ตัวอย่างของการกำหนด พฤติกรรมเช่น

- คำนวณค่าของ.....
- แยกชิ้นส่วนของคาร์บูเรเตอร์
- สร้างสะพานจำลอง
- บอกส่วนประกอบของ.....

การเขียนพฤติกรรมในวัตถุประสงค์ภาคปฏิบัติจะต้องหลีกเลี่ยงคำที่เป็นนามธรรมเช่น รู้, เข้าใจ, ซึ่งไม่สามารถสังเกตได้

2. กำหนดเงื่อนไข (Conditions) วัตถุประสงค์ภาคปฏิบัติที่ดี จะต้องบอกเงื่อนไขที่พึงปฏิบัติสำหรับผู้เรียน ภายใต้เงื่อนไขนี้จะกลายเป็นเกณฑ์ที่สำคัญของวัตถุประสงค์ ซึ่งมีความสัมพันธ์กับการปฏิบัติ เงื่อนไขจะเป็นสิ่งที่บอกว่าผู้เรียนจะสามารถปฏิบัติงานได้ ภายใต้สถานการณ์เช่นใด ซึ่งเป็นตัวแทนของเงื่อนไขสถานการณ์ที่เกิดขึ้นในการปฏิบัติงานจริง เช่น วัตถุประสงค์หนึ่งต้องการให้ช่างบักกรีทำการบักกรีท่อทองแดง ในการปฏิบัติดังกล่าว สถานการณ์บางอย่างอาจเกิดขึ้นในขณะที่ปฏิบัติจริง เช่น ท่อทองแดงนั้นอาจอยู่ใต้ถนนบ้านที่มีช่องว่างให้คลานในขณะที่ปฏิบัติงาน ดังนั้นการบักกรีจะต้องถูกสังเกตหรือวัดโดยมีเงื่อนไขเหล่านี้ด้วย ตัวอย่างเงื่อนไขที่รวมกับทักษะปฏิบัติ เช่น

- สามารถตอบคำถามได้โดยไม่ต้องใช้หนังสืออ้างอิง
- กำหนดเครื่องมือการทดลองให้ 3 ชิ้น.....
- กำหนด ค้อน ไม้กระดาน 3 ชิ้น ตะปู 3 ตัว ให้.....

3. กำหนดเกณฑ์หรือระดับที่จะยอมรับ (Criteria) การประเมินความสำเร็จของการปฏิบัติจะต้องระบุมาตรฐานการทำงานของผู้เรียนที่จะยอมรับได้สำหรับงานนั้น เช่น

- ถูกต้องแม่นยำ 100 %
- คำนวณค่าผิดพลาดไม่เกิน ± 0.05
- ทำงานที่กำหนด ได้สำเร็จ 8 ใน 10 ชิ้น
- ทำงานสำเร็จภายใน 20 นาที

ค่ามาตรฐานความถูกต้องแม่นยำคือ 100 % อาจละไว้ในฐานที่เข้าใจโดยไม่ระบุได้ เกณฑ์การกำหนดเวลาในการปฏิบัติงานเป็นสิ่งจำเป็นอีกประการหนึ่งสำหรับการวัดภาคปฏิบัติของงานบางอย่าง เช่น การปฏิบัติงานด้านอุตสาหกรรม จำเป็นต้องกำหนดเวลาในการจุดระเบิดเครื่องยนต์ผู้เรียนที่ไม่สามารถตั้งเวลาได้อย่างรวดเร็วจะไม่สามารถทำงานให้ประสบความสำเร็จได้ ดังนั้นการเขียนจุดประสงค์การเรียนรู้เรื่องนี้จึงจำเป็นต้องกำหนดมาตรฐานเกี่ยวกับเวลาไว้ด้วย ตัวอย่างวัตถุประสงค์การวัดด้านปฏิบัติ

: ผู้เรียนทุกคนสามารถแยกคาร์บูเรเตอร์โดยใช้คู่มือเทคนิคอ้างอิงที่ถูกต้อง 100 %

: ผู้เรียนทุกคนสามารถเตรียมเครื่องมือในการติดตั้งเครื่องปรับอากาศได้ถูกต้องภายในเวลาไม่เกิน 5 นาที

: ผู้เรียนทุกคนสามารถจัดทำอาหารเลี้ยงแขก 5 คน โดยใช้งบประมาณไม่เกิน 400 บาท การกำหนดจุดประสงค์ทำให้มองเห็นงานการปฏิบัติและวิธีการประเมินผลการปฏิบัติงาน กล่าวได้ว่าจุดประสงค์เป็นแนวทางนำไปสู่กระบวนการเรียนการสอนและการจัดกิจกรรม เป็นเป้าหมาย ที่นำไปสู่การสอบวัด ทำให้รู้ว่าหลังจากจบบทเรียนแล้วผู้เรียนควรมีความสามารถใด ดังนั้นจึงควรบอกจุดประสงค์การเรียนรู้ให้ผู้เรียนทราบก่อนด้วย

ลักษณะงานภาคปฏิบัติและแนวคิดในการวัดภาคปฏิบัติ

การวัดด้านการปฏิบัติ เหมาะสำหรับใช้วินิจฉัยพฤติกรรมการปฏิบัติของผู้เรียน ทำให้เห็นจุดเด่นจุดด้อยในการปฏิบัติของผู้เรียน สุมาลี จันทรชลอ (สุมาลี จันทรชลอ. 2542 : 173 – 174) กล่าวว่า การวัดด้านนี้ มีความเกี่ยวข้องกับการวัด 2 ส่วนคือ

1. กระบวนการ (Process)
2. ผลของงาน (Product)

ในการปฏิบัติงานบางอย่างจะเน้นที่กระบวนการมากกว่าผลงาน แต่อาจมีงานปฏิบัติบางอย่างที่ให้ความสนใจผลงานมากกว่าวิธีการทำงาน และในงานบางประเภทให้ความสำคัญทั้ง 2 ส่วน ใกล้เคียงกัน ทั้งนี้ขึ้นอยู่กับธรรมชาติของงานและจุดประสงค์ที่จะสอบวัด

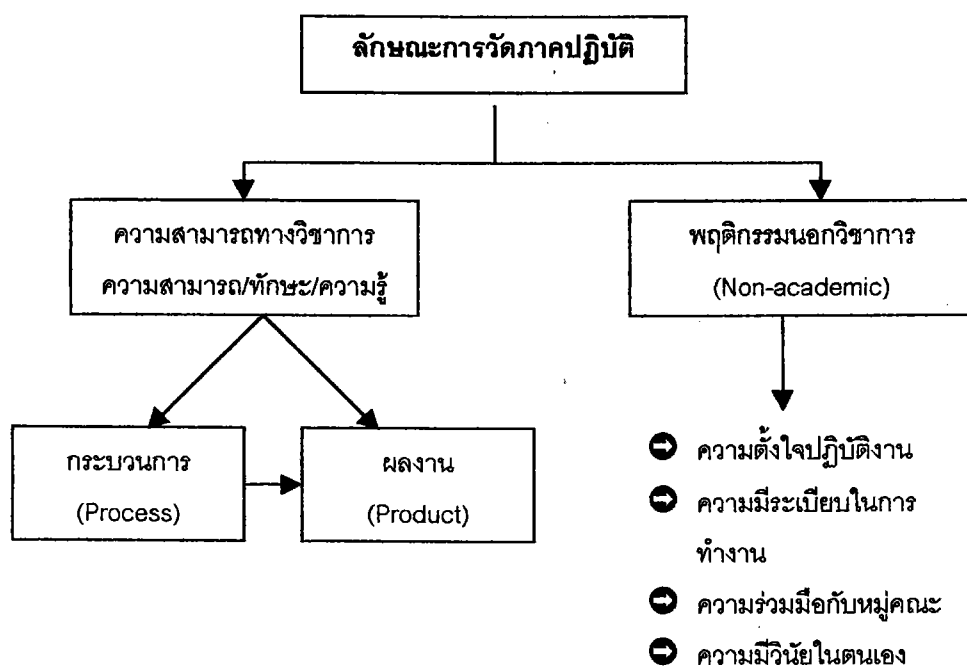
กระบวนการและผลงานมีความสัมพันธ์เกี่ยวข้องกันแม้ว่าจะมีขั้นตอนต่อเนื่องในการทำงานตามลำดับ แต่ก็สามารถแยกแยะการสังเกตและประเมินแยกจากกันได้ เช่น การพิมพ์ดีด อาจสังเกตท่าทาง การวางมือที่แป้นพิมพ์ การมองข้อความที่พิมพ์มากกว่าการมองแป้นพิมพ์ การเคาะแป้นพิมพ์ การวัดความสามารถในการรำไทย การวัดท่ารำจะสังเกตกระบวนการรำ ประเด็นที่จะวัดถูกกำหนดขึ้นจากการวิเคราะห์ท่ารำไทย เช่น การตั้งท่ารำ ความสอดคล้องระหว่างท่ารำและดนตรี การเคลื่อนไหวร่างกาย เป็นต้น ทั้งนี้อาจรวมทั้งพฤติกรรมที่เป็นคุณธรรม เช่น ความขยันหมั่นเพียรในการฝึกซ้อม เมื่อกำหนดประเด็นที่จะวัดแล้ว ต้องให้น้ำหนักความสำคัญของแต่ละประเด็นและกำหนดเกณฑ์การให้คะแนน เมื่อสร้างแบบสังเกตการรำตามประเด็นที่กำหนดแล้วต้องฝึกสังเกตโดยใช้เกณฑ์ที่กำหนด

ผลผลิต หมายถึง ผลสุดท้ายหรือผลงานที่เนื่องมาจากการกระทำ เช่น ในการพิมพ์ จดหมาย ผลผลิตได้แก่จดหมาย ซึ่งอาจพิจารณาจากรูปแบบการพิมพ์ จำนวนคำผิด เป็นต้น ในแง่ของวรรณกรรม ผลผลิต หมายถึงแนวคิดที่เขียนเสนอมา การวัดผลงานต้องกำหนดประเด็นที่จะประเมินและเกณฑ์มาตรฐานที่จะยอมรับในการผ่านเกณฑ์นั้นๆ การกำหนดเกณฑ์ควรคำนึงถึงอายุ วุฒิภาวะของผู้เรียน การตัดสินผลงานหรือผลผลิตจะต้องเลือกเกณฑ์ที่เหมาะสม สำหรับการประเมินผลการจัดพานไหว้ครู ขั้นตอนของการจัดซึ่งเป็นกระบวนการอาจมีความสำคัญไม่มากนัก แต่ผลการจัดจะมีความสำคัญมากกว่านอกจากนี้ผลของงานที่ปฏิบัติอาจแสดงให้เห็น

ขั้นตอนการทำงาน เช่น การออกแบบ การเลือกใช้วัสดุอุปกรณ์ ความคิดสร้างสรรค์ สามารถสังเกตได้จากผลของงานเช่นเดียวกัน ในบางครั้งอาจมีความสับสนที่จะแยกกระบวนการออกจากผลผลิต เช่นการร้องเพลง หรือการร่าง จัดว่าเป็นกระบวนการหรือเป็นผลงาน การกำหนดประเภทของการวัดต้องพิจารณาจากจุดประสงค์ของการสอนประกอบด้วยว่า เกี่ยวข้องกับผลของงานหรือกระบวนการกระทำ จะเห็นได้ว่าการวัดผลการภาคปฏิบัติมีความเป็นอัตนัยมากกว่าการวัดผลสัมฤทธิ์ภาคความรู้ความจำและมีความซับซ้อนและยุ่งยากในการจัดการ การให้คะแนนและแปลความหมายของคะแนน

การสังเกตการปฏิบัติงานบางอย่างของกระบวนการและผลงานมีความสัมพันธ์เกี่ยวข้องกันมาก กรอนลันด์ (Gronlund, 1981) แนะนำให้ประเมินกระบวนการในตอนต้นของการเรียน และประเมินผลงานภายหลังจากที่ผู้เรียนได้เรียนรู้ทักษะพื้นฐานแล้ว

การวัดผลภาคปฏิบัติ ควรต้องสนใจที่จะวินิจฉัยจุดอ่อนของระบบการเรียนหรือกระบวนการเรียนรู้ด้วย นอกจากจะสังเกตกระบวนการปฏิบัติงานและผลของงานแล้ว ควรสังเกตการปฏิบัติในส่วนอื่นๆ ของผู้เรียนซึ่งเป็นพฤติกรรมนอกเหนือจากวิชาการ (non-academic) ในขณะปฏิบัติงานของผู้เรียนได้ด้วย เช่นความตั้งใจปฏิบัติงาน ความมีระเบียบเรียบร้อยในการทำงานความร่วมมือกับหมู่คณะ ความมีวินัยในตนเอง ทักษะคิดต้องงาน เป็นต้น พฤติกรรมดังกล่าวส่งเสริมความเป็นคนที่สมบูรณ์ ดังนั้นจึงควรให้ความสนใจวัดประเด็นต่าง ๆ นี้ด้วย การวัดผลภาคปฏิบัติสรุปเป็นภาพประกอบได้ดังนี้



ภาพประกอบ 2 แนวทางการวัดผลภาคปฏิบัติ

วิธีการวัดภาคปฏิบัติ

พฤติกรรมการปฏิบัติ เป็นพฤติกรรมที่ละเอียดอ่อน ยากต่อการสังเกตและประเมินทั้งนี้เพราะพฤติกรรมการปฏิบัติ เป็นที่รวมของความสามารถหลายอย่างประกอบกัน ต้องการเครื่องมือวัดเฉพาะด้าน จึงสามารถแสดงผลของความแตกต่างได้ นักการศึกษาหลายท่านได้กำหนดลักษณะเกี่ยวกับพฤติกรรม และวิธีการตรวจสอบไว้ ดังนี้

พริสทลีย์ (สุมาลี จันทรชลอ. 2542 อ้างอิงจาก Priestley. 1982) ได้เสนอแนะเทคนิคการวัดเป็น 6 ประเภท ได้แก่

1. ทดสอบผลการปฏิบัติ (Performance test)
2. ทดสอบจากสถานการณ์จำลอง (Situation test)
3. ทดสอบจากแบบสังเกต (Observation test)
4. ทดสอบปากเปล่า (Oral test)
5. ทดสอบจากข้อเขียน (Written test)
6. ทดสอบโดยใช้โปรแกรมการประเมิน (Program requirements)

พวงแก้ว ปุณยกนก และสุวิมล ว่องวานิช (พวงแก้ว ปุณยกนก และสุวิมล ว่องวานิช. 2524) ได้กล่าวว่า การวัดภาคปฏิบัติสามารถทำได้หลายวิธี

1. การวัดโดยการเขียนตอบ เช่น การคัดไทย การวาดรูป หรืองาน ที่มีความเสี่ยงอันตรายสูง ต้องทำการวัดความรู้เกี่ยวกับการปฏิบัติ ด้วยวิธีการสอบข้อเขียนก่อน เช่น การกระโดดร่ม การวัดแรงดันไฟฟ้าในระบบแรงสูง
2. การวัดความจำ (Recognition Test) เป็นการตรวจสอบความรู้ของผู้เรียน ในการทำงานที่มอบหมาย เช่น ให้ตรวจสอบชิ้นส่วนที่ทำงานผิดปกติของเครื่องมือ
3. การวัดโดยการสร้างสถานการณ์จำลอง (Simulated Conditions) เป็นการจำลองสถานการณ์ที่คล้ายคลึงตามสถานการณ์ที่เป็นจริง เช่น การฝึกหัดขับรถ
4. การวัดตัวอย่างของงานที่ได้จากการปฏิบัติงานจริง

ชูศักดิ์ เปลี้นนุ (ชูศักดิ์ เปลี้นนุ. 2537) ได้กล่าวถึงเรื่องการวัดผลภาคปฏิบัติ ว่าการประเมินผลทางภาคปฏิบัติ นั้น มี 2 ลักษณะ คือ

การประเมินผลการใช้ทักษะในการปฏิบัติงาน (Workshop Skill)

การประเมินผลการปฏิบัติทดลอง (Laboratory)

การประเมินผลการใช้ทักษะในการปฏิบัติงาน มุ่งที่จะเน้นพิจารณาฝีมือในการดำเนินงานตามขั้นตอน ความเร็วในการปฏิบัติงาน คุณภาพของงาน ส่วนการประเมินผล การปฏิบัติทดลองนั้นจะเน้นเรื่องประสิทธิภาพ การใช้เครื่องมือในการตรวจสอบอุปกรณ์ การแก้ปัญหาการดำเนินงานตามขั้นตอน ความปลอดภัยในการทำงาน

การประเมินผลการเรียนทั้ง 2 ระบบนี้ มีองค์ประกอบอื่นๆ อีกมากมาย ดังนั้น การวัดผลที่เที่ยงตรง ต้องมีความละเอียดและรอบคอบในการดำเนินงาน ต้องมีการวิเคราะห์ เนื้อหาวิเคราะห์งาน ดังนั้นรูปแบบของการประเมินจะมีหลายรูปแบบ พอสรุปได้ ดังนี้

1. การทดสอบเครื่องมือ วัสดุ และอุปกรณ์ (Object Test) (ชูศักดิ์ เปลี่ยนภู. 2537)

เป็นการวัดความเข้าใจในคุณลักษณะของ อุปกรณ์ เครื่องมือ วัสดุที่เกี่ยวข้องเป็น การวัดโดยตรงซึ่งใช้วัดความสามารถ ดังต่อไปนี้ คือ

1.1 การแยกแยะ (Identify) ชนิด ขนาด คุณภาพ และส่วนประกอบ ที่สำคัญของเครื่องมือและวัสดุ

1.2 ความเข้าใจในกระบวนการบางอย่าง

1.3 ความรู้ในการใช้ เครื่องมือ และวัสดุ การเลือกใช้

1.4 ความรู้ความสามารถในการปรับแต่งเครื่องมือ การปรับแต่ง เครื่องมือในงานพิเศษ

2. การวัดทักษะในการปฏิบัติงาน (Manipulative Performance Test)

เป็นการปฏิบัติงานภายใต้สถานการณ์ที่สร้างขึ้นมา ปกติจะวัด ความสามารถในการด้านต่อไปนี้

2.1 ความเร็วในการปฏิบัติงาน เปรียบเทียบกับเวลาที่กำหนด

2.2 คุณภาพของผลงานที่ทำ ความประณีต ความถูกต้อง

2.3 วัดกระบวนการทำงาน เช่น การเลือกใช้เครื่องมือ อุปกรณ์ การระมัดระวังความปลอดภัย ความสนใจ การสังเกตและความมั่นใจในการปฏิบัติงาน

3. การประเมินด้วยการสังเกต (Observation and Evaluation)

เป็นวิธีการพิจารณาความเจริญงอกงามของนักศึกษาในด้านต่างๆ ที่ไม่สามารถวัดได้ด้วยการสอบข้อเขียน (written test) หรือการสอบภาคปฏิบัติ (Performance test) มีหลักการพิจารณา ดังต่อไปนี้ คือ ต้องมีจุดมุ่งหมายที่แน่นอนว่าจะมีการสังเกต และวัดอะไร และทราบลำดับความสำคัญของขั้นตอนต่างๆ ที่ต้องการจะวัดโดยทั่วไปควรสังเกตในสิ่งต่อไปนี้

3.1 กิจนิสัยในการทำงาน (Work Habit) เช่น ความขยันหมั่นเพียร ทำงานอย่างถูกต้องแม่นยำประณีต มีขั้นตอน ตรงต่อเวลา และมีความมั่นใจในการทำงาน ตั้งใจในการทำงาน ทำงานด้วยความสม่ำเสมอ

3.2 คุณลักษณะต่อส่วนร่วม(Group Attitude)เป็นผู้ที่ได้รับความไว้วางใจ เชื่อถือได้การร่วมมือกับผู้อื่น ความเป็นผู้นำเสียสละมีน้ำใจเป็นนักกีฬา เคารพสิทธิของผู้อื่น

3.3 คุณลักษณะส่วนตัว (Personal Characteristics) เป็นกันเอง สุภาพเคารพผู้อื่นความประณีตของตัวเอง มีความมั่นใจตัวเอง ควบคุมตัวเองให้อยู่ในกฎระเบียบ

4. การสอบด้วยข้อเขียน (Written Examination for Practical Work)

เป็นการทดสอบความรู้เรื่องทั่วไป ในการดำเนินงานประกอบไปด้วย หลักการและความสามารถในการนำหลักการหรือความรู้ในภาคปฏิบัติไปใช้งานมักนิยมใช้ในการประเมินผลการปฏิบัติการทดลอง (Laboratory) มากกว่าการฝึกทักษะ (Workshop) โดยรูปแบบแล้วแบ่งออกเป็น 2 ส่วนคือ

4.1 การทดสอบความรู้ในหลักการที่ใช้ในการปฏิบัติงาน (Practical Theory Test) เกณฑ์การพิจารณา คือ ความรู้ในขีดความสามารถของ เครื่องมือ วัสดุ อุปกรณ์ ความรู้ระบบกระบวนการของการทดลอง สรุปข้อมูล ตัวอย่างของการประเมินด้านนี้ คือการแยกแยะเครื่องมือวัสดุที่ใช้ในห้องปฏิบัติการ การใช้งานของเครื่องมือ อุปกรณ์ การทำงานหรือใช้งานของเครื่องมือ วงจร หรือหลักการทางทฤษฎีต่างๆ ที่เกี่ยวข้อง การใช้ส่วนประกอบของเครื่องมือช่วยงาน ข้อระมัดระวังในการปฏิบัติงาน ขั้นตอนการปฏิบัติงาน ความสามารถในการแก้ปัญหาเมื่อมีเหตุขัดข้อง

4.2 การวัดความรู้ในทฤษฎีที่เรียนมาจากชั้นเรียน (Theoretical Test) ซึ่งเป็นวิธีการสอบข้อเขียนที่ทำกันเป็นประจำ การทดสอบแบบนี้มุ่งที่จะวัดความสามารถในด้านความคิดรวบยอดของเรื่องที่เรียน ความสามารถในการแก้ปัญหา

5. การประเมินผลของชิ้นงาน (Ended-Product Evaluation)

เป็นการประเมินผลของชิ้นงาน ที่ผู้เรียนได้สร้างขึ้นมา มีเกณฑ์ของการพิจารณา คือ เวลาที่ใช้ในการทำงาน ลักษณะการเลือกใช้วัสดุ ความคิดริเริ่มสร้างสรรค์ คุณภาพของชิ้นงานถูกต้องตามแบบ ความประณีต ความถูกต้องของขนาด ซึ่งรายละเอียดของเกณฑ์นี้มีมากมาย ทั้งนี้ขึ้นอยู่กับลักษณะของงานแต่ละอย่าง

จากข้อมูลของนักการศึกษาที่กล่าวมาแล้วนี้ สามารถสรุปได้ว่า การวัดภาคปฏิบัตินั้นสามารถกระทำได้หลายวิธี เช่น การสอบด้วยข้อเขียน การเขียนรายงาน การทดลองงานให้ปฏิบัติในสถานการณ์จำลอง การปฏิบัติตามสภาพที่เป็นจริง หรือการสอบภาคปฏิบัติ การเลือกวิธีวัดภาคปฏิบัติไม่ว่าจะด้วยวิธีการใด หรือรูปแบบใดก็ตาม ขึ้นอยู่กับลักษณะงานที่ต้องการวัดและจุดมุ่งหมายว่า ต้องการวัดพฤติกรรมการเรียนรู้ระดับใด ด้วยเหตุนี้ เมื่อต้องการสร้างแบบทดสอบ จะต้องคำนึงและพิจารณาถึงสิ่งที่เกี่ยวข้อง ดังที่ได้กล่าวมาแล้วด้วย

h 12117
155704 8.3

การสร้างแบบทดสอบการปฏิบัติงาน(Construction of Performance Test)

แบบทดสอบการปฏิบัติงาน เชาวนา ขวลิขำรง (เชาวนา ขวลิขำรง. 2537) สรุป
รวมได้เป็น 2 ลักษณะคือ

1. การสร้างแบบทดสอบการจำแนก (Identification Test)

การสร้างแบบทดสอบการจำแนก (Identification Test) การสร้างแบบทดสอบการ
จำแนก มีลักษณะไม่ต่างจากแบบทดสอบเขียนทั่ว ๆ ไปมากนัก ชนิดของคำถามที่ใช้ก็อาจเป็น
แบบบรรยาย เติมคำ เลือกตอบ จับคู่ หรือ ถูกผิด แต่ที่นิยมใช้กันมากคือ แบบเติมคำ ส่วนการ
ตอบอาจกำหนดให้นักเรียนผู้สอบตอบปากเปล่า หรือเขียนตอบก็ได้ สำหรับงานที่ให้ทำบาง
อย่างก็ให้นักเรียนฟังแล้วตอบ เช่น ให้ฟังดนตรีแล้วบอกระดับเสียง ให้ฟังเสียงคำที่มี “s” “sh”
และ “ch” แล้วบอกว่าคำไหนเป็นเสียงอะไร หรือบางอย่างก็ให้ดูของจริง เช่น ให้ดูสไลด์ด้วย
กล้องจุลทรรศน์แล้วบอกชื่อสิ่งที่เห็น หรือบางอย่างก็ให้ดูจากภาพที่เขียนไว้ในกระดาษข้อสอบ
เช่น ให้บอกชื่อส่วนประกอบของเครื่องพิมพ์ดีดจากภาพร่างโครงสร้างของเครื่องพิมพ์ดีด หรือ
บางอย่างก็เป็นคำถามแบบข้อสอบเขียนตอบทั่ว ๆ ไป เช่นถามให้บอกวิธีการเลือกซื้อผักสด
 เป็นต้น แต่ไม่ว่าจะกำหนดให้นักเรียนทำงานโดยวิธีใดก็ตามข้อสอบก็ยังคงมีลักษณะเป็น
แบบบรรยาย เติมคำ เลือกตอบ จับคู่ หรือถูกผิด นั่นเอง

ดังนั้นหลักในการเขียนข้อสอบการจำแนก จึงใช้หลักการเดียวกับการเขียนข้อสอบ
แบบบรรยาย เติมคำ เลือกตอบ จับคู่ หรือถูกผิด ซึ่งอาจสรุปหลักการสำคัญที่ต้องคำนึงถึงร่วม
กันในการเขียนข้อสอบการจำแนกไม่ว่าจะใช้คำถามชนิดใด ได้ดังนี้

1. คำถามต้องชัดเจน และมีคำตอบที่ถูกต้องเพียงคำตอบเดียวในแต่ละข้อ
2. คำถามต้องสั้น รัดกุม ไม่คลุมเครือ และถูกหลักไวยากรณ์
3. ไม่ควรใช้คำฟุ่มเฟือย และศัพท์เทคนิค ถ้าไม่จำเป็น
4. สำหรับแบบมีคำตอบให้เลือก(เลือกตอบ จับคู่ ถูกผิด)ตัวเลือกต้อง

ไม่ถูกผิดเด่นชัดซึ่งจะเป็นการชี้แนะคำตอบ

5. การเขียนคำชี้แจงสำหรับแบบทดสอบต้องให้แจ่มแจ้ง เกี่ยวกับวิธีการ
ตอบหรือการทำงานที่กำหนดให้ และการบันทึกคำตอบ

2. การสร้างแบบทดสอบสถานการณ์จำลองและตัวอย่างงาน (Simulated Situation and Work Sample Test)

การสร้างแบบทดสอบสถานการณ์จำลองและตัวอย่างงาน (Simulated Situation
and Work Sample Test) เนื่องจากการสร้างแบบทดสอบทั้งสองชนิดนี้ มีลักษณะที่ร่วมกันอยู่
เป็นอันมาก เช่นการเลือกงานที่เหมาะสมกับพฤติกรรมที่จะวัด การควบคุมสถานการณ์ใน
การทดสอบ การกำหนดวิธีการให้คะแนนที่เหมาะสม เป็นต้น ดังนั้นในการสร้างข้อสอบ จึงขอ
กล่าวรวมพร้อมกันทั้งสองชนิด

ลำดับขั้นในการสร้างแบบทดสอบสถานการณ์จำลองและตัวอย่างงานมีดังนี้

1. การวิเคราะห์งานการวิเคราะห์งาน เป็นการพิจารณาว่างานทั้งหมด (Jobs) ตามที่หลักสูตรกำหนดประกอบด้วยชิ้นงานย่อย ๆ (Tasks) อะไรบ้างและแต่ละชิ้นงาน จะต้องมีความเฉพาะอะไรบ้างที่ครูต้องการให้นักเรียนรู้ และฝึกปฏิบัติงานที่ต้องการให้นักเรียนรู้ และฝึกปฏิบัติได้จริง ๆ สิ่งนี้จะช่วยให้ครูผู้สร้างข้อสอบมองเห็นลำดับขั้นที่เกี่ยวข้อง และภาพรวมของการปฏิบัติงานที่ต้องการให้นักเรียนทำได้จริง ๆ

การวิเคราะห์ชิ้นงาน(Task Analysis) แต่ละขั้น อาจทำโดยกำหนดทักษะ หรือความสามารถปลายทางที่ต้องการให้นักเรียนทำได้ (จากจุดประสงค์ของวิชา) แล้วตั้งคำถามว่า “นักเรียนจะต้องมีความสามารถหรือทักษะอะไรบ้าง ในการที่จะนำไปสู่ความสามารถปลายทางที่ต้องการ” จากการตอบคำถามนี้เราก็จะได้ความสามารถหรือทักษะเฉพาะต่าง ๆ ซึ่งอาจมีมากหรือน้อย ทั้งนี้ขึ้นอยู่กับลักษณะของชิ้นงานแต่ละขั้น แล้วตั้งคำถามต่อไปว่า “สำหรับความสามารถหรือทักษะเฉพาะแต่ละอย่างนั้นนักเรียนต้องมีทักษะพื้นฐานอะไรบ้าง” จากนั้นก็ตั้งคำถามไล่ต่อไปเรื่อย ๆ ตามลำดับจนถึงทักษะพื้นฐานขั้นต่ำสุด

2. การเลือกกลุ่มตัวอย่างของทักษะที่จะนำมารวมไว้ใน แบบทดสอบ จากการวิเคราะห์งานจะได้ทักษะต่าง ๆ จำนวนมาก ครูผู้สร้างข้อสอบก็ต้องเลือกทักษะที่สำคัญ และมีความจำเป็นสำหรับงานนั้นๆ มาเป็นกลุ่มตัวอย่างในแบบทดสอบสิ่งที่ควรพิจารณาในการเลือกกลุ่มตัวอย่างของทักษะที่จะนำมารวมไว้ในแบบทดสอบ มีดังนี้

1. เป็นทักษะที่ร่วมกันในการปฏิบัติงานแต่ละขั้น และเป็นตัวแทนของการปฏิบัติงานทั้งหมดได้
2. เป็นทักษะสำคัญที่มีผลต่อคุณภาพของงานที่ปฏิบัติ
3. เป็นทักษะที่สะท้อนให้เห็นถึงสิ่งที่ต้องการเน้นในการสอน
4. เป็นทักษะที่สามารถวัดได้อย่างมีประสิทธิภาพ คือเสียเวลาและค่าใช้จ่ายน้อยแต่ได้ผลดี

3. การกำหนดชิ้นงานหรืออนุกรมของชิ้นงาน หลังจากเลือกทักษะที่สำคัญและจำเป็นอย่างระมัดระวัง เพื่อให้ได้กลุ่มตัวอย่างที่เป็นตัวแทนที่ดีของทักษะทั้งหลายแล้ว ขั้นตอนต่อไปก็เป็นการกำหนดชิ้นงาน หรืออนุกรมของชิ้นงานที่ครอบคลุมทักษะเหล่านั้นได้ทั้งหมด ชิ้นงานที่กำหนดขึ้นนี้ต้องพยายามทำให้เป็นจริงตามธรรมชาติมากที่สุดเท่าที่จะทำได้

สิ่งที่ควรพิจารณาในการกำหนดชิ้นงานมีดังนี้

1. เป็นงานที่ไม่ยากหรือไม่ง่ายเกินไป เพื่อที่จะสามารถจำแนกความสามารถของนักเรียนได้ เพราะถ้ายากหรือง่ายเกินไปก็จะทำให้ผลงานออกมาเหมือนกันหมดทุกคน ไม่สามารถบอกระดับทักษะหรือความสามารถของนักเรียน
2. เป็นงานที่ครอบคลุมทักษะทั้งหลายที่ต้องการวัดนักเรียนจะต้องใช้ทักษะด้านต่าง ๆ ประกอบกันในการปฏิบัติงานนั้น

3. เป็นงานที่อยู่ในวิสัยที่นักเรียนสามารถปฏิบัติได้ และครูสามารถจัดสถานการณ์เพื่อการปฏิบัติงานนั้นได้อย่างเหมาะสม

4. เป็นงานที่เหมาะสมจะปฏิบัติรายบุคคลและสามารถปฏิบัติเป็นกลุ่มได้เพื่อประหยัดทั้งเวลาและแรงงาน โดยที่ผลการวัดนั้นยังคงถูกต้อง และสามารถให้คะแนนอย่างมีความเชื่อมั่น

5. เป็นงานที่สามารถชี้แจงการปฏิบัติได้อย่างชัดเจน โดยบอกจุดมุ่งหมายและขอบข่ายของงาน ผลที่ต้องการ เครื่องมือที่จะใช้ เงื่อนไขต่าง ๆ ในการปฏิบัติ เวลาที่ใช้ในการปฏิบัติ รวมทั้งเกณฑ์การพิจารณาในการตรวจให้คะแนน

4. การกำหนดสิ่งที่จะวัดหลังจากกำหนดชิ้นงานที่จะให้นักเรียนทำในการทดสอบ แล้วก็พิจารณาว่าในชิ้นงานนั้นเราจะวัด กระบวนการ (Process) หรือผลงาน (Product) อย่างใดอย่างหนึ่ง หรือวัดทั้งสองอย่างร่วมกัน ทั้งนี้ย่อมขึ้นอยู่กับลักษณะของชิ้นงานที่ให้นักเรียนทำ เพราะงานบางอย่างไม่สามารถแยกกระบวนการกับผลงานได้เด่นชัด เช่น การพูด การร้องเพลง การเดินร่า การเล่นบาสเกตบอล การเล่นฟุตบอล ฯลฯ ก็จะต้องพิจารณาวัดทั้งกระบวนการ และผลงานร่วมกันไป

5. การสร้างแบบสังเกตการณ์เมื่อกำหนดว่าจะวัดกระบวนการ หรือผลงานหรือทั้งกระบวนการและผลงานแล้วก็สร้างแบบสังเกตการณ์สำหรับใช้บันทึกผลการสังเกตให้สอดคล้องกับสิ่งที่จะวัด ในการสร้างแบบสังเกตการณ์ควรจะนำทักษะต่าง ๆ ที่ได้จากการวิเคราะห์ชิ้นงานมาพิจารณาแยกแยะคุณลักษณะที่จะวัดในลักษณะนั้น ๆ เช่น เกี่ยวกับทักษะการตอกตะปู คุณลักษณะที่วัดก็มี ความตรงของตะปู รอยฉ้อรอบตะปู รอยแตกของเนื้อไม้รอบตะปู เป็นต้น

6. การกำหนดแผนการวัดการกำหนดแผนการวัด หรือการกำหนดน้ำหนักคะแนนในแต่ละส่วน เป็นการวางแผนว่าในการปฏิบัติงานนั้น ๆ จะให้คะแนนส่วนใดบ้าง แต่ละส่วนจะให้คะแนนเท่าไร อาจทำเป็น 2 ขั้นตอนคือ

6.1 กำหนดสัดส่วนของคะแนนระหว่างส่วนหรือขั้นตอนใหญ่ ๆ เช่น ระหว่างกระบวนการ กับผลงาน หรือระหว่างขั้นปฏิบัติ ผลงาน และเวลา การกำหนดสัดส่วนหรือน้ำหนักคะแนนจะต้องเป็นไปตามจุดประสงค์ของงานนั้นว่าเน้นด้านใด แล้วให้คะแนนมากในด้านนั้น

6.2 กำหนดน้ำหนักคะแนนของแต่ละรายข้อ (Item) ในแต่ละส่วนหรือแต่ละขั้นตอนจะประกอบไปด้วยรายข้อย่อย ๆ แต่ละรายข้ออาจมีความสำคัญทัดเทียมหรือไม่ทัดเทียมกันก็ได้ การกำหนดคะแนนแต่ละรายข้ออาจต่างกันตามความสำคัญของข้อนั้น ๆ แต่เมื่อนำคะแนนจากรายข้อเหล่านั้นมารวมกันแล้วจะเท่ากับคะแนนของส่วนหรือขั้นตอนใหญ่ ๆ ตามสัดส่วนที่กำหนดไว้

7. การวางแผนบริหารการทดสอบขั้นสุดท้าย ในการสร้างแบบทดสอบการปฏิบัติงาน คือ การเขียนคำสั่งชี้แจง ข้อแนะนำและแผนงานในการดำเนินการสอบทั้งหมด ซึ่งอาจทำเป็นทางการหรือไม่เป็นทางการก็ได้ตามสภาพการณ์และความเหมาะสม ถ้าผลจากการทดสอบนั้นไม่สำคัญมากนัก และเป็นการทดสอบกลุ่มเล็ก ๆ การดำเนินการสอบอาจไม่เป็นทางการ แต่ถ้าผลการทดสอบนั้นสำคัญ และ/หรือเป็นการทดสอบกลุ่มใหญ่ ๆ ก็จำเป็นต้องทำเป็นทางการ และเตรียมแบบบันทึกผลการสังเกตไว้ให้พร้อม ในบางกรณีอาจต้องเตรียมการสำหรับบันทึกสิ่งที่จะเกิดขึ้นโดยบังเอิญ หรือสิ่งที่เกิดขึ้นระหว่างการปฏิบัติงานอันอาจมีผลต่อการประเมิน การปฏิบัติงานนั้นด้วย

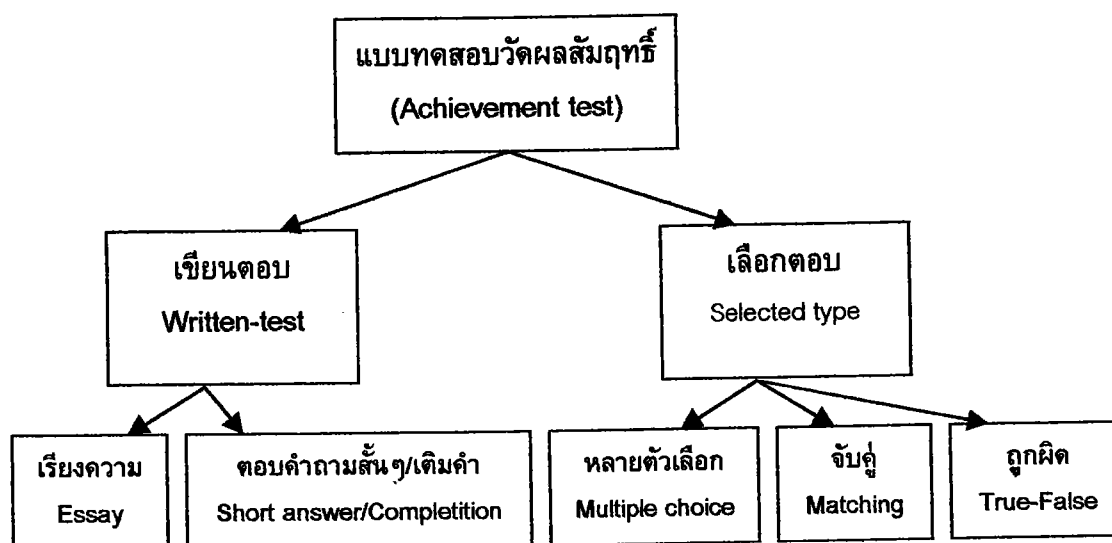
ไม่ว่าการดำเนินการสอบจะเป็นทางการหรือไม่ก็ตาม สถานการณ์ในการดำเนินการสอบจะต้องเป็นมาตรฐานเดียวกันเท่าที่จะสามารถทำได้ อันได้แก่การแบ่งสรรหรือกำหนดเวลา การทำงานตามลำดับขั้น การจัดวางเครื่องมือเครื่องใช้และวัสดุ การกำหนดจำนวนครั้งที่ให้ทดลองทำ ฯลฯ ในการเขียนคำสั่งชี้แจงจึงต้องระบุให้ชัดเจน เพื่อนำให้นักเรียนผู้สอบทุกคนมีความเข้าใจตรงกันในการทำงานชิ้นนั้น จะได้ไม่มีการได้เปรียบเสียเปรียบซึ่งกันและกัน นอกจากนี้ยังจะต้องมีข้อแนะนำสำหรับผู้ที่จะดำเนินการสอบว่า จะต้องทำอะไร จะต้องพูดอะไร จะต้องสังเกตสิ่งใดเป็นพิเศษ จะต้องบันทึกผลการสังเกตอย่างไร เป็นต้น คำสั่งชี้แจงและข้อแนะนำเหล่านี้ จะต้องประกอบด้วยแผนการปฏิบัติงานทั้งหมดในการดำเนินการสอน แผนงานเหล่านี้จะต้องเขียนออกมาอย่างละเอียดเป็นขั้นตอนเมื่อการทดสอบนั้นมีความซับซ้อน หรือเมื่อสอบนักเรียนกลุ่มใหญ่ หรือเมื่อมีผู้ดำเนินการสอบหลายคน แต่ถ้าเป็นการสอบนักเรียนกลุ่มเล็ก ๆ จำนวนน้อย ๆ และมีผู้ดำเนินการสอบเพียงคนเดียว อาจทำเป็นโครงร่างหรือบันทึกสั้น ๆ โดยไม่ต้องเขียนรายละเอียดมากนักก็เป็นการเพียงพอ

3. ชนิดของแบบทดสอบ

แบบทดสอบ(Test) (ล้วน สายยศ และอังคณา สายยศ. 2539) หมายถึง ชุดของข้อคำถาม หรือปัญหา ที่ออกแบบสร้างขึ้นอย่างมีระบบและกระบวนการ เพื่อค้นหาตัวอย่างของพฤติกรรมของผู้ที่สอบ ภายใต้เงื่อนไขเฉพาะอย่าง แบบทดสอบที่ใช้ในห้องเรียนแบ่งเป็นประเภทต่าง ๆ ดังต่อไปนี้

1. แบบทดสอบชนิดให้เขียนตอบ แบ่งเป็นประเภทต่าง ๆ ได้แก่ แบบทดสอบชนิดความเรียง(Essay – type) และแบบทดสอบชนิดให้ตอบคำถามสั้น ๆ หรือให้เติมคำตอบ
2. แบบทดสอบชนิดเลือกตอบ(Selected – type) แบ่งเป็น แบบทดสอบชนิดจับคู่(Matching) ถูก-ผิด (True – False) และชนิดหลายตัวเลือก (Multiple – choice)

ประเภทของแบบทดสอบ ดังแสดงในภาพประกอบต่อไปนี้(สุมาลี จันทรชลด.2542)



ภาพประกอบ 3 ประเภทต่าง ๆ ของแบบทดสอบวัดผลสัมฤทธิ์

1.1 แบบทดสอบความเรียง(Essay Test)(ล้วน สายยศ และอังคณา สายยศ.2539 :86 -120) แบบทดสอบแบบนี้มีจุดประสงค์วัดความสามารถในการบรรยาย อธิบายและแสดงเหตุผลตามความคิดเห็นของตน อาจจำกัดความยาวหรือให้เขียนตอบได้ตามสบายก็ได้ การวัดแบบนี้ถ้าตรวจให้คะแนนทั้งด้านใช้ภาษาและความมีเหตุผล หรือบรรยาย อธิบายดี แต่บางวิชาไม่ได้มองด้านภาษา ดังนั้นการตอบในวิชานั้นอาจให้เหตุผล หรือบรรยาย อธิบายดี แต่เขียนภาษาผิด ๆ ถูก ๆ คะแนนจะให้อย่างไร ผู้ตรวจจะต้องคิดให้รอบคอบอย่าให้เกิดความลำเอียง (Bias) ในการพิจารณาข้อสอบข้อนั้น ในการตรวจให้คะแนนข้อสอบความเรียงจึงต้องสร้างเกณฑ์ไว้ให้ดี มีแนวการตรวจตรงกัน ในการสร้างข้อคำถามแบบความเรียงมีหลักสำคัญดังนี้

1. ดูจุดประสงค์ของการสอบก่อน แล้วจึงเขียนข้อคำถาม เช่น จุดประสงค์ต้องการให้นักเรียนอธิบายทฤษฎีที่สอนมาแล้ว ว่านักเรียนสามารถจำได้ถูกต้องหรือไม่ การเขียนข้อสอบก็ต้องเน้นให้นักเรียนจำทฤษฎีที่สอนมา ไม่ใช่ถามให้หาเหตุผลเปรียบเทียบ ซึ่งจะเป็นการถามที่ผิดจุดประสงค์ สรุปก็คือ ถามให้ตรงจุดประสงค์การเรียนรู้ที่ต้องการวัด

2. การใช้คำถามที่มีความกระชับ ด้วยหลักการถามและหลักภาษา ทั้งนี้ก็เพื่อไม่ให้ผู้สอบเข้าใจผิดพลาด หรือมัวตีความหมายของคำถามที่เป็นได้หลายแง่หลายมุมจนเกินไปเพราะคำถามเพื่อถามวิชานั้น ๆ ไม่ใช่เขียนคำถามเพื่อหลอกล่อนทางภาษา

3. คำถามหนึ่ง ๆ ควรเป็นเรื่องเดียว ทั้งนี้ก็เพื่อให้ผู้ตอบสามารถตอบตรงเป้าหมายที่ ผู้ถามต้องการ เพราะถ้าผู้ตอบตอบผิดก็ผิดไปเลย ถ้ามีหลายเรื่องอาจผิดบ้าง ถูกบ้าง การให้คะแนนจะลำบาก การอธิบายในเรื่องเดียวสามารถวัดได้ดีกว่าในข้อเดียวกัน แต่มีหลายเรื่อง

4. คำถามควรคำนึงถึงเวลาให้ผู้ตอบทำการตอบ คำถามบางอย่างใช้เวลาหลายชั่วโมงขณะที่ผู้ให้สอบมีเวลาแค่ 1 ชั่วโมงผลการตอบก็จะไม่สมบูรณ์ทำให้เกิดความคลาดเคลื่อนในการให้คะแนน

5. คำถามทุกคำถามผู้สอบควรทำเฉลยไว้ และวางแผนการให้คะแนนแต่ละส่วนว่าเป็นเท่าไร ในเวลาที่จำกัดไว้คำตอบที่สมบูรณ์ที่สุดควรเป็นอย่างไร เพื่อใช้เปรียบเทียบกับ นอกจากนั้นจะต้องพิจารณาคำตอบที่มีโอกาสเป็นไปได้ ที่ไม่จำเป็นจะต้องตรงกับเฉลยทุกตัว แต่ก็ถูก สามารถให้คะแนนได้ด้วย

1.2 แบบทดสอบเติมคำ (Completion test) แบบทดสอบแบบนี้เป็นการวัดความสามารถในการหาคำ หรือข้อความ มาเติมลงในช่องว่างของประโยคที่กำหนดได้ถูกต้องแม่นยำ โดยไม่มีคำตอบใดมาชี้ไว้ก่อน นอกจากข้อความหรือประโยคที่ให้ไว้เท่านั้น โดยธรรมชาติเป็นการวัดความจำ แต่ถ้าออกดี ๆ ก็สามารถวัด ความคิดได้

การเขียนข้อสอบเติมคำ มักเป็นของความมากกว่าคำถาม แต่ละข้อความหรือประโยคเว้นที่ให้เติม 1 หรือ 2 แห่ง แต่ถ้ากำหนดข้อความยาวเป็นสถานการณ์ สามารถเว้นให้เติมได้หลายแห่ง เป็นลักษณะโคลเซทสท์ (Cloze Test) ไปในตัว แต่แบบทดสอบโคลสนั้น กำหนด เติมคำที่ 5 หรือ 7 หรือ อื่น ๆ แล้วแต่ผู้ออกข้อสอบกำหนดนิยมใช้ในข้อสอบภาษาอังกฤษ ในการเขียนข้อสอบเติมคำควรพิจารณาดังนี้

1. ข้อความควรสามารถบังคับคำตอบได้ อย่าให้สามารถเติมได้ถูกต้องหลาย ๆ อย่าง

2. ไม่ควรเอาข้อความโดยตรงจากตำรามาเขียนเป็นข้อสอบ

3. ถ้าไม่จำเป็นไม่ควรเว้นให้เติมตอนข้อความ

4. ถ้าคำตอบเติมคำให้ตอบเป็นตัวเลข ให้กำหนดหน่วยด้วยว่าต้องการให้ตอบเป็นหน่วยอะไร และทศนิยมกี่ตำแหน่ง

5. ช่องให้เติมควรมีขนาดยาวพอให้ผู้ตอบสามารถเติมได้หมด

6. การให้เติมมากกว่า 1 ช่อง เช่น 2 ช่องหรือมากกว่านี้ แต่ละช่องควรมีความสัมพันธ์กัน หรือมีความเกี่ยวพันกัน ไม่ควรเอาคนละเรื่องมาให้เติมในแต่ละช่อง

7. ในวิชาที่เกี่ยวกับจิตวิทยาคำนวณแล้วได้ค่ามา 1 ตัว เช่น วิชคณิตศาสตร์สามารถใช้แบบทดสอบเติมคำได้ดี แต่อย่าลืมหาคำหนดจำนวนทศนิยมว่าต้องการกี่ตำแหน่ง ดังที่กล่าวในข้อ 4

1.3 แบบทดสอบถูกผิด (True-False test) แบบทดสอบแบบนี้วัดความสามารถในการพิจารณาข้อความที่กำหนดให้ว่าถูกหรือผิดใช่หรือไม่ใช่ จากความสามารถที่เรียนรู้มาแล้วโดยทั่วไปจะเป็นการวัดความสามารถด้าน ความจำ แต่ถ้าสามารถพลิกแพลงข้อความให้ดีหน่อย อาจจะสามารถวัดด้านความคิดที่สูงขึ้นได้บางข้อควรระวังในการเขียนข้อความมีดังนี้

1. ข้อความควรแจ่มชัดว่าถูกหรือผิดแน่ๆ
2. ข้อความไม่ควรถามตรง ๆ จากที่สอนในตำรา ควรพลิกแพลงบ้าง ทั้งนี้เพราะถ้าข้อความตรงในตำราทุกประการ เด็กที่ท่องจำมากก็จะตอบได้โดยไม่คิด จะทำให้ข้อสอบไม่มี คุณภาพ
3. ข้อเดียวควรถามเพียงเรื่องเดียว หมายความว่า ข้อความที่จะสร้างเป็นข้อสอบควรเป็นข้อ
4. ความในเรื่องเดียวกัน เช่น เรื่องต้นไม้ก็ให้เป็นเรื่องของต้นไม้เรื่องนกก็ข้อให้เกี่ยวข้องกับนกแต่อย่างเดียว
5. ตัดคำขยายที่ฟุ่มเฟือยออกบางกรณี ผู้เขียนข้อสอบอยากทำให้ข้อความสมบูรณ์แบบ แต่จริง ๆ แล้วพอขยายคำหรือประโยคมากเกินไปกลายเป็นข้อความที่ฟุ่มเฟือยจนเกินไป ถือว่าเป็นข้อสอบไม่ดี
6. ไม่ควรใช้คำปฏิเสธซ้อน
7. ควรใช้ข้อความที่สามารถวัดได้สูงกว่าความรู้ความจำแต่เพียงอย่างเดียว

1.4 แบบทดสอบจับคู่ (Matching test) แบบทดสอบแบบนี้เป็นลักษณะการวางข้อเท็จจริง เงื่อนไข คำ ตัวเลข หรือสัญลักษณ์ไว้ 2 ด้านขนานกัน เป็นแถวตั้ง ก. กับแถวตั้ง ข. แล้วให้อ่านดูข้อเท็จจริงในแถวตั้ง ก. ก่อน ต่อจากนั้นพิจารณาดูว่าจะไปเกี่ยวข้องกับ จับคู่กันได้พอดีกับข้อเท็จจริงไหนในแถวตั้ง ข. ที่กำหนดไว้ ตามธรรมชาติ แล้วแถวตั้ง ก. มักจะน้อยกว่าแถวตั้ง ข. เพื่อให้ได้ใช้ความสามารถในการจับคู่ให้มากขึ้น ถ้ามีจำนวนเท่ากันพอข้อที่ใกล้เคียง ๆ จะหมด ไม่ได้ใช้ความสามารถเลย

ในแถวตั้ง ก (Column ก.) มักจะถือว่าเหตุหรือหลักฐานในการพิจารณา

ส่วนแถวตั้ง ข (Column ข.) ถือเป็นคำตอบ ดังนั้นคำตอบจึงมักเขียนไว้เกินตัวที่เป็นเหตุหรือโจทย์เสมอ

1.5 แบบเลือกตอบ (Multiple Choice) ข้อสอบแบบเลือกตอบพัฒนามาจากข้อสอบแบบความเรียงและข้อสอบแบบเติมคำ ในข้อสอบดังกล่าวเมื่อมีคำถามหนึ่งคำถาม จะมีผลคำตอบแตกต่างกันไปตามความคิดเห็นแต่ละคนที่ตอบข้อคำถาม คำตอบถูกจะมีอยู่เพียงคำตอบเดียวเท่านั้น ดังนั้นผู้ที่ตอบเบี่ยงเบนไปจากคำตอบถูกก็ถือว่าเป็นคำตอบผิด การเลือกตอบผิดมาหลาย ๆ คำตอบมาวางเรียงกับคำตอบถูก แล้วให้ผู้สอบเลือก จึงเป็นลักษณะคำตอบหลาย ๆ คำตอบ หรือจะเรียกว่า ตัวเลือกหลาย ๆ ตัวเลือกก็ได้

๒.๕ ลักษณะของข้อสอบเลือกตอบประกอบด้วยส่วนสำคัญที่สุดสองส่วนคือ ส่วนข้อคำถาม(Stem) และส่วนตัวเลือก(Alternative หรือ Chice) ตัวเลือกยังแยกออกเป็นสองส่วนคือ ตัวเลือกที่เป็นตัวถูก (Key) กับตัวเลือกที่เป็นตัวลวง (Foils หรือ Distractors) ในการเขียนข้อสอบแบบเลือกตอบนั้น รูปแบบของคำถามมีส่วนสำคัญในการที่จะเสริมสร้างคุณภาพของข้อสอบให้ดีขึ้น ส่วนใหญ่แล้วรูปแบบของคำถามจะประกอบด้วย 3 แบบคือ

1. ประเภทคำถามโดดๆ (Single item) ประเภทนี้เป็นการเขียนข้อสอบแต่ละข้อวัดอิสระในตัวของมันเอง ไม่จำเป็นจะต้องอาศัยแหล่งการถามร่วมกัน หรือประกอบกันจากจุดใด จุดหนึ่ง คำถามที่นิยมใช้กันมากได้แก่คำถามลักษณะนี้

1.1 ชนิดคำตอบถูกตัวเดียว การเขียนข้อสอบแบบนี้ผู้เขียนนิยมกันมาก เพราะง่ายแก่การสร้าง คุณภาพไม่ดีมากนัก แต่ก็สามารถวัดพฤติกรรมบางอย่างที่ต้องการระลึกออกมาได้อย่างง่าย ๆ

1.2 ชนิดคำตอบถูกต้องที่สุดหรือดีที่สุด การเขียนข้อสอบแบบนี้ยากขึ้นอีกหน่อยตรงที่ต้องหาตัวเลือกให้เป็นเอกพันธ์ และเป็นตัวเลือกที่แต่ละตัวมีโอกาสถูกต้องกันทั้งนั้นแต่ น้ำหนักของการถูกต้องแตกต่างกันไป เวลาเขียนตัวลวงจึงต้องระวังอย่างยิ่ง ตัวเลือกที่เป็น ตัวถูกต้องพยายามเขียนให้มีน้ำหนักการถูกมากกว่าตัวเลือกอื่น มิฉะนั้นจะเกิดปัญหาในการพิจารณาทำให้ข้อสอบขาดความเป็นปรนัยไป

1.3 ชนิดเติมแห่งเดียว ลักษณะนี้เป็นแบบเติมคำดีๆ นั้นเอง แต่มาตัดแปลงให้เป็นข้อสอบแบบเลือกตอบ การใช้คำถามประเภทนี้ต้องถามหลายข้อ แล้วมีคำชี้แจงให้แจ่มชัด มิฉะนั้นผู้สอบจะไม่สามารถทราบได้ว่าข้อสอบต้องการให้ทำอะไรกันแน่ แบบนี้วิชาภาษานิยมใช้กันมาก

1.4 ชนิดเติมหลายแห่ง ลักษณะนี้คล้ายข้อ 1.3 ที่ต่างกันตรงที่เติมเพิ่มมากขึ้น อาจจะเป็น 2 แห่ง หรือมากกว่า 2 แห่งก็ได้ ยิ่งเติมมากแห่งก็ยิ่งเพิ่มความยากมากขึ้น

1.5 ชนิดหาจุดที่ผิดจากประโยคหรือข้อความ ลักษณะนี้โจทย์จะกำหนดข้อความมาให้ ในข้อความนั้นจะมีจุดที่ผิดอยู่ 1 แห่ง ผู้เขียนข้อสอบจะต้องขีดที่จุดอื่นอีก 4 จุด เพื่อเป็นจุดลวงในการขีดจุดใดจุดหนึ่งนั้นให้ใส่ ก, ข, ค, ง, และ จ เอาไว้เลย

1.6 ชนิดคำตรงข้าม การถามตรงไปตรงมาบางอย่างอาจจะไม่เกิดประโยชน์เท่ากับการถามคำตรงข้าม ในกรณีนี้การเขียนข้อสอบก็ควรจะมีการเขียนคำตรงข้ามกันบ้าง เป็นการทดลองดูความคิดอีกด้านหนึ่ง คำตรงข้ามที่ใช้ควรเป็นคำตรงข้ามที่มีเหตุผลพอที่จะหาตัวลวงและตัวถูกได้ เช่น ถ้าถามว่า คน ตรงข้ามกับคำใด แบบนี้รับรองหาคำตรงข้ามเท่าไรก็ไม่สมเหตุผล การเขียนข้อสอบจึงต้องระวังอย่างมากในเรื่องนี้

1.7 ชนิดเรียงอันดับ แบบนี้วิธีการเขียนเพียงหาข้อปัญหาที่เกี่ยวกับการเรียงลำดับ เหตุการณ์ เรื่องราว เวลา คุณลักษณะ วิธีการ หรือเหตุผล เพื่อใช้ถามผู้สอบว่าจะลำดับสิ่งเหล่านี้ได้มากน้อยเพียงใด

1.8 ชนิดอนุกรม แบบนี้ส่วนใหญ่จะเป็นคณิตศาสตร์มากกว่าอย่างอื่น เพราะลักษณะของอนุกรมนั้น เป็นตัวเลขที่เรียงลำดับค่าอย่างมีกฎเกณฑ์ที่แน่นอน

1.9 ชนิดจำแนกประเภท ข้อสอบประเภทนี้มุ่งวัดว่าผู้ตอบสามารถจำแนกประเภทสิ่งที่อยู่ลักษณะเดียวกันหรือต่างกันได้หรือไม่ การเขียนข้อสอบจึงเน้นอยู่ 2 วิธี คือ หาระหว่างที่อยู่ในพวกเดียวกัน กับหาสิ่งที่ไม่อยู่ในพวกเดียวกัน

1.10 ชนิดถามความสัมพันธ์ แบบนี้ผู้เขียนข้อสอบจะต้องพยายามหาสิ่งที่เกี่ยวกันหรือความสัมพันธ์จะต้องพิจารณาให้ถี่ มีเหตุมีผลที่ยอมรับด้วย จึงจะถือว่าปรนัย มิฉะนั้นแล้วก็เป็นปัญหาเดียวกัน

1.11 ชนิดขาดหรือเกิน การเขียนข้อสอบประเภทนี้ มุ่งให้ผู้ตอบวินิจฉัยความสมบูรณ์ของข้อคำถามดูว่าขาดไปหรือเกินไปโดยไม่จำเป็น ผู้ที่ทำถูกจึงเป็นผู้ที่มีความสามารถในการพิจารณาอย่างดี

1.12 ชนิดอุปมาอุปไมย คำถามแบบนี้คล้ายกับการหาความสัมพันธ์ แต่เป็นอีกรูปแบบหนึ่งที่ควรออกเป็นคำถาม เพราะวัดด้านวิเคราะห์หรือเหตุผลอย่างดี ลักษณะข้อคำถามเป็นการกำหนดความสัมพันธ์ของสิ่งคู่หนึ่ง แล้วให้ขยายอิงไปยังคู่อื่น ๆ

1.13 ชนิดหาตัวร่วม คำถามประเภทนี้เป็นการฝึกให้ผู้ตอบสามารถเก็บใจความสำคัญหรือหัวใจของสิ่งนั้น ซึ่งเป็นคุณสมบัติอย่างหนึ่งที่ร่วมกัน

1.14 ชนิดหาตัวต่าง ลักษณะนี้ข้อคำถามตรงข้ามกับข้อ 1.13 คือให้หาดูว่าสิ่งที่มีอยู่นั้นต่างกันเพราะสิ่งสำคัญใด ตัวรวมอะไรที่ทำให้สิ่งนั้นๆ ต่างกันได้

1.15 แบบสรุปความ แบบนี้อาศัยหลักตรรกวิทยาใช้ในการเขียนข้อสอบผู้เขียนจะต้องนึกหาเหตุใหญ่ที่เป็นจริงก่อน แล้วหาเหตุเล็กๆ ที่เป็นจริงอีก ต่อจากนั้นให้ ผู้ตอบสรุป

1.16 แบบเลือกตอบถูกผิด ในบางสาขาวิชาเขียนข้อสอบเลือกตอบเพื่อให้มีคำตอบถูกเพียงตัวเดียว หรือถูกเพียงตัวเลือกเดียว ผิดหลักวิชาของเขา จำเป็นจะต้องถูกอย่างน้อย 2 ตัวเลือก จึงจะถือว่าถูก เช่น วิชาแพทย์ ดังนั้นแบบทดสอบเลือกตอบถูกผิดจึงพัฒนาขึ้นมาจากกลุ่มนักวัดผลที่อยู่ตามมหาวิทยาลัยแพทย์ก่อน

1.17 แบบเลือกตอบซ้อน แบบทดสอบแบบนี้เป็นลักษณะเลือกตอบ 2 ตอบ ถัดจากโจทย์ถาม ตอนแรกเป็นลักษณะของเงื่อนไข ควรจะมีอย่างน้อย 3 เงื่อนไข ในเงื่อนไขของวิชานั้นแต่ละเงื่อนไขอาจจะผิดบ้างถูกบ้าง ตอนที่ 2 เป็นการเขียนตัวเลือกโดยทั่วไปก็จะเป็นการนำเอาเงื่อนไขที่กำหนดให้มาเขียนใหม่ เพื่อให้ผู้ตอบพิจารณาจากเงื่อนไขหลายๆ ตัว

ซึ่งเปรียบเสมือนตัวเลือกครั้งที่ 1 แล้วมาพิจารณาให้ตัวเลือกครั้งที่ 2 อีกที จึงเรียกว่าแบบเลือกตอบซ้อน

2. ประเภทตัวเลือกคงที่ (Constant choice) ตัวเลือกคงที่เกิดจากตัวเลือกข้อใดในคำถามใดๆ ซ้ำกันอยู่บ่อย ๆ ดังนั้นเพื่อให้คำถามและตัวเลือกมีประสิทธิภาพขึ้น จึงเอาตัวเลือกที่ซ้ำกันมาเป็นตัวเลือกคงที่ แล้วเขียนคำถามเป็นข้อๆ เท่านั้น คำชี้แจงในการทำข้อสอบสำคัญมากจะต้องชัดเจนที่สุด มิฉะนั้นแล้วผู้ตอบจะสับสน ไม่สามารถทำข้อสอบได้ตามจุดประสงค์

3. การสร้างสถานการณ์ (Situational Test) การเขียนข้อสอบแบบนี้จำเป็นจะต้องเลือกสถานการณ์จำลอง ข้อความหรือภาพ มาก่อน แล้วผู้ออกข้อสอบจะต้องถามล้วงลึกเฉพาะในสถานการณ์เท่านั้น จะอาศัยส่วนภายนอกมาตอบถูกไม่ได้ ดังนั้นการเขียนข้อสอบประเภทนี้พึงระวังเป็นพิเศษ จะต้องชี้แนะผู้สอบให้เข้าใจว่าการตอบแต่ละข้อ ใช้สถานการณ์ที่ให้เป็นหลัก ถึงจะผิดหรือแปลกจากความเป็นจริงก็ต้องตอบตามนั้น เพราะถือว่าเป็นสถานการณ์จำลอง

การวิเคราะห์ข้อสอบรายข้อ

การวิเคราะห์ข้อสอบ (Item Analysis) เป็นการหาคุณภาพเป็นรายข้อของแบบทดสอบ โดยการเอาผลที่ได้จากการสอบของนักเรียน มาทำการวิเคราะห์หาคุณภาพเป็นรายข้อ

1. ดัชนีค่าความยากของข้อสอบหรือดัชนีค่าความง่ายของข้อสอบ เป็นดัชนีที่แสดงถึงระดับความง่ายของข้อสอบ ข้อสอบที่เหมาะสมจะต้องมี ค่า P_D หรือ $P_E = 0.50$ และในการคัดเลือกข้อสอบนั้นจะพิจารณาค่า P_D หรือ P_E อยู่ระหว่าง 0.20 ถึง 0.80

ตาราง 1 แสดงการแปลความหมายของค่าความง่าย (P_E) ของข้อสอบ

ค่าดัชนี P_E	ความหมาย
มากกว่า 0.80	ง่ายมาก (ปรับปรุงหรือตัดทิ้ง)
0.60 – 0.80	ค่อนข้างง่าย
0.40 – 0.59	ปานกลาง
0.20 – 0.39	ค่อนข้างยาก
ต่ำกว่า 0.20	ยากมาก (ปรับปรุงหรือตัดทิ้ง)

2. ดัชนีค่าอำนาจจำแนก (Discriminant index) หมายถึง ดัชนีที่บ่งบอกถึงว่าข้อสอบนั้นสามารถจำแนกนักเรียนออกเป็นสองกลุ่ม คือ กลุ่มที่ได้คะแนนสูงหรือกลุ่มเก่ง กับกลุ่มที่ได้

คะแนนต่ำหรือกลุ่มอ่อน ค่าอำนาจจำแนกนี้จะมีความหมายในรูปค่าสัมประสิทธิ์สหสัมพันธ์ ซึ่งมีค่าอยู่ระหว่าง -1 ถึง +1 โดยทั่วไปแล้วข้อสอบที่มีค่าอำนาจจำแนกใช้ได้จะมีค่ามากกว่าหรือเท่ากับ 0.20 และถ้าข้อสอบนั้นมีค่าอำนาจจำแนกใกล้ +1 ก็แสดงว่าข้อสอบนั้นสามารถจำแนกคนเก่งและคนอ่อนได้ถูกต้องสูงมาก

ตาราง 2 แสดงค่าอำนาจจำแนกจะพิจารณาระดับอำนาจจำแนก ดังนี้

ค่าอำนาจจำแนก	ความหมาย
มากกว่า 0.40	ดีมาก
0.30 – 0.39	ดี
0.20 – 0.29	ปานกลาง
0.00 – 0.19	ปรับปรุง
ต่ำกว่า 0.00	ตัดทิ้ง

4. ความเชื่อมั่นของแบบทดสอบ (Reliability)

ความหมายของความเชื่อมั่น

นักการศึกษาและนักจิตวิทยา ได้ให้ความหมายของความเชื่อมั่นไว้ต่างๆ กัน ดังนี้

นันทนอลลี (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Nunnally. 1964 : 59)

กล่าวว่า ความเชื่อมั่นเป็นสัดส่วนระหว่าง ความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการทดสอบ

ลินด์วอลล์ และนิคโค(บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Lindvall and Nitko. 1967 : 126) กล่าวว่า ความเชื่อมั่นเป็นค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการทดสอบสองครั้ง โดยใช้แบบทดสอบฉบับเดียวกันสอบในเวลาต่างกัน

อนาสตาซี (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Anastasi. 1968 L 105) กล่าวว่า ความเชื่อมั่นเป็นความคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันแต่ต่างเวลาและโอกาสกัน

บุญเชิด ภิญโญนนตพงษ์ (2537 : 269) ได้ให้ความหมายไว้ว่า เป็นความคงที่แน่นอนของคะแนนซึ่งได้จากการวัดนักเรียนกลุ่มเดียวกันด้วยแบบทดสอบฉบับเดียวกันหลายๆ ครั้งหรือด้วยแบบทดสอบสองฉบับที่มีลักษณะเสมอเหมือนกัน หรือภายใต้เงื่อนไขตัวแปรอื่นๆ ในการวัดนั้น

สำเร็จ บุญเรืองรัตน์ (2529 : 61) ได้กล่าวว่า ความเชื่อมั่นเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนผลการสอบ

จากความหมายข้างต้นผู้วิจัย ได้ประมวลความหมายของความเชื่อมั่น และสรุปได้ว่า ความเชื่อมั่น เป็นความคงที่แน่นอนของคะแนนในการสอบทุกครั้ง จากผู้สอบกลุ่มเดียวกันด้วย แบบทดสอบฉบับเดียวกัน หรือคู่ขนานกันที่เป็นสัดส่วนระหว่างความแปรปรวนของคะแนนที่ได้จากการทดสอบ ความเชื่อมั่นนี้มีวิธีหลายวิธี ซึ่งงานวิจัยนี้ผู้วิจัยหาความเชื่อมั่นของ แบบทดสอบแบบสัมประสิทธิ์ของความสอดคล้องภายใน (Coefficient of Internal Consistency) ซึ่งแนวคิดของวิธีนี้ กำหนดไว้ว่าแบบทดสอบที่ดีต้องมีเอกภาพภายในการวัด (Functional Unity) กล่าวคือส่วนย่อยของแบบทดสอบฉบับหนึ่ง ๆ จะต้องมีความสัมพันธ์ซึ่งกันและกันเพื่อให้เกิดความเป็นเอกภาพในการที่จะวัดคุณลักษณะหนึ่ง ๆ ความหมายของสัมประสิทธิ์ ความสอดคล้องภายในนี้คือข้อคำถามแต่ละข้อ หรือส่วนย่อยจะมีความเสมอเหมือนกันทุกข้อ หรือทุกส่วน สำหรับวิธีการนั้นผู้วิจัยใช้วิธีการวิเคราะห์ส่วนย่อย วิธีนี้ใช้แบบทดสอบสามฉบับ ไปทำการทดสอบกับบุคคลสามกลุ่ม แล้วนำมาคำนวณหาความเชื่อมั่นของแต่ละฉบับ ด้วย สูตร r_{FR} ของเฟลด์ต์-ราชู (Feldt – Raju) นำค่าความเชื่อมั่นที่ได้ ไปทดสอบความแตกต่าง ด้วย สูตร UX_1 เพื่อสนับสนุนผลการวิจัยว่ารูปแบบคำถามของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง มีผลต่างค่าความเชื่อมั่นมากน้อยเพียงใด

การประมาณค่าความเชื่อมั่น

นักการศึกษาและนักวิจัย ได้กล่าวถึง การประมาณค่าความเชื่อมั่นในหลาย ๆ วิธีที่ แตกต่างกันไป ดังนี้

เฟอร์กูสัน (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Ferguson. 1966 : 365-366) และสแตนเลย์ และฮอปกินส์ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Standley and Hopkins. 1972 : 122-127) ได้กล่าวในทำนองเดียวกันว่า วิธีการสัมประสิทธิ์ความเชื่อมั่น มี 4 วิธีดังนี้

1. วิธีสอบซ้ำ (Test retest method) หรือบางครั้งเรียกว่า สัมประสิทธิ์ของความคงตัว (Coefficient of Stability) เป็นการนำแบบทดสอบฉบับเดียวไปทำการทดสอบกับ บุคคลเดียวกันซ้ำสองครั้งในช่วงเวลาที่แตกต่างกันพอสมควร คะแนนที่ได้จากการทำ แบบทดสอบทั้งสองครั้งมีสหสัมพันธ์กัน ค่าสหสัมพันธ์ที่ได้เป็นค่าสัมประสิทธิ์ความเชื่อมั่นของ แบบทดสอบ

2. วิธีใช้แบบทดสอบคู่ขนาน (Parallel – forms method) เป็นการนำแบบทดสอบที่มีลักษณะคู่ขนานกันหรือเท่าเทียมกัน โดยมีเนื้อหา ค่าเฉลี่ยและความแปรปรวนเท่ากัน ไป ทดสอบในเวลาเดียวกันหรือเวลาที่แตกต่างกันก็ได้ คะแนนที่ได้จากการทดสอบทั้งสองฉบับ มีสหสัมพันธ์ ค่าสหสัมพันธ์ที่ได้เป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

3. วิธีแบ่งครึ่งแบบทดสอบ (Split – half method) เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับบุคคลกลุ่มเดียว แล้วแบ่งครึ่งแบบทดสอบเป็นชุดคะแนนของข้อคู่และชุดคะแนนของข้อคี่ นำคะแนนที่ได้จากการแบ่งครึ่งแบบทดสอบไปหาสหสัมพันธ์กัน จากนั้นปรับขยายด้วยสูตรของสเปียร์แมน – บราวน์ เป็นสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับ

4. วิธีวัดความคงที่ภายในของแบบทดสอบ (Internal – consistency method) เป็นการนำแบบทดสอบฉบับเดียวไปทดสอบกับกลุ่มบุคคลเดียวกัน และนำไปหาสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบโดยวิธีของ กูเดอร์ – ริชาร์ดสัน (Kuder – Richardson) ซึ่งการหาความเชื่อมั่นของแบบทดสอบโดยวิธีของกูเดอร์ – ริชาร์ดสัน นี้ เครื่องมือจะต้องมีลักษณะที่วัดองค์ประกอบร่วมกัน และคะแนนแต่ละข้อต้องอยู่ในลักษณะที่ถ้าทำถูกได้ 1 คะแนน ทำผิดได้ 0 คะแนนเท่านั้น วิธีนี้จะมีสูตรที่ใช้ในการหาค่าความเชื่อมั่นอยู่ 2 สูตร คือ KR – 20 กับ KR-21

อนาสตาซี (บุญเชิด ภิญญอนันตพงษ์. 2537; อ้างอิงจาก Anastasi. 1968 : 105-133) กล่าวว่า การประมาณค่าความเชื่อมั่นมี 4 แบบ คือ

1. สัมประสิทธิ์ของความคงตัว เป็นค่าที่ได้จากการนำแบบทดสอบฉบับเดียวไปทดสอบซ้ำในเวลาที่แตกต่างกัน ได้คะแนนสองชุด นำคะแนนสองชุดไปหาสหสัมพันธ์โดยวิธีอย่างง่าย (Product Moment Correlation) ซึ่งค่าสหสัมพันธ์ที่ได้เป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

2. สัมประสิทธิ์ของความสมมูล เป็นค่าที่ได้จากการนำแบบทดสอบสองฉบับที่มีลักษณะเป็นคู่ขนานกัน คือมีเนื้อหา ค่าเฉลี่ย และความแปรปรวนของคะแนน จากคะแนนแบบทดสอบทั้งสองฉบับมาหาค่าสหสัมพันธ์ ค่าที่ได้เป็นสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

3. สัมประสิทธิ์ของความคงตัวและความสมมูล เป็นค่าที่ได้จากการนำแบบทดสอบสองฉบับที่มีลักษณะคู่ขนานกัน คือมีเนื้อหา ค่าเฉลี่ย และความแปรปรวนของคะแนนจากแบบทดสอบเท่ากัน ไปทดสอบกับนักเรียนกลุ่มเดียวกัน ในเวลาที่ต่างกัน โดยเว้นช่วงเวลาระหว่างการทำแบบทดสอบฉบับที่ 1 และฉบับที่ 2 พอสมควร จากนั้นนำคะแนนจากแบบทดสอบทั้งสองมาหาค่าสหสัมพันธ์ ค่าที่ได้เป็นสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ

4. สัมประสิทธิ์ของความคงที่ภายใน เป็นค่าที่ได้จากการนำแบบทดสอบฉบับเดียวสอบกับนักเรียนกลุ่มหนึ่งเพียงครั้งเดียว และแบ่งครึ่งซึ่งนิยมแบ่งข้อคี่และข้อคู่ นำคะแนนจากการแบ่งครึ่งแบบทดสอบทั้งสองชุดมาหาค่าสหสัมพันธ์ แล้วปรับขยายเป็นความเชื่อมั่นของแบบทดสอบทั้งฉบับ โดยใช้สูตรปรับขยายของสเปียร์แมน – บราวน์

เมห์เรนส์ และเลห์แมน (บุญเชิด ภิญญอนันตพงษ์. 2537; อ้างอิงจาก Mehrens and Lehmann. 1984 : 271-271) ได้กล่าวถึงการประมาณค่าความเชื่อมั่นว่ามีวิธีการ ดังนี้

1. วิธีวัดความคงตัว (Measures of stability)

2. วิธีวัดความสมมูล (Measures of equivalence)
3. วิธีใช้วัดความสมมูลและความคงตัว (Measures of equivalence and stability)
4. วิธีวัดความสอดคล้องภายใน (Measures of Internal – consistency)
 - 4.1 วิธีแบ่งครึ่งแบบทดสอบ (Split – half)
 - 4.2 วิธีของคูเดอร์ – ริชาร์ดสัน (Kuder – Richardson estimates)
 - 4.3 วิธีสัมประสิทธิ์แอลฟา (Coefficient alpha)
 - 4.4 วิธีวิเคราะห์ความแปรปรวนของฮอยท์ (Hoyt 's analysis of variance

procedure)

5. ความเชื่อมั่นของผู้ให้คะแนน (Score Judge reliability)

บุญเชิด ภิญญอนันตพงษ์(2521 :278-300) ได้กล่าวถึงวิธีการประมาณค่าความเชื่อมั่นของแบบทดสอบไว้ 2 แนวทาง ดังนี้

1. แบบสัมประสิทธิ์ของความคงตัว (Coefficient of Stability) เป็นวิธีการคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสองชุด ซึ่งได้จากแบบทดสอบคนละฉบับ หรือฉบับเดียวกันแต่เป็นการสอบต่างเวลากัน ซึ่งจำแนกเป็น 2 วิธี คือ

1.1 วิธีสอบซ้ำ (Test retest Method) วิธีนี้จะหาสัมประสิทธิ์ของความคงตัวของคะแนนที่ได้จากการสอบนักเรียนกลุ่มเดียวกันสองครั้ง โดยทิ้งช่วงเวลาให้ห่างกันพอประมาณ แล้วนำคะแนนที่สอบวัดแต่ละครั้งมาคำนวณหาสัมประสิทธิ์สหสัมพันธ์โดยใช้สูตรของเพียร์สัน

1.2 วิธีคู่ขนาน (Parallel Form Method) วิธีการนี้เป็นการคำนวณสัมประสิทธิ์ของความเสมอเหมือน (Coefficient of equivalence) ของคะแนนการทดสอบตั้งแต่สองฉบับ การประมาณค่าตามวิธีนี้อาศัยแนวคิดที่ว่า แบบทดสอบที่สร้างทั้งสองฉบับไปทดสอบกับนักเรียนกลุ่มเดียวกัน ในเวลาเดียวกัน แล้วนำคะแนนที่ได้จากการสอบมาคำนวณสัมประสิทธิ์สหสัมพันธ์ของคะแนนทั้งสองฉบับ โดยใช้สูตรของเพียร์สัน

2. แบบสัมประสิทธิ์ของความสอดคล้องภายใน (Coefficient of Internal Consistency) แนวคิดของวิธีนี้กำหนดว่าแบบทดสอบที่ดีจะต้องมีเอกภาพภายในการวัด (Functional Unity) กล่าวคือ ส่วนย่อยของแบบทดสอบฉบับหนึ่งๆจะต้องมีความสัมพันธ์ซึ่งกันและกัน เพื่อให้เกิดความเป็นเอกพันธ์ในการที่จะวัดคุณลักษณะหนึ่งๆ สัมประสิทธิ์ของความสอดคล้องภายในนี้หมายความว่า ข้อคำถามแต่ละข้อหรือส่วนย่อย จะมีความเสมอเหมือนกันทุกข้อ หรือทุกส่วน แบ่งออกเป็น

2.1 วิธีแบ่งครึ่งแบบทดสอบ วิธีนี้จะนำแบบทดสอบไปทดสอบกับนักเรียนกลุ่มหนึ่ง แล้วจึงแบ่งแบบทดสอบออกเป็น 2 ส่วนเท่าๆกัน โดยให้ทั้งสองส่วนมีข้อคำถามที่ถามคล้ายคลึงกัน และความยากง่ายของแต่ละข้อคำถามของทั้งสองส่วนมีค่าเท่าๆกัน โดยอาจจะแบ่งออกเป็น

ข้อคู่กับข้อคี่ แล้วนำคะแนนของทั้งสองส่วนไปวิเคราะห์หาสัมประสิทธิ์ความเชื่อมั่นโดยใช้สูตรของ สเปียร์แมน – บราวน์ ,ฮอยต์ หรือรูลอน เป็นต้น

2.2 วิธีวิเคราะห์ส่วนย่อย เนื่องจากวิธีแบ่งครึ่งแบบทดสอบ ไม่สามารถคำนวณค่าความสอดคล้องได้อย่างแท้จริง เพราะการแบ่งครึ่งแบบทดสอบนั้น ลักษณะของความเชื่อมั่นจะเป็นความเสมอเหมือนระหว่างคะแนนข้อคู่กับข้อคี่มากกว่า จึงมีผู้คิดวิธีวิเคราะห์คะแนนแบบทดสอบจากส่วนย่อยอย่างแท้จริง วิธีนี้จะใช้แบบทดสอบฉบับเดียวไปทดสอบกับนักเรียนกลุ่มเดียว แล้วนำคะแนนที่ได้จากการทดสอบมาคำนวณ ค่าความเชื่อมั่นด้วยวิธีต่างๆ เช่น วิธีของคูเดอร์ – ริชาร์ดสัน วิธีสัมประสิทธิ์แอลฟา วิธีวิเคราะห์ความแปรปรวนของฮอยท์ เป็นต้น

ทฤษฎีความเชื่อมั่น (Theory of Reliability)

ในการวัดและประเมินผลการศึกษาตั้งแต่อดีตจนถึงปัจจุบันมีผู้เสนอทฤษฎีของการวัดไว้หลายทฤษฎี เช่น

1. ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical test theory) คือความเชื่อมั่นสามารถอธิบายได้โดยเริ่มต้นจากคะแนนที่สังเกตได้ (Observed Score) สำหรับคะแนนที่สังเกตได้นี้จะประกอบด้วย คะแนนจริง(Ture Score) และคะแนนความคลาดเคลื่อน (Error Score) ดังสมการ

	X	$=$	$T + E$
เมื่อ	X	แทน	คะแนนที่สังเกตได้
	T	แทน	คะแนนจริง
	E	แทน	คะแนนความคลาดเคลื่อน

คะแนนจริง (Ture Score) หมายถึง คะแนนที่ผู้สอบได้รับจากการวัดด้วยเครื่องมือที่มีคุณภาพสูงปราศจากความคลาดเคลื่อน หรือหมายถึง คะแนนเฉลี่ยของผู้สอบซึ่งได้จากการทำแบบทดสอบฉบับเดิมหลาย ๆ ครั้ง โดยมีข้อตกลงว่า จะต้องไม่มีอิทธิพลจากการฝึกฝนความเมื่อยล้า และการเรียนรู้ในการทดสอบซ้ำ

คะแนนความคลาดเคลื่อน (Error Score) หมายถึง ค่าความผิดพลาดที่เกิดขึ้นจากการวัด ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นในลักษณะสุ่ม(Random Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยบังเอิญกับความคลาดเคลื่อนอย่างมีระบบ(Systematic Error) สำหรับความคลาดเคลื่อนประเภทหลังจะไม่มีผลกระทบต่อความเชื่อมั่น ความคลาดเคลื่อนที่เกิดขึ้นโดยสุ่มอาจเป็นไปได้ทั้งทางบวกและทางลบ

ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory) เน้นความสำคัญที่ความเชื่อมั่นและความเที่ยงตรงของการวัด หัวใจสำคัญของความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม คือ ความคิดรวบยอด (Concept) ของการวัดที่ “คู่ขนาน” (Cronbach and others. 1963 : 137) ซึ่งนำไปสู่การพัฒนาวิธีประมาณค่าความเชื่อมั่นของการวัดทางจิตวิทยา

อย่างมากมาย และได้รับการอ้างอิงในหนังสือตำราและวารสารการวัดผลอย่างแพร่หลาย (Feldt and Brennan. 1985 : 108) การประมาณค่าความเชื่อมั่นของการวัด อาศัยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ของการวัดหลาย ๆ ครั้ง จากการสอบซ้ำด้วยแบบทดสอบฟอร์มเดียวกัน หรือสองฟอร์มที่คู่ขนานกัน หรือวิเคราะห์ความแปรปรวนของคะแนนสอบภายในฉบับ จากแต่ละส่วนที่สามารถเปรียบเทียบกันได้ (Comparable Parts)(บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Kristof. 1974 : 491)

2. ทฤษฎีการตอบสนองข้อสอบ(Item Response Theory : IRT) นี้มีพื้นฐานความเชื่อว่าพฤติกรรม การตอบสนองต่อข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงว่าถูกหรือผิด จะถูกกำหนดโดยคุณลักษณะภายในหรือความสามารถที่มีอยู่ในตัวบุคคล ซึ่งเป็นสิ่งที่ไม่สามารถสังเกตได้โดยตรง ทฤษฎีการตอบสนองข้อสอบนี้จะอธิบายความสัมพันธ์ของคุณลักษณะภายในหรือความสามารถที่มีอยู่ในตัวบุคคล กับพฤติกรรม การตอบสนองข้อสอบในรูปของฟังก์ชันทางคณิตศาสตร์ ซึ่งเรียกว่า โมเดลการตอบสนองข้อสอบซึ่งมีหลายรูปแบบ เช่น โมเดล 1, 2, 3, พารามิเตอร์ เป็นต้น

3. ทฤษฎีการสรุปอ้างอิงทางการทดสอบ(Generalizability Theory : GT) ทฤษฎีนี้สามารถที่จะช่วยให้นักวัดผลใช้ข้อมูลจากกลุ่มตัวอย่าง สรุปอ้างอิงถึงนัยทั่วไปของค่าความเที่ยงตรงของการวัดภายใต้เงื่อนไขของการวัดต่าง ๆ กัน อันเป็นจุดอ่อนของทฤษฎี Classical Reliability เพื่อตอบคำถามว่าความเที่ยงตรงควรจะเป็นเท่าไรเมื่อสภาพการวัดมีความหลากหลายในระดับต่าง ๆ กันของเงื่อนไขการวัด เช่น ความยาวของข้อสอบ ลักษณะเนื้อหาที่มุ่งวัด จำนวนครั้งของการสอน วิธีการวัดจำนวนผู้ตรวจให้คะแนน ระดับการวัด เป็นต้น ซึ่งจะช่วยให้นักวัดผลในการตัดสินใจทางการทดสอบว่า ถ้าต้องการใช้เครื่องมืออย่างมีคุณภาพ ควรจะใช้ภายใต้สถานการณ์การวัดลักษณะใด แนวทางของทฤษฎีนี้จะนำไปสู่การสรุปอ้างอิงนัยทั่วไปของความเชื่อมั่นต่อไป

การประมาณค่าความเชื่อมั่นของการวัดนั้น บุญเชิด ภิญโญนนตพงษ์ (2537 : 1-3) ได้จำแนกตามวิธีการคำนวณได้สามแนวทางสำหรับสองแนวทางแรกเป็นการประมาณค่าความเชื่อมั่นจากการวัดซ้ำถ้าสอบซ้ำด้วยแบบทดสอบฉบับเดียวกัน สัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้จากแบบทดสอบทั้งสองครั้งเรียกว่าสัมประสิทธิ์ของความคงที่ (Coefficient of Stability) ถ้าสอบซ้ำด้วยแบบทดสอบที่คู่ขนานสองฟอร์ม(Parallel Test Forms) หรือแบบสลับฟอร์ม (Alternate Forms) ซึ่งอาจสอบติดต่อกันทันที หรือสอบทิ้งช่วงเวลา สัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้จากแบบทดสอบทั้งสองฟอร์ม เรียกว่าสัมประสิทธิ์ของความสมมูล (Coefficient of Equivalent) การประมาณค่าความเชื่อมั่นจากการสอบทิ้งช่วงเวลาด้วยฟอร์มที่คู่ขนาน เป็นวิธีที่ให้ค่าประมาณที่ดีที่สุด เพราะสัมประสิทธิ์ชนิดนี้สามารถสะท้อนถึงผลกระทบจากแหล่งความคลาดเคลื่อนในการวัดหมดทุกแหล่ง อย่างไรก็ตามแนวทางทั้งสองดังกล่าวต้องทำการสอบอย่างน้อยสองครั้ง หรือต้องใช้แบบทดสอบอย่างน้อยสองฟอร์ม โดยเฉพาะแนวทาง

ที่สองมักจะไม่สามารถสร้างแบบทดสอบสองฟอร์มให้คู่ขนานกันอย่างแท้จริงได้ จึงไม่เหมาะสมในทางปฏิบัติ สำหรับแนวทางสุดท้ายเป็นวิธีที่หลีกเลี่ยงการสอบซ้ำ โดยอาศัยคะแนนของแบบทดสอบเพียงฉบับเดียวจากการสอบเพียงครั้งเดียวแล้วคำนวณสัมประสิทธิ์ความสอดคล้องภายในแบบทดสอบ(Coefficient of Internal Consistency)

การประมาณค่าความสอดคล้องภายในได้รับความสนใจเป็นอย่างมากจากนักทฤษฎีทางการวัด นับตั้งแต่สเปียร์แมน (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Spearman.1910) และบราวน์ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Brown.1910) ครอนบัค (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Cronbach and others.1963 : 138 – 139) เริ่มต้นศึกษาเรื่องนี้มาจนถึงปัจจุบัน ได้มีการเสนอสัมประสิทธิ์ในการประมาณค่าความเชื่อมั่นไว้หลายวิธี ซึ่งอาจจัดกลุ่มตามข้อตกลงของระดับความคู่ขนานได้สามรุ่น ดังนี้

รุ่นแรกเป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบที่แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts) มีข้อตกลงว่า แต่ละส่วนของแบบทดสอบที่แบ่งต้องมี

1. ความเป็นเอกพันธ์ในเนื้อหา หรือวัตถุคุณลักษณะเดียวกัน
2. มีคะแนนจริงเท่ากันและความแปรปรวนคลาดเคลื่อนเท่ากัน
3. มีคะแนนสอบเฉลี่ยเท่ากัน
4. มีความแปรปรวนของคะแนนสอบเท่ากัน
5. มีความแปรปรวนร่วมของคะแนนสอบเท่ากัน
6. มีความแปรปรวนร่วมของคะแนนสอบกับคะแนนเกณฑ์ภายนอกเท่ากัน

(บุญเชิด ภิญโญนนตพงษ์. 2537 : 2; อ้างอิงมาจาก Gulliksen. 1950 : 41; Cronbach. 1963 : 139; Feldt and Brennan.1989 : 110) จากข้อตกลงดังกล่าวทำให้ไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิมได้ใกล้เคียงที่ตกลงไว้ ซึ่งทำให้เกิดการละเมิดข้อตกลงอยู่เสมอ อันเป็นผลให้ไม่สามารถประมาณค่าความเชื่อมั่นได้ถูกต้องแม่นยำ ดังนั้นนักทฤษฎีการทดสอบจึงได้พัฒนาวิธีประมาณค่าความเชื่อมั่นรุ่นที่สองที่เหมาะสมขึ้นใหม่เป็นการประมาณค่าความเชื่อมั่นตามแบบจำลองคะแนนจริงสมมูล(Essentially Tau-Equivalent Mode) วิธีนี้ได้ผ่อนปรนเงื่อนไข ข้อ 2. มีคะแนนจริงเท่ากันและความแปรปรวนคลาดเคลื่อนเท่ากัน 3. มีคะแนนสอบเฉลี่ยเท่ากัน และ 4. มีความแปรปรวนของคะแนนสอบเท่ากัน ให้มีความเป็นไปได้ในทางปฏิบัติมากขึ้น โดยผ่อนปรนให้มีคะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วน และผ่อนปรนให้แต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วน และผ่อนปรนให้แต่ละส่วนให้มีคะแนนสอบเฉลี่ยต่างกัน และความแปรปรวนต่างกันเล็กน้อย แต่ยังเป็นไปตามเงื่อนไขข้อ 5. และข้อ 6. แต่ในทางปฏิบัติมีแบบทดสอบบางชนิดอาจต้องแบ่งส่วนให้เหมาะสมตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีขนาดความยาวหรือจำนวนข้อไม่เท่า

กัน ซึ่งมีผลกระทบต่อเงื่อนไข 5. มีความแปรปรวนร่วมของคะแนนสอบเท่ากัน และ 6. มีความแปรปรวนร่วมของคะแนนสอบกับคะแนนเกณฑ์ภายนอกเท่ากัน แม้แต่ส่วนประกอบด้วยจำนวนข้อที่เท่ากันก็ตาม แต่เมื่อนำไปสอบกับกลุ่มตัวอย่างแล้ว ปรากฏว่าแต่ละส่วนการกระจายของคะแนนมากน้อยต่างกันมาก ๆ แสดงว่าความยาวที่แท้จริงของมัน (Function Lengths) หรือความยาวที่เป็นผลมาจากการสอบ (Functionally) ของแต่ละส่วนมีขนาดไม่เท่ากัน (Feldt. 1975 : 558) ดังนั้นจึงไม่สอดคล้องตามแบบจำลองคะแนนสมมูล นักทฤษฎีทางการทดสอบจึงได้นิยามความคู่ขนานขึ้นมาอีกระดับหนึ่ง เรียกว่า ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)

แบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)

บุญเชิด ภิญโญนันตพงษ์ (2537 : 12-16) ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ เป็นความคู่ขนานระดับต่ำที่สุดเมื่อเทียบกับความคู่ขนานสองระดับแรก เพราะได้ผ่อนปรนเงื่อนไขต่าง ๆ เกือบหมด โดยคงไว้เฉพาะเงื่อนไขข้อแรกที่ว่าแต่ละส่วนของแบบทดสอบต้องมีเนื้อหาที่เป็นเอกพันธ์กัน หรือวัดคุณลักษณะเดียวกัน (บุญเชิด ภิญโญนันตพงษ์. 2537; อ้างอิงจาก Kristof. 1974 : 492) ซึ่งถือได้ว่าเป็นข้อตกลงที่สำคัญของการวัดที่ใช้อยู่ทั้งในทฤษฎีการทดสอบมาตรฐานเดิม และทฤษฎีการตอบข้อคำถาม (Item Response Theory) ส่วนข้อตกลงข้อ 2. ได้ผ่อนปรนจากแต่ละส่วนต้องมีคะแนนจริงต่างกันเท่ากับค่าคงที่ (หรือ $T_{gh} = T_h + C_{gh}$ เมื่อ C_{gh} ไม่จำเป็นต้องเท่ากับศูนย์เสมอไป) มาเป็นคะแนนจริงของแต่ละส่วนต้องมี สหสัมพันธ์สมบูรณ์แบบ(perfectly Correlated) นั่นคือ คะแนนจริงของแต่ละส่วนมีค่าเท่ากับหนึ่ง หรือมีความสัมพันธ์เชิงเส้นตรง ดังที่เฟลด์ต์(Feldt and Brennan. 1989 : 110 – 111) เขียนเป็นรูปสมการทั่วไปได้ดังนี้ “ $T_{gh} = (b_{gh}) T_h + C_{gh}$ ” เมื่อ b_{gh} และ C_{gh} เป็นค่าคงที่ ซึ่ง b_{gh} ไม่จำเป็นต้องเท่ากับ 1.0 เสมอ และ C_{gh} ก็ไม่จำเป็นต้องเท่ากับ 0 เสมอไป ค่าคงที่ทั้งสองจะแปรเปลี่ยนไปตามความสัมพันธ์ของแต่ละส่วนของแบบทดสอบ แต่จะไม่ขึ้นกับการเปลี่ยนแปลงของแต่ละบุคคลผลของ b_{gh} เหมือนกับผลขององค์ประกอบที่เกิดจากการเพิ่มหรือลดความยาว(lengthening or Shortening factor) เช่น สมมติว่าส่วนย่อย g มีความยาวเป็น 1.2 เท่าของส่วนย่อย h คะแนนเฉลี่ยของส่วนย่อย g จึงไม่ได้มากเป็น 1.2 เท่าของส่วนย่อย h เพียงอย่างเดียว แต่จะขึ้นอยู่กับค่าคงที่ C_{gh} ที่ไม่เท่ากับศูนย์ด้วย ดังนั้นความแตกต่างของความแปรปรวนของคะแนนสอบของส่วนย่อย g และ h จึงไม่ได้ขึ้นอยู่กับความแตกต่างของความแปรปรวนของคะแนนคลาดเคลื่อนเพียงอย่างเดียว แต่ยังขึ้นอยู่กับความแตกต่างของความแปรปรวนของคะแนนจริงอีกส่วนหนึ่งด้วย ซึ่งไม่เหมือนกับความคู่ขนานแบบคะแนนจริงจำเป็นต้องสมมูล ลักษณะสำคัญดังกล่าวข้างต้น ก็คือการแบ่งแบบทดสอบออกเป็น ส่วน ๆ ที่มีขนาดความยาวไม่เท่ากัน หรือมีขนาดความยาวเท่ากันแต่การกระจายของคะแนนในแต่ละส่วนแตกต่างกันมาก ๆ (หรือความยาวที่ทำหน้าที่มีขนาดไม่เท่า

กัน) ซึ่งนักวัดผลนิยมเรียกว่า ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model) ซึ่งย่อมาจากข้อตกลงที่ว่า $T_{ig} = (b_{gh}) T_h + C_{gh}$ เมื่อ $b_{gh} = 1.0$ แล้ว $T_{ig} = T_h + C_{gh}$ และถ้า $b_{gh} = 1.0$ และ $C_{gh} = 0$ แล้ว $T_{ig} = T_h$ แสดงให้เห็นว่าแบบจำลองคะแนนจริงจำเป็นต้องสมมูล เป็นกรณีเฉพาะของ แบบจำลองคะแนนจริงสัมพันธ์ และแบบจำลองความคู่ขนานแบบมาตรฐานเดิม เป็นกรณีเฉพาะของความคู่ขนานแบบคะแนนจริงจำเป็นต้องสมมูล(บุญเชิด ภิญโญอนันตพงษ์. 2537; อ้างอิงจาก. Feldt Gilmer. 1983 : 99)

แบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model) สมมติว่า คะแนนจริงของส่วนย่อยของแบบทดสอบ K ส่วน มีหน่วยการวัดที่แตกต่างกัน (บุญเชิด ภิญโญอนันตพงษ์. 2537; อ้างอิงจาก Gilmer & Feldt. 1983 : Kristof. 1674) คะแนนจริงของแต่ละส่วน T_g ($g = 1, \dots, K$) มีความสัมพันธ์เชิงเส้นตรงกับคะแนนจริงรวมทั้งฉบับ นั่นคือ $\tau_g T + C_g$ โดยมีข้อจำกัดสองประการคือ ค่า C_g เป็นค่าคงที่โดย $\sum C_g = 0$, และ τ_g เป็นจำนวนจริงบวกโดย $\sum \tau_g = 1$ ถ้า τ_g ทุกค่าเท่ากันหมดแล้ว แบบจำลองคะแนนจริงจำเป็นต้องสมมูล (essentially tau equivalent) จะเป็นกรณีเฉพาะของ แบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model) ถ้าให้ S_T^2 แทนความแปรปรวนของคะแนนจริงรวมทั้งฉบับและให้ S_{Eq}^2 ($g = 1, \dots, K$) แทนความแปรปรวนของคะแนนความคลาดเคลื่อนแต่ละส่วน และให้ $S_g^2 = \tau_g S_T^2 + S_{Eq}^2$ แทน ความแปรปรวนของคะแนนสอบ และให้ $S_{gh} = \tau_g \tau_h S_T^2$ ($g = h$) แทนความแปรปรวนร่วมของคะแนนแต่ละส่วน แบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model) ประกอบด้วยสมการทั้งหมด $K(K + 1) / 2$ สมการ (เช่น S_g^2 และ S_{gh} และกำหนดค่าว่า $X_g = 1$) และมีตัวไม่ทราบค่า ทั้งหมด $2K + 1$ ตัว (เช่น $X_g S_{Eq}^2$ และ S_T^2) เมื่อ $K > 3$ แล้วจะมีจำนวนคำตอบของตัวไม่ทราบค่าจำนวนมากมาย การประมาณค่าความเชื่อมั่นได้ของแบบทดสอบทั้งฉบับในกรณีนี้จะต้องวิเคราะห์ด้วย สถิติกำลังสองน้อยที่สุด และความเป็นไปได้สูงสุด (Maximum Likelihood : ML) (Joreskog. 1971) อย่างไรก็ตาม วิธีการทางสถิติดังกล่าวมีข้อจำกัดอยู่ที่จำเป็นต้องใช้การคำนวณด้วยเครื่องคอมพิวเตอร์ (บุญเชิด ภิญโญอนันตพงษ์. 2537; อ้างอิงจาก.Liou. 1989 : 153 – 154)

การประมาณค่าความเชื่อมั่นได้ของแบบทดสอบตามแบบจำลองคะแนนจริงสัมพันธ์ นอกจากใช้สถิติขั้นสูงวิเคราะห์ด้วยเครื่องคอมพิวเตอร์แล้ว นักทฤษฎีการวัดยังได้ศึกษาค้นคว้าและเสนอเทคนิคที่ใช้เครื่องคำนวณแบบธรรมดาหาค่าความเชื่อมั่นไว้หลายเทคนิคดังนี้

ฮอริสต (บุญเชิด ภิญโญอนันตพงษ์. 2537; อ้างอิงจาก.Horst. 1951 : 368-371) ได้เสนอสูตรความเชื่อมั่นของการวัด เมื่อแบ่งแบบทดสอบเป็นสองส่วนที่มีขนาดความยาวหรือจำนวนข้อไม่เท่ากัน ไว้ตั้งแต่ปี พ.ศ. 2494(ค.ศ. 1951) โดยอาศัยสัดส่วนจำนวนข้อของสองส่วนมากำหนดเป็นสูตรที่ใช้ในการคำนวณถ้าสัดส่วนจำนวนข้อของแต่ละส่วนเท่ากันแล้วจะกลายเป็นสูตรของสเปียร์แมน-บราวน์(Spearman-Brown. 1910) หลังจากนั้นการพัฒนาสูตรความเชื่อมั่นที่แบ่งส่วนความยาวภายในฉบับไม่เท่ากันนี้ก็หยุดชะงักไปจนถึงปี พ.ศ. 2517

(ค.ศ. 1974) คริสทอฟ(บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Kristof. 1974 : 491-499) ได้คิดเทคนิคประมาณค่าความเชื่อมั่นของแบบทดสอบที่วัดเนื้อหาที่เป็นเอกพันธ์กันโดยแบ่งเป็นสามส่วนมีขนาดความยาวไม่เท่ากันและแต่ละส่วนต้องมีสหสัมพันธ์ของคะแนนจริงเป็นเชิงเส้นตรง โดยไม่ใช้จำนวนข้อของแต่ละส่วนมาใช้ใน การคำนวณ อย่างไรก็ตามเทคนิคนี้มีข้อจำกัดที่ต้องแบ่งแบบทดสอบเป็นสามส่วนเท่านั้น และถ้ามีการละเมิดข้อตกลงเกี่ยวกับความเป็นเอกพันธ์ของเนื้อหาแล้วจะทำให้การประมาณค่าต่ำกว่าความเป็นจริง หลังจากนั้นอีกหนึ่งปี เฟลด์ต์ (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Feldt. 1975 L 557-561) ได้พัฒนาแนวคิดของการประมาณค่าความเชื่อมั่นเมื่อแบบทดสอบประกอบด้วยส่วนย่อยที่มีขนาดไม่เท่ากันของคริสทอฟ (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Kristof. 1974) ให้สามารถวิเคราะห์กับ แบบทดสอบที่แบ่งเป็นสองส่วนไม่เท่ากัน โดยกำหนดข้อตกลงเกี่ยวกับความแปรปรวน คลาดเคลื่อนตามทฤษฎีมาตรฐานเดิมเพิ่มขึ้นอีกหนึ่งข้อ คือ ให้ $S_{Eg}^2 = X_g - S_E^2$ ซึ่งข้อตกลงที่เพิ่มขึ้นนี้มีได้ทำให้เกิดข้อจำกัดที่สำคัญประการใดตามมา เนื่องจากได้เพิ่มข้อตกลงดังกล่าว เทคนิคนี้จึงจัดอยู่ในประเภทคะแนนจริงสัมพันธ์ แบบมาตรฐานเดิม (Classical Congeneric) อีกสองปีถัดมา ราฐู (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Raju. 1977 : 549-565) ได้พิจารณาเห็นว่าสูตรสัมประสิทธิ์-แอลฟาของครอนบาค(บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจากCronbach. 1951) แม้จะสามารถใช้วิเคราะห์ ความเชื่อมั่นได้ของแบบทดสอบที่แบ่งเป็นหลาย ๆ ส่วนได้ก็ตาม แต่ถ้าส่วนย่อย ๆ เหล่านั้นมีขนาดความยาวไม่เท่ากันแล้ว สัมประสิทธิ์-แอลฟาจะมีปัญหาทำให้ประมาณค่าที่ต่ำกว่าความเป็นจริง ดังนั้นราฐู (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Raju. 1977)จึงได้พัฒนาสูตรปรับแก้สูตรสัมประสิทธิ์-แอลฟา ให้เป็นสูตรทั่วไปที่ใช้กับแบบทดสอบที่ส่วนย่อยมีความยาวขนาดไม่เท่ากันเรียกว่าสัมประสิทธิ์-เบต้า (Coefficient-β) ถ้าส่วนย่อยมีขนาดความยาว หรือจำนวนข้อเท่ากันแล้ว สัมประสิทธิ์-เบต้า จะให้ค่าเท่ากับสัมประสิทธิ์-แอลฟา เทคนิคนี้สามารถคำนวณจากแบบทดสอบที่แบ่งเป็นกี่ส่วนก็ได้แต่ที่สำคัญต้องทราบจำนวนข้อของแต่ละส่วนย่อย

ต่อมาในปี พ.ศ. 2526 กิลเมอร์ และเฟลด์ต์ (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Gilmer and Feldt. 1983 : 99-111) ได้เสนอผลงานการค้นคว้าในการประมาณค่าความเชื่อมั่นตามแบบจำลองคะแนนจริงสัมพันธ์(Congeneric model) ซึ่งเป็นปริญาณิพนธ์ระดับปริญญาเอกที่ กิลเมอร์(Gilmer) ทำภายใต้คำแนะนำของศาสตราจารย์ ดร.เฟลด์ต์ (Feldt) โดยพยายามลดจำนวนไม่ทราบค่า $K(K+1)/2+1$ ให้เหลือเท่ากับความสัมพันธ์ของตัวแปรจำนวน $2K + 1$ ความสัมพันธ์ แล้วแก้สมการตัวไม่ทราบค่า $2K + 1$ ซึ่งมีเทคนิคในการวิเคราะห์สองสูตร คือ r_{F1} และ r_{F2} สามารถใช้วิเคราะห์แบบทดสอบที่แบ่งเป็นหลาย ๆ ส่วน ($K > 3$) ให้ได้คำตอบที่เป็นค่าเดียวออกมาได้ และสรุปว่า สูตร r_{F2} ง่ายต่อการนำไปใช้ จนกระทั่งปี พ.ศ.2532 เลี้ยว (บุญเขต ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Liou. 1989 : 153 – 163) จากมหาวิทยาลัยแห่งชาติไต้หวัน ได้แก้ปัญหาเทคนิคของคริสทอฟ(บุญเขต ภิญโญนนตพงษ์.

2537; อ้างอิงจาก Kristof. 1974) ให้สามารถนำไปใช้กับแบบทดสอบที่แบ่งส่วนเป็นหลายๆ ส่วน ($K > 3$) ได้สำเร็จ และยังสรุปอีกว่า ถ้ายุบส่วนย่อยของแบบทดสอบให้เหลือสามส่วนแล้วค่าความเชื่อมั่นที่คำนวณด้วยสูตรของคริสทอฟ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Kristof. 1974) จากการยุบเป็นสามส่วนที่เป็นไปได้ทุกวิธีเฉลี่ยแล้วจะเท่ากับค่าความเชื่อมั่นที่คำนวณจากส่วนย่อยทั้งหมดที่คำนวณด้วยสูตรของเลียว (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Lion 1989) ดังนั้นจึงถือได้ว่าเทคนิคของเลียว (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Liou. 1989) เป็นเทคนิคกรณีทั่วไปของสูตรของคริสทอฟ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Kristof. 1974) และในรายงานวิจัยเดียวกันนี้ เลียว (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Liou. 1989) ยังได้ปรับขยายสูตรความเชื่อมั่นของเฟลด์ต์ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Feldt. 1975) ซึ่งเป็นการวิเคราะห์ความเชื่อมั่นตามแบบจำลองคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม (Classical Congeneric) ให้สามารถใช้กับแบบทดสอบที่แบ่งเป็นหลายๆ ส่วน ($K > 3$) ได้ ซึ่งเทคนิคนี้เป็นวิธีที่สามารถคิดคำนวณได้ง่ายกว่าเทคนิคอื่น ๆ และในปีเดียวกันนี้ เฟลด์ต์ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Feldt. 1989) ได้ขยายสูตรที่วิเคราะห์ความเชื่อมั่นตามแบบจำลองคะแนนจริงสัมพันธ์แบบมาตรฐานเดิมจากการแบ่งแบบทดสอบสองส่วนมาเป็นหลายๆ ส่วน (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Feldt and Brennan. 1989 : 115) โดยไม่ได้แสดงที่มาของสูตรไว้ให้

สรุปการประมาณค่าความเชื่อมั่นแบบ Congeneric เป็นแนวคิดใหม่ที่น่าสนใจแก้ปัญหาในกรณีที่ไม่สามารถแบ่งส่วนย่อยแบบทดสอบให้แต่ละตอนมีจำนวนข้อเท่ากัน และในกรณีที่แต่ละตอนมีจำนวนข้อเท่ากันแต่ไม่เป็นไปตามข้อตกลงของความคู่ขนานแบบคลาสสิก และแบบคะแนนจริง-สมมูล สูตรที่สามารถนำไปใช้ได้กว้างขวางได้แก่สูตรของ Raju และ Feldt

แบบทดสอบคะแนนจริงสัมพันธ์ : สัมประสิทธิ์ r_B (Congeneric Test : Coefficient r_B)

บุญเชิด ภิญโญนนตพงษ์ (2537 : 57-63) เสนอวิธีการใหม่สำหรับประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์จากแบบทดสอบ K ข้อที่ให้คะแนนในระบบ (0,1) โดยพัฒนามาจากสัมประสิทธิ์ r_{12} ให้สามารถประมาณค่าความเชื่อมั่นจากแบบทดสอบที่มีเงื่อนไขเฉพาะ ซึ่งมีความเป็นมาดังนี้ คือครอนบัค (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Cronbach. 1951) ได้นำเสนอสัมประสิทธิ์ α เป็นสูตรสำหรับคำนวณค่าความสอดคล้องภายใน (Internal Consistency) ของแบบทดสอบ ลงในบทความของท่านตั้งแต่ปี 1951 และได้แสดงให้เห็นว่าสูตรที่ 20 ของคูเดอร์ และริชาร์ดสัน (Kuder – Richardson) (ที่นิยามเรียนว่า KR-20) เป็นกรณีเฉพาะของสัมประสิทธิ์ α แต่ก็ไม่ได้ระบุข้อตกลง (Assumption) ของสูตรให้ชัดเจน

อย่างไรก็ตามทั้งสัมประสิทธิ์ α และ KR-20 ต่างก็ได้รับความนิยมาใช้ประมาณค่าความเชื่อมั่นของแบบทดสอบอย่างกว้างขวางจนถึงปัจจุบัน

ในช่วงเวลาที่ผ่านมาผู้สนใจหลายท่านได้แสดงที่มาของสูตรด้วยวิธีการต่าง ๆ กัน เช่น เฟลด์ต์ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Feldt. 1975), กัตแมน (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Gulliksen 1950), โนวิกและ เลวิส (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Novick and lewis 1967) และ (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Lord and Novick 1968) เป็นต้น โดยเฉพาะบทความของ โนวิก และเลวิส (Novick and lewis) ได้แสดงที่มาของสัมประสิทธิ์ α โดยใช้วิธีการทางคณิตศาสตร์มากกว่าคนอื่น ๆ และได้แสดงให้เห็นว่าสัมประสิทธิ์ α ต้องมีเงื่อนไขที่จำเป็นและเพียงพอ (Necessary and Sufficient) จึงจะทำให้สัมประสิทธิ์ α มีค่าเท่ากับค่าความเชื่อมั่นของแบบทดสอบ นั่นก็คือสูตรสัมประสิทธิ์ α กับสูตรความเชื่อมั่นในทางคณิตศาสตร์แล้วไม่ใช่สูตรเดียวกัน (Not Mathematical Identity) สัมประสิทธิ์ α จะเท่ากับค่าความเชื่อมั่นของแบบทดสอบที่ประกอบด้วยส่วนย่อยได้ ก็ต่อเมื่อส่วนย่อยของแบบทดสอบ หรือข้อสอบแต่ละข้อเป็นไปตามข้อตกลงที่ว่า คะแนนจริงของแต่ละส่วนมีความสัมพันธ์เป็นเส้นตรง ซึ่ง โนวิก และ เลวิส (บุญเชิด ภิญโญนนตพงษ์. 2537; อ้างอิงจาก Novick and lewis 1967) ได้ขยายความในรูปคณิตศาสตร์ว่าคะแนนจริงของแต่ละส่วนย่อย มีความสัมพันธ์เป็นเส้นตรง หมายความว่า $T_i = T_j + C_{ij}$ ทุกค่าของ i, j และ C_{ij} เป็นค่าคงที่ ขึ้นอยู่กับการจับคู่ของแต่ละส่วนย่อยซึ่ง Novick and Lewis. (1967) และ Lord and Novick (1968) เรียกข้อตกลงนี้ว่า Essentially Tau Equivalent นั่นก็คือสัมประสิทธิ์ α จะมีค่าเท่ากับค่าความเชื่อมั่นของแบบทดสอบก็ต่อเมื่อส่วนย่อยของแบบทดสอบแต่ละส่วนจำเป็นต้องมีคะแนนจริงสมมุติกัน(Essentially Tau Equivalent) นอกจากนั้นผลการวิจัยล่าสุดในปี 1993 Zimmerman, Zumbo และ Lalonde ได้ศึกษาค่าแนวโน้มเข้าส่วนกลางและการกระจายของสัมประสิทธิ์ α ที่คำนวณมาจากคะแนนซึ่งจำลองด้วยคอมพิวเตอร์ (Computer Simulation) ให้ละเมิดข้อตกลงของสูตรสัมประสิทธิ์ α สองข้อคือ

1. คะแนนที่ได้ไม่เป็นไปตามข้อตกลงของคะแนนจริงจำเป็นต้องสมมุติ และ
2. คะแนนคลาดเคลื่อน ไม่เป็นไปตามข้อตกลงที่ว่าต้องมีความอิสระกัน

ผลการวิจัยพบว่าเมื่อคะแนนเป็นไปตามข้อตกลง แล้วค่าเฉลี่ยของสัมประสิทธิ์ α จะมีค่าใกล้เคียงกับสัมประสิทธิ์ α ที่คำนวณจากประชากร แต่ทว่าค่าที่ได้มีการกระจายสูงมาก สัมประสิทธิ์ α จะมีค่าต่ำกว่าค่าความเชื่อมั่น (Underestimated Reliability) เมื่อละเมิดข้อตกลงเกี่ยวกับคะแนนจริงจำเป็นต้องสมมุติและสัมประสิทธิ์ α จะมีสูงกว่าค่าความเชื่อมั่น เมื่อละเมิดข้อตกลงเกี่ยวกับคะแนนคลาดเคลื่อนต้องอิสระจากกัน และทั้งสองกรณีความลำเอียงของการประมาณค่า มีความผันแปรขึ้นอย่างเป็นระบบตามระดับของการละเมิดข้อตกลง ข้อค้นพบนี้

สะท้อนให้เห็นในทางปฏิบัติว่าสัมประสิทธิ์ α ที่ปรากฏอยู่ในหนังสือคู่มือดำเนินการสอบ รายงานการวิจัย และที่อื่น ๆ บางครั้งอาจมีค่าสูงกว่า และบางครั้งอาจมีค่าต่ำกว่าสัมประสิทธิ์ ความเชื่อมั่นของประชากร

นักทฤษฎีการทดสอบได้ตระหนักถึงประเด็นดังกล่าวข้างต้นมานานแล้ว พร้อมทั้งหาทางผ่อนปรนข้อตกลงเกี่ยวกับคะแนนของแบบทดสอบจากข้อตกลงของคะแนนจริงสมมูล มาเป็นข้อตกลงของคะแนนจริงสัมพันธ์ (Congeneric) ซึ่งมีแบบจำลองคะแนน ดังนี้

$$T_i = \lambda_i T_i + C_i$$

นั่นก็คือ ตามข้อตกลงของคะแนนจริงจำเป็นต้องสมมูล คะแนนจริงของแต่ละส่วนคู่ใด ๆ จะมีความแตกต่างกันคงที่เท่ากับ C_i เท่านั้นแต่ตามข้อตกลงของคะแนนจริงสัมพันธ์ คะแนนจริง แต่ละส่วนของคู่ใด ๆ จะแตกต่างกันไปตามค่าคงที่ C_i และแตกต่างกันไปตามลักษณะเฉพาะ ของคะแนนแต่ละส่วนคู่ นั่นคือ λ_i นั่นเอง และสามารถประมาณค่าความเชื่อมั่นทั้งฉบับของ แบบทดสอบที่มีข้อตกลงแบบคะแนนจริงสัมพันธ์ โดยอาศัยสถิติขั้นสูง ซึ่ง Joreskog (1971) ใช้ เทคนิคความน่าจะเป็นสูงสุดและกำลังสองน้อยสุด (Least Square and Maximum Likelihood) แต่เป็นที่น่าเสียดายที่วิธีการดังกล่าวจำเป็นต้องคำนวณจากเครื่องคอมพิวเตอร์ อย่างไรก็ตาม มี นักวัดผลหลายท่าน อาทิ คริสทอฟ (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Kristof 1974) เฟลด์ต์ (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Feldt 1975) ราจู (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Raju 1977) กิลเมอร์ และ เฟลด์ต์ (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Gilmer and Feldt 1989) และ เลี้ยว (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Liou 1989) ต่างก็ได้พัฒนาสูตรสำหรับประมาณค่า ความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ไว้หลายสูตร แต่สูตรที่ให้ผลคงเส้นคงวามากที่สุด และมีการคำนวณที่ค่อนข้างง่ายมากสามารถคิดคำนวณจากเครื่องคิดคำนวณธรรมดา คือ สัมประสิทธิ์ r_{12} ซึ่ง เลี้ยว (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Liou 1989) ได้ ปรับปรุงสูตรมาจากสัมประสิทธิ์ β_K ของ ราจู (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Raju, 1977) และสูตรของ เฟลด์ต์ (บุญเชิด ภิญโญอนันตพงษ์, 2537; อ้างอิงจาก Feldt 1975) ให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยเป็น K ส่วน โดยไม่จำเป็นต้องทราบจำนวนข้อของแบบทดสอบ อย่างไรก็ตามแบบทดสอบส่วนใหญ่ที่ใช้กันอยู่ในชั้นเรียน มักจะเป็นแบบทดสอบปรนัยที่มีระบบการให้คะแนนเป็น (0,1) หรือให้คะแนนคำตอบที่ถูกเป็น 1 คะแนน และคำตอบที่ผิดเป็น 0 นั่นเอง ดังนั้นถ้าแบ่งแบบทดสอบเป็นส่วน ๆ ส่วนย่อยที่เล็กที่สุดก็คือข้อสอบแต่ละข้อ และแต่ละข้อจะมีคะแนนเป็น 0 หรือ 1 เท่านั้น ในกรณีเช่นนี้การจะใช้ สูตร r_{12} คำนวณค่าความเชื่อมั่นก็สามารถทำได้ แต่ต้องใช้เวลามากเกินไป ด้วยเหตุผล

ดังกล่าว จึงนำเสนอวิธีการใหม่ สำหรับประมาณค่าความเชื่อมั่น แบบคะแนนจริงสัมพันธ์จากแบบทดสอบ K ข้อที่ให้คะแนนในระบบ (0,1) โดยพัฒนามาจากสัมประสิทธิ์ r_{12} ให้สามารถประมาณค่าความเชื่อมั่นจากแบบทดสอบที่มีเงื่อนไขเฉพาะดังกล่าวข้างต้น และให้ชื่อว่าสัมประสิทธิ์ r_B เป็นสัมประสิทธิ์กรณีเฉพาะของสัมประสิทธิ์ r_{12} ซึ่งมีที่มาของของสัมประสิทธิ์ r_B ดังนี้

บุญเชิด ภิญโญนนตพงษ์ (2537 : 57-63) กำหนดให้แบบทดสอบประกอบด้วยข้อสอบ K ข้อ คะแนนสอบ (Observed Scores) คะแนนจริง และ คะแนนคลาดเคลื่อนของแต่ละข้อ และของทั้งฉบับแทนด้วยสัญลักษณ์ ดังนี้

คะแนนสอบแทนด้วย $X_1, X_2, X_3, \dots, X_K$ และ X

คะแนนจริงแทนด้วย $T_1, T_2, T_3, \dots, T_K$ และ T

คะแนนคลาดเคลื่อนแทนด้วย $E_1, E_2, E_3, \dots, E_K$ และ E ตามลำดับดังนี้

$$(1) \quad X = X_1 + X_2 + X_3 + \dots + X_K$$

$$(2) \quad T = T_1 + T_2 + T_3 + \dots + T_K$$

$$(3) \quad E = E_1 + E_2 + E_3 + \dots + E_K$$

จากแบบจำลองคะแนนของทฤษฎีการทดสอบแบบมาตรฐานเดิม(Classical Test Theory Model) ได้คะแนนแต่ละข้อดังนี้

$$(4) \quad X_i = T_i + E_i \quad i = 1 \text{ ถึง } K$$

และกำหนดข้อตกลงว่า แบบทดสอบประกอบด้วยข้อสอบที่วัดคุณลักษณะเดียวกัน และเป็นไปตามข้อตกลงแบบคะแนนจริงสัมพันธ์ กล่าวคือคะแนนจริงของแต่ละข้อมีความสัมพันธ์เป็นเส้นตรงกับคะแนนจริงของทั้งฉบับ

$$(5) \quad T_i = \lambda_i T + C_i \quad i = 1 \text{ ถึง } K$$

เมื่อ λ_i เป็นลักษณะประจำข้อของข้อที่ i ซึ่งกำหนดให้มีค่าเป็นจำนวนจริงบวก

C_i เป็นค่าคงที่ของข้อสอบ i

โดยกำหนดเงื่อนไขว่า $\sum \lambda_i = 1$ และ $\sum C_i = 0$ นำสมการ (5) แทนลงในสมการ (4) ได้

$$(6) \quad X_i = \lambda_i T + C_i + E_i \quad i = 1 \text{ ถึง } K$$

ถ้านำแบบทดสอบไปสอบกับกลุ่มตัวอย่างจำนวนหนึ่ง และหาค่าความแปรปรวนของข้อสอบแต่ละข้อ (σ_i^2) และหาค่าความแปรปรวนของข้อสอบแต่ละคู่ (σ_{ij}) ได้ดังนี้

$$(7) \quad \sigma_i^2 = \lambda_i^2 \sigma_T^2 + \sigma_{Ei}^2$$

$$(8) \quad \sigma_{ji} = \lambda_j \lambda_i \sigma_T^2 \quad (j \neq i)$$

ผลรวมของความแปรปรวนของคะแนนสอบทุกข้อคือ

$$(9) \quad \sum \sigma_i^2 = \sum \lambda_i^2 \sigma_T^2 + \sum \sigma_{Ei}^2$$

นำ σ_x^2 ไปหาสมการ (9) และนิยามความเชื่อมั่นของแบบทดสอบ $P_{xx'} = \sigma_T^2 / \sigma_x^2$ จะได้สูตรความเชื่อมั่น ซึ่ง Raju (1977) เป็นผู้เสนอไว้เป็นคนแรก ดังนี้

$$(10) \quad P_{xx'} = \frac{1}{1 - \sum \lambda_i^2} \left\{ 1 - \frac{\sum \sigma_i^2}{\lambda_x^2} \right\}$$

เลียว (Liou 1989) ได้อาศัยวิธีการของ เฟลด์ต์ (Feldt 1975) ที่จะประมาณค่า λ_i โดยพิจารณาว่าข้อสอบแต่ละข้อก็คือแบบทดสอบที่ลดจำนวนข้อลงจนเหลือข้อสอบเพียงข้อเดียว ดังนั้นข้อสอบก็คือแบบทดสอบฟอร์มหนึ่งนั่นเอง ซึ่งตามทฤษฎีการทดสอบมาตรฐานเดิม ความแปรปรวนของคะแนนคลาดเคลื่อนแต่ละข้อ และทั้งฉบับจะมีความสัมพันธ์กันดังนี้ $\sigma_{Ei}^2 = \lambda_i \sigma_E^2$ และถ้านำคะแนนสอบของแต่ละข้อมาสร้างเป็นเมทริกซ์ความแปรปรวน - ความแปรปรวนร่วมแล้ว จะสามารถหาค่า λ_i ของแต่ละข้อ จากแต่ละแถว (Row) ของเมทริกซ์ดังกล่าวได้โดยนำค่า σ_x^2 ไปหารผลรวมของความแปรปรวน - ความแปรปรวนร่วมในแต่ละแถว ดังนี้

จากสมการ(11) นำมาเขียนใหม่โดยแยกความแปรปรวนร่วมของคะแนนตัวมันเองออกมา ซึ่งก็คือค่าความแปรปรวนของข้อนั้น เช่น

$$(11) \quad \left\{ \begin{array}{l} \lambda_1 = \sum \sigma_{1i} / \sigma_x^2 \\ \lambda_2 = \sum \sigma_{2i} / \sigma_x^2 \\ \vdots \\ \lambda_K = \sum \sigma_{Ki} / \sigma_x^2 \end{array} \right.$$

$$(12) \quad \lambda = \frac{\sigma_1^2 + \sum \sigma_{li}}{\sigma_x^2} \quad (i \neq 1)$$

ถ้าสามารถหาค่าความแปรปรวนของคะแนนข้อ 1 กับข้ออื่น ๆ ทั้งหมดได้ ก็สามารถ หาค่า λ_i ได้การหาค่าความแปรปรวนร่วมของข้อสอบแต่ละคู่จากสูตรโดยตรง มีความยุ่งยากและใช้เวลามากวิธีที่ง่ายและรวดเร็วกว่า คืออาศัยค่านิยามของสัมประสิทธิ์สหสัมพันธ์ $P_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ ดังนั้นความแปรปรวนร่วมของข้อสอบ ข้อ 1 และข้อ 2 ได้จากการนำค่าส่วนเบี่ยงเบนมาตรฐานคูณค่าสัมประสิทธิ์สหสัมพันธ์คือ $\sigma_{12} = \sigma_1 \sigma_2 \sigma_{12}$ แต่เนื่องจากข้อสอบแต่ละข้อให้คะแนนในระบบ (0,1) ดังนั้น

$$(13) \quad \sigma_1 = \sqrt{P_1(1-P_1)} \text{ และ } \sigma_2 = \sqrt{P_2(1-P_2)}$$

เมื่อ P เป็นสัดส่วนของผู้ตอบถูก หรือค่าความยากง่ายของข้อสอบ สำหรับสัมประสิทธิ์สหสัมพันธ์จะต้องใช้สูตรเฉพาะกรณีเฉพาะของสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน คือ สัมประสิทธิ์ ϕ ซึ่งมีสูตรดังนี้

$$(14) \quad \phi = \frac{BC - AD}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

เมื่อ A, B, C และ D เป็นสัดส่วนจำนวนผู้ตอบข้อสอบสองข้อ ถูกหรือผิดดังแสดงในตาราง

		ข้อ 2		
		0	1	
ข้อ 1	1	A	B	A+B
	0	C	D	C+D
		A+C	B+D	1.0

จากตารางจะได้ว่า

$$A + B = P_1$$

$$B + D = P_2$$

$$C + D = 1 - P_1$$

$$A + C = 1 - P_2$$

และจัดค่าต่าง ๆ อยู่ในพจน์ของ P_1 , P_2 และ B จะได้

$$A = P_1 - B, D = P_2 - B, C = 1 - P_1 - P_2 + B$$

นำค่าดังกล่าวแทนในสมการ (14) ได้

$$\phi_{12} = \frac{B_{12}(1 - P_1 - P_2 + B_{12}) - (P_1 - B_{12})(P_2 - B_{12})}{\sqrt{P_1(1 - P_1)}(1 - P_2)P_2}$$

$$(15) \quad \phi_{12} = \frac{B_{12} - P_1P_2}{\sqrt{P_1(1 - P_1)}\sqrt{P_2(1 - P_2)}}$$

ดังนั้นจากสมการ (13) และ (15) ได้ความแปรปรวนร่วมดังนี้

$$\sigma_{12} = \sqrt{P_1(1 - P_1)}\sqrt{P_2(1 - P_2)} \frac{B_{12} - P_1P_2}{\sqrt{P_1(1 - P_1)}\sqrt{P_2(1 - P_2)}}$$

$$\sigma_{12} = B_{12} - P_1P_2$$

ทำนองเดียวกัน

$$\sigma_{13} = B_{13} - P_1P_3$$

$$\sigma_{14} = B_{14} - P_1P_4$$

หรือเขียนเป็นรูปทั่วไปได้ดังนี้

$$\sigma_{1i} = B_{1i} - P_1P_i$$

$$\sigma_{2i} = B_{2i} - P_2P_i$$

$\vdots$

$$\sigma_{Ki} = B_{Ki} - P_KP_i$$

ดังนั้นจากสมการ(11) และ (12) นำมาเขียนใหม่ได้

$$\begin{aligned}
 \lambda_i &= \frac{P_i(1-P_i) + \sum(B_{li} - P_i P_i)}{\sigma_x^2} & i \neq 1 \\
 &= \frac{P_i(1-P_i) + \sum(B_{li} - P_i \sum P_i)}{\sigma_x^2} \\
 &= \frac{P_i(1-P_i - \sum P_i) + \sum B_{li}}{\sigma_x^2} \\
 &= \frac{P_i(1-\bar{X}) + \sum B_{li}}{\sigma_x^2}
 \end{aligned}$$

$$\lambda_1 = \frac{\sum B_{li} - P_1(\bar{X}-1)}{\sigma_x^2} \quad i \neq 1$$

ทำนองเดียวกัน

$$\begin{aligned}
 \lambda_2 &= \frac{\sum B_{2i} - P_2(\bar{X}-1)}{\sigma_x^2} & i \neq 2 \\
 \vdots & \quad \quad \quad \vdots \\
 \lambda_K &= \frac{\sum B_{Ki} - P_K(\bar{X}-1)}{\sigma_x^2} & i \neq K
 \end{aligned}$$

ดังนั้นการประมาณค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์แบบมาตรฐานเดิม (Classical Congeneric) จากแบบทดสอบที่ให้คะแนนระบบ (0,1) จำนวน K ข้อ โดยใช้สัมประสิทธิ์ r_B ที่สร้างขึ้นใหม่มีสูตรดังนี้

$$(15) \quad r_B = \frac{1}{1 - \sum \lambda_i^2} \left\{ 1 - \frac{\sum P_i(1-P_i)}{S_x^2} \right\}$$

เมื่อ P_i คือ ค่าความยากง่ายของข้อสอบแต่ละข้อ
 S_x^2 คือ ความแปรปรวนของคะแนนทั้งฉบับ

$$\lambda_1 = [\Sigma B_{1i} - P_1(\bar{X} - 1)] / S_x^2$$

$$\lambda_2 = [\Sigma B_{2i} - P_2(\bar{X} - 1)] / S_x^2$$

$$\vdots \quad \quad \quad \vdots$$

$$\lambda_k = [\Sigma B_{ki} - P_k(\bar{X} - 1)] / S_x^2$$

$$\bar{X} = \Sigma P_i$$

B_{ji} คือ สัดส่วนจำนวนผู้ตอบข้อสอบสองข้อใด ๆ ถูกทั้งคู่

สรุปแบบทดสอบหรือเครื่องมือวัดสำหรับการวิจัยหรือการวัดและประเมินผลอาจจำแนกเป็น 3 ชนิด ตามข้อตกลงของระดับความคล้ายคลึงกันของแบบทดสอบ คือแบบทดสอบคะแนนจริงสัมพัทธ์ แบบทดสอบคะแนนจริงสมมูล และแบบทดสอบคู่ขนานซึ่งแบบทดสอบชนิดแรกมีข้อจำกัดของข้อตกลงน้อยที่สุด คือ แต่ละฟอร์มต้องวัดคุณลักษณะเดียวกัน หรือแต่ละคู่ต้องมีคะแนนจริงสัมพัทธ์กับเป็นเส้นตรงเท่านั้น เมื่อเทียบกับอีกสองชนิดแล้วแบบทดสอบคะแนนจริงสัมพัทธ์จะมีความผ่อนปรนของข้อตกลงมากที่สุดซึ่งทำให้สามารถนำไปใช้ในสถานการณ์ทดสอบต่าง ๆ ได้มากมาย โดยไม่มีการละเมิดข้อตกลง ดังนั้นการประมาณค่าความเชื่อมั่นของแบบทดสอบโดยใช้ข้อตกลงแบบคะแนนจริงสัมพัทธ์ จึงมีความปลอดภัยหรือมีความเสี่ยงน้อยที่สุดที่จะให้ผลคลาดเคลื่อน การวิจัยครั้งนี้ผู้วิจัยได้ใช้แบบทดสอบคะแนนจริงสัมพัทธ์ โดยหาค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างจากสูตรของเฟลด์ต์-ราซ

5. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องภายในประเทศ

งานวิจัยที่ศึกษาความเชื่อมั่นของแบบทดสอบ ได้แก่

สนั่น สิทธิวัง. (2512) ได้เปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบที่ครูสร้างตัวเองขึ้นเองกับข้อสอบที่ใช้ตัวเลือกจากคำตอบของนักเรียนในวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 4 ชั้นประถมศึกษาปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3 จำนวน 1,950 คน ผลการวิจัยพบว่า ความเที่ยง

ตรงของแบบทดสอบฉบับที่ครูได้ตัวเลือกจากนักเรียนสูงกว่าฉบับที่ครูคิดสร้างตัวเลือกขึ้นเอง อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ในชั้นประถมศึกษาปีที่ 4 และมีนัยสำคัญทางสถิติที่ระดับ .01 ในชั้นประถมศึกษาปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3 เมื่อพิจารณาค่าอำนาจจำแนกของแบบทดสอบ พบว่าฉบับที่ครูได้ตัวเลือกจากเด็กสูงกว่าฉบับที่ครูสร้างตัวเลือกขึ้นเองอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ค่าความยากมาตรฐานของแบบทดสอบเลือกตอบฉบับที่ครูได้รับตัวเลือกจากเด็ก ยากกว่าฉบับที่ครูคิดสร้างตัวเลือกขึ้นเอง

สุธรรม จันทร์หอม. (2513) ได้ศึกษาเปรียบเทียบการใช้แบบทดสอบถูกผิด เลือกตอบ และเติมคำ วัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ชั้นประถมศึกษาปีที่ 7 ในโรงเรียนสังกัดเทศบาลโรงเรียนราษฎร์ และสังกัดกรมสามัญศึกษาในจังหวัดเชียงใหม่ จำนวน 1,350 คน เครื่องมือที่ใช้ในการศึกษา คือแบบทดสอบวิชาคณิตศาสตร์ 3 ฉบับ คือ คณิตศาสตร์ทักษะ คณิตศาสตร์ปัญหา และคณิตศาสตร์ความเข้าใจ การเก็บรวบรวมข้อมูลให้นักเรียนสอบคนละ 3 ฉบับโดยให้นักเรียนแต่ละคนได้รับรูปแบบข้อสอบไม่ซ้ำกัน แบบทดสอบทั้ง 3 แบบวัดเนื้อหา เดียวกันและมีจำนวนข้อเท่ากัน ผลจากการวิจัย ได้ผลดังนี้

1. ค่าความเชื่อมั่นของแบบทดสอบแบบถูกผิด เลือกตอบ และเติมคำแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยแบบเติมคำมีค่าความเชื่อมั่นสูงสุด แบบเลือกตอบมีค่าปานกลาง และแบบถูกผิดมีค่าต่ำสุด

2. ค่าความเที่ยงตรงของแบบทดสอบแบบถูกผิด เลือกตอบ และเติมคำ ของวิชา คณิตศาสตร์ทักษะ และคณิตศาสตร์ปัญหาไม่แตกต่างกัน แต่วิชาคณิตศาสตร์ความเข้าใจ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ความยากของแบบทดสอบแบบถูกผิด เลือกตอบ และเติมคำ แตกต่างกันอย่างมีนัยสำคัญ ทางสถิติที่ระดับ .01 โดยแบบทดสอบถูกผิดง่ายที่สุด แบบเลือกตอบยากปานกลาง และแบบเติมคำยากที่สุด

4. ค่าอำนาจจำแนกของแบบทดสอบแบบถูกผิด เลือกตอบ และเติมคำของวิชา คณิตศาสตร์ทักษะไม่แตกต่างกัน แต่วิชาคณิตศาสตร์ปัญหาและคณิตศาสตร์ความเข้าใจ แตกต่างกัน โดยแบบทดสอบแบบเติมคำมีค่าอำนาจจำแนกสูงสุดแบบเลือกตอบมีค่าอำนาจจำแนกปานกลาง และแบบถูกผิดมีค่าอำนาจจำแนกต่ำสุด

ไพบุลย์ จิตรโต. (2514) ได้ศึกษาเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบ ชนิดต่างๆ และแบบเติมคำ ในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 และชั้นประถมศึกษาปีที่ 7 จำนวน 1,104 คน ผลการวิจัยพบว่าแบบทดสอบส่วนใหญ่มีค่าความเชื่อมั่น ความเที่ยงตรง และค่าอำนาจจำแนกเฉลี่ยพอ ๆ กันและเมื่อพิจารณาโดยทั่ว ๆ ไปแล้ว แบบทดสอบเติมคำมีค่าความเชื่อมั่น ความเที่ยงตรง ค่าอำนาจจำแนก และความยากมาตรฐานเฉลี่ยสูงกว่าแบบทดสอบเลือกตอบ และแบบทดสอบเลือกตอบให้ค่าคะแนนเฉลี่ยสูงกว่าแบบทดสอบเติมคำ

นภาพร อมรเลิศสินไทย. (2514) ได้ศึกษาผลการใช้แบบทดสอบเลือกตอบแบบตอบสั้น ๆ และแบบผสมระหว่างแบบเลือกตอบกับแบบตอบสั้น ๆ วัดผลสัมฤทธิ์ทางการเรียนวิชาสังคมศึกษา ชั้นประถมศึกษาปีที่ 7 จังหวัดพระนครศรีอยุธยา จำนวน 558 คน ผลการวิจัยพบว่าแบบเลือกตอบให้ค่าคะแนนเฉลี่ยสูงสุด และแบบตอบสั้น ๆ ยากที่สุด ค่าอำนาจจำแนกความเที่ยงตรง และความเชื่อมั่นของแบบทดสอบทั้ง 3 แบบไม่แตกต่างกัน แต่แบบตอบสั้น ๆ ให้ค่าความเชื่อมั่นสูงสุด ด้านการประเมินค่าแบบทดสอบเลือกตอบเหมาะกับนักเรียนทั้งกลุ่มที่มีผลสัมฤทธิ์ทางการเรียนสูงและต่ำ

สวัสดี ประทุมราชและคณะ. (2524) ได้ศึกษาผลการใช้แบบทดสอบเลือกตอบ และแบบเติมคำของคณะกรรมการสอบคัดเลือกนิสิตเข้าเรียนระดับบัณฑิตศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2523 ในวิชาภาษาอังกฤษ จำนวน 2,000 คน ผลการวิจัยพบว่า

1. แบบทดสอบเลือกตอบ จำนวน 50 ข้อ ให้ค่าความเชื่อมั่น .428
2. แบบทดสอบเติมคำ จำนวน 25 ข้อ ให้ค่าความเชื่อมั่น .777

กรรทอง เทพศิริอำนวย. (2525) ได้ศึกษาเปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบ และแบบเติมคำในวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 5 เขตกรุงเทพมหานคร จำนวน 420 คน ผลการวิจัยพบว่า

1. ในด้านค่าความยากของแบบทดสอบ แบบทดสอบเลือกตอบง่ายกว่าแบบทดสอบเติมคำ
2. ค่าอำนาจจำแนกของแบบทดสอบ แบบทดสอบเติมคำ มีค่าอำนาจจำแนกมากกว่าแบบทดสอบเลือกตอบ
3. ค่าความเชื่อมั่นของแบบทดสอบ แบบทดสอบเลือกตอบมีค่าความเชื่อมั่นน้อยกว่าแบบทดสอบเติมคำ
4. ค่าความเที่ยงตรงของแบบทดสอบ แบบทดสอบเลือกตอบมีค่าความเที่ยงตรงน้อยกว่าแบบทดสอบเติมคำ

จินดา โตอนันต์. (2527) ทำการศึกษาเปรียบเทียบคุณลักษณะของแบบทดสอบวัดผลสัมฤทธิ์แบบตอบถูกผิดหลายตัวเลือก กับแบบทดสอบเลือกตอบแบบตอบตัวเลือกเดียว กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในภาคเรียนที่ 2 ปีการศึกษา 2525 ของเขตกรุงเทพมหานคร จำนวน 432 คน เครื่องมือที่ใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์วิชาวิทยาศาสตร์ (ว102) ชั้นมัธยมศึกษาปีที่ 1 เรื่องสิ่งมีชีวิตและสิ่งแวดล้อม จำนวน 2 ฉบับ ที่มีเนื้อหาของ ข้อคำถามเหมือนกันข้อต่อข้อแต่มีรูปแบบในการตอบต่างกันดังนี้

แบบทดสอบฉบับ ก ประกอบด้วย แบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือก 24 ข้อ และแบบทดสอบเลือกตอบแบบตัวเลือกเดียว 24 ข้อ

แบบทดสอบฉบับ ข ประกอบด้วย แบบทดสอบเลือกตอบแบบตอบตัวเลือกเดียว 24 ข้อ และแบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือก 24 ข้อ

ผลการวิจัยพบว่า แบบทดสอบเลือกตอบตัวเลือกเดียวยากกว่าแบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือก และมีค่าอำนาจจำแนกสูงกว่าแบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือกทั้งในแบบทดสอบฉบับ ก. และแบบทดสอบฉบับ ข. ส่วนค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือก ฉบับ ก และฉบับ ข มีค่าเท่ากับ .42 และ .64 ตามลำดับ ในขณะที่ค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบเลือกตอบแบบตอบตัวเลือกเดียว ฉบับ ก และฉบับ ข มีค่าเท่ากับ .49 และ .38 ตามลำดับ

จักรกฤษณ์ สำราญใจ. (2531) ได้ศึกษาประสิทธิภาพของการใช้แบบทดสอบแบบตอบถูกผิดหลายตัวเลือกเปรียบเทียบกับแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียวกลุ่มตัวอย่างเป็นนักศึกษาวิทยาลัยครู 15 แห่ง จำนวน 1,672 คน ที่ลงทะเบียนเรียนวิชา ศึกษา 251 ในภาคเรียนที่ 2 ปีการศึกษา 2529 เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียว และแบบตอบถูกผิดหลายตัวเลือกจำนวน 63 ข้อ ที่มีข้อคำถามเหมือนกันข้อต่อข้อต่างกันตรงจำนวนตัวเลือกที่เป็นคำตอบถูกเท่านั้น ผลการวิจัยพบว่า ประสิทธิภาพของการใช้แบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือก จะสูงกว่าแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียว

วรรณพิศ นามเย็น. (2532) ได้ทำการศึกษาเปรียบเทียบค่าความยาก อำนาจจำแนก และค่าความเชื่อมั่น ของแบบทดสอบเลือกตอบวิชาคณิตศาสตร์ที่มีรูปแบบตัวเลือกต่างกัน 4 รูปแบบ คือ แบบธรรมดา แบบปลายเปิด แบบคำตอบย่อ และแบบค่าเปรียบเทียบ โดยศึกษากับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ปีการศึกษา 2531 ในจังหวัดปราจีนบุรีจำนวน 576 คน ผลการศึกษาพบว่า แบบทดสอบเลือกตอบที่มีตัวเลือกเป็นค่าเปรียบเทียบ มีค่าความยากมาตรฐานสูงสุด แบบตัวเลือกธรรมดามีค่าความยากมาตรฐานต่ำสุด เมื่อทดสอบความแตกต่างปรากฏว่าแบบทดสอบเลือกตอบตัวเลือกธรรมดากับแบบทดสอบเลือกค่าเปรียบเทียบมีความยากมาตรฐานแตกต่างกัน นอกจากนั้นมีความยากไม่แตกต่างกัน ในด้านค่าอำนาจจำแนกแบบทดสอบเลือกตอบตัวเลือกธรรมดามีค่าอำนาจจำแนกสูงสุด และแบบค่าเปรียบเทียบมีค่าอำนาจจำแนกต่ำสุด เมื่อทดสอบความแตกต่างปรากฏว่า ไม่มีความ แตกต่างกันทั้ง 4 รูปแบบ ส่วนในด้านค่าความเชื่อมั่น แบบทดสอบที่มีรูปแบบตัวเลือกธรรมดา มีค่าความเชื่อมั่นสูงสุด และแบบค่าเปรียบเทียบมีค่าความเชื่อมั่นต่ำสุด เมื่อทดสอบความ แตกต่างปรากฏว่าแบบทดสอบที่มีค่าความเชื่อมั่นแตกต่างกัน คือ แบบทดสอบที่มีตัวเลือกธรรมดาแบบค่าเปรียบเทียบ และแบบทดสอบที่มีตัวเลือกปลายเปิดกับค่าเปรียบเทียบนอกจากนั้น มีค่าความเชื่อมั่นไม่แตกต่างกัน

นพพร พรหมมา. (2536) ได้เปรียบเทียบคุณภาพของแบบทดสอบปรนัยที่มีรูปแบบตัวเลือกต่างกัน 3 รูปแบบ คือ แบบธรรมดา แบบปลายเปิดและแบบให้เติมคำตอบ

โดยเปรียบเทียบคุณภาพใน 3 ด้าน คือ ค่าความยาก ค่าอำนาจจำแนก และค่าความเที่ยง กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2535 จำนวน 90 คน จากโรงเรียนในกลุ่มโรงเรียนวังทอง สำนักงานการประถมศึกษาอำเภอวังทอง จังหวัดพิษณุโลก ผลการวิจัยพบว่า

1. แบบทดสอบปรนัยตัวเลือกแบบธรรมดา กับแบบปลายเปิด มีค่าความยากเฉลี่ยทั้งฉบับ ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับและค่าความเที่ยงทั้งฉบับไม่แตกต่างกัน
2. แบบทดสอบปรนัยตัวเลือกแบบธรรมดา กับแบบให้เติมคำตอบ มีค่าความยากเฉลี่ยทั้งฉบับ ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับและค่าความเที่ยงทั้งฉบับไม่แตกต่างกัน
3. แบบทดสอบปรนัยตัวเลือกแบบปลายเปิด กับแบบให้เติมคำตอบ มีค่าความยากเฉลี่ยทั้งฉบับ ค่าอำนาจจำแนกเฉลี่ยทั้งฉบับและค่าความเที่ยงทั้งฉบับไม่แตกต่างกัน

จุฑารัตน์ ดันเจริญ. (2538) ได้เปรียบเทียบคุณภาพของแบบทดสอบเรขาคณิต แบบเติมคำ แบบเลือกตอบซ้อน และแบบเลือกตอบถูกผิดซ้อน กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ปีการศึกษา 2536 ของโรงเรียนสังกัดกรมสามัญศึกษา ในท้องที่การศึกษา 2 กรุงเทพมหานคร จำนวน 960 คน ซึ่งเลือกมาโดยวิธีการสุ่มอย่างง่าย

จากการศึกษาพบว่า ความยากของแบบทดสอบเติมคำกับแบบเลือกตอบซ้อน, แบบทดสอบเลือกตอบซ้อนกับแบบเลือกตอบถูกผิดซ้อน มีค่าความยากมาตรฐานแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และแบบทดสอบเติมคำกับแบบเลือกตอบถูกผิดซ้อน มีค่าความยากมาตรฐาน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับค่าอำนาจจำแนก แบบทดสอบเติมคำกับแบบเลือกตอบซ้อนและแบบทดสอบเติมคำกับแบบเลือกตอบถูกผิดซ้อนมีค่าอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สำหรับแบบทดสอบเลือกตอบซ้อน กับแบบเลือกตอบถูกผิดซ้อน มีค่าอำนาจจำแนกไม่แตกต่างกัน ส่วนค่าความเชื่อมั่นของแบบทดสอบเติมคำกับแบบเลือกตอบซ้อน และแบบทดสอบเติมคำกับแบบเลือกตอบถูกผิดซ้อน แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 สำหรับแบบทดสอบเลือกตอบซ้อนกับแบบเลือกตอบถูกผิดซ้อนค่าความเชื่อมั่นไม่แตกต่างกัน

บังอร กมลวัฒน์. (2542) ได้เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2541 ของโรงเรียนในสังกัดสำนักงานคณะกรรมการการศึกษาเอกชน กรุงเทพมหานคร จำนวน 1,600 คนได้มาโดยวิธีการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลเป็นแบบทดสอบวัดความสามารถในการแก้ปัญหาตามกระบวนการแก้ปัญหาของเวียร์ ซึ่งผู้วิจัยสร้างขึ้นจำนวน 4 ฉบับ ดังนี้ ฉบับที่ 1 แบบทดสอบแบบเลือกตอบที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก ฉบับที่ 2 แบบทดสอบแบบเลือกตอบที่เรียงข้อปัญหาโดยวิธีสุ่ม ฉบับที่ 3 แบบทดสอบแบบตอบสั้นที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก ฉบับที่ 4 แบบทดสอบแบบตอบ

สั้นที่เรียงข้อปัญหาโดยวิธีสุ่ม ซึ่งในแต่ละฉบับมีสถานการณ์ปัญหาจำนวน 10 สถานการณ์ โดยแต่ละสถานการณ์มีข้อคำถามให้ตอบตามกระบวนการแก้ปัญหาจำนวน 4 ข้อ รวมแต่ละฉบับมีข้อคำถามทั้งหมด 40 ข้อ

ผลการวิจัยพบว่า ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาทั้ง 4 ฉบับ คือ แบบเลือกตอบที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก แบบเลือกตอบที่เรียงข้อปัญหาโดยวิธีการสุ่ม แบบตอบสั้นที่เรียงข้อปัญหาตามค่าความยากมาตรฐานเฉลี่ยจากง่ายไปยาก และแบบตอบสั้นที่เรียงข้อปัญหาโดยวิธีสุ่ม มีค่าเท่ากับ .788 .793 .804 และ .816 ตามลำดับ เมื่อเปรียบเทียบความแตกต่างของค่าความเชื่อมั่นแล้วพบว่า

1. ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีการจัดเรียงปัญหาต่างกันเมื่อมีรูปแบบการตอบเหมือนกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบต่างกันเมื่อมีการจัดเรียงปัญหาเหมือนกัน มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

สมพงษ์ อิศวฤกษ์. (2542) ได้เปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจและความยาวของแบบทดสอบต่างกัน ด้วยแบบทดสอบเลือกตอบถูกผิด วิชาสถิติ ธุรกิจ 1 จำนวน 2 ฉบับ และนำมาตรวจให้คะแนนด้วยวิธี 0-1, L-5, L-33, L-25, และ L-2 ใน การศึกษาครั้งนี้กลุ่มตัวอย่างเป็นนักศึกษาประกาศนียบัตรวิชาชีพชั้นสูง ปีการศึกษา 2541 สถาบันเทคโนโลยีราชมงคล แผนกบริหารธุรกิจ (พาณิชยกรรม) ส่วนกลาง จำนวน 5 วิทยาเขต โดยสุ่มตัวอย่างแบบแบ่งชั้น (Stratified Random Sampling) มีวิทยาเขตเป็น Strata และห้องเรียนเป็นหน่วยการสุ่มด้วยความเชื่อมั่น 95 % จำนวน 728 คน แบ่งเป็นสองกลุ่ม กลุ่มละ 364 คน เพื่อทำการทดสอบแบบทดสอบเลือกตอบถูกผิด 2 ฉบับ

ผลการวิจัยพบว่า

1. ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีจำนวนข้อต่างกัน และตรวจให้คะแนนวิธีเดียวกันด้วยสูตรสัมประสิทธิ์แอลฟาและสูตรของราซ จะมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติระดับ .01

2. ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีวิธีการตรวจให้คะแนนต่างกัน และจำนวนข้อเท่ากันด้วยสูตรสัมประสิทธิ์แอลฟาและสูตรของราซ จะมีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อรอนงค์ บำรุง (2542) เพื่อแสดงหลักฐานความเที่ยงตรง และความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ โดยมีจุดมุ่งหมายเพื่อ 1) แสดงหลักฐานความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ 2 วิธี คือหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait – Multimethod) และหาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบ ที่คำนวณด้วยสูตรของ ไฮเซนและบอร์นสเตดท์(Heise and Bohmstedt) และ 2) หาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ 2 วิธี คือ วิเคราะห์

นำนักองค์ประกอบ ที่คำนวณด้วยสูตรของอัลเลน(Allen) และวิเคราะห์แบบคะแนนจริงสัมพันธ์ ที่คำนวณด้วยสูตรของเลียว(Liou) กลุ่มตัวอย่างที่ใช้ในการศึกษาค้างนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 ของโรงเรียนสังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร จำนวน 300 คน โดยใช้การสุ่มแบบ 2 ขั้นตอน

ผลการวิจัยพบว่า

1. หลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่หาด้วยวิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait – Multimethod) พบว่าสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .4862 ถึง .7316 และสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงจำแนก มีค่าตั้งแต่ .1508 ถึง .4753

2. หลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของ ไฮเซนและบอร์นสเตดท์(Heise and Bohrnstedt) มีค่าเท่ากับ .8993

3. ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของอัลเลน(Allen) มีค่าเท่ากับ .8755

4. ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของเลียว(Liou) มีค่าเท่ากับ .6580

อัมพร วิชัยศรี. (2541) การศึกษาค้างนี้มีจุดมุ่งหมายเพื่อ 1.แสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบที่สร้างขึ้นตามแนวโครงสร้างของแบบทดสอบความสามารถทางสมองของโอติส - -ลนนอน ฟอรั่ม เจ ระดับสูง ที่มีรูปแบบการจัดเรียงแตกต่างกัน 3 ฟอรั่ม ได้แก่ ฟอรั่ม 1 เรียงจากง่ายไปหายาก ฟอรั่ม 2 เรียงเนื้อหาเป็นด้านๆ และฟอรั่ม 3 แยกเป็นฉบับย่อย ๆ ผู้วิจัยได้ทำการวิเคราะห์โดยใช้วิธีหาความเที่ยงตรงเชิงโครงสร้าง 3 วิธีคือวิธีการวิเคราะห์หาความเที่ยงตรงโดยใช้สูตรของไฮเซนและบอร์นสเตดท์วิธีหาความสัมพันธ์ระหว่างแบบทดสอบที่สร้างขึ้นกับแบบทดสอบความสามารถทางสมองของโอติส-ลนนอน ฉบับแปล โดยหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation Coefficient) และวิธีการเปรียบเทียบคะแนนเฉลี่ยของนักเรียนจำแนกตามระดับชั้น และจำแนกตามระดับอายุ โดยใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว(One-way analysis of variance) 2. ศึกษาค่าความเชื่อมั่นของแบบทดสอบ 3 ฟอรั่มโดยคำนวณหาค่าความเชื่อมั่นด้วยสูตร Feldt-Raju และสูตร Ω_w 3. เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบ 3 ฟอรั่ม ที่แบ่งส่วนย่อยไม่เท่ากัน 4 ส่วน ที่คำนวณด้วยสูตรเดียวกัน ด้วยสูตร Ux_1

กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3-6 ในระดับอายุ 13 – 19 ปี ปีการศึกษา 2540 จาก 17 โรงเรียน จำนวน 5,400 คน ในเขตกรุงเทพมหานคร จากการสุ่มแบบหลายขั้นตอน โดยแบ่งเป็น 3 กลุ่มกลุ่มๆ ละ 1,800 คน

ผลการวิจัยสรุปได้ดังนี้

1. แบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 ที่ผู้วิจัยสร้างขึ้น มีความเที่ยงตรงเชิงโครงสร้างเท่ากับ .8122, .9249 และ .8388 ซึ่งอยู่ในเกณฑ์สูง คะแนนของแบบทดสอบ 3 ฟอร์มที่สร้างขึ้นมีความสัมพันธ์ทางบวกกับคะแนนของแบบทดสอบความสามารถทางสมองของ โอติส-เลนนอนฟอร์ม เจ ระดับสูง ฉบับแปลอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นอกจากนี้ยังพบว่าคะแนนเฉลี่ยของแบบทดสอบฟอร์ม 1, ฟอร์ม 2, และฟอร์ม 3 แต่ละฟอร์มที่จำแนกตามระดับชั้นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 คะแนนเฉลี่ยของแบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 แต่ละฟอร์มที่จำแนกตามระดับอายุแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อทำการทดสอบความแตกต่างรายคู่ด้วยการเปรียบเทียบพหุคูณด้วยวิธีของเชฟเฟ้(Scheffe' s Test)ปรากฏว่านักเรียนในระดับชั้นสูงกว้างมีคะแนนเฉลี่ยของแบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 สูงกว่านักเรียนที่อยู่ชั้นต่ำกว่า อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และนักเรียนที่มีระดับอายุมากกว่ามีคะแนนเฉลี่ยของแบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 สูงกว่านักเรียนที่มีระดับอายุน้อยกว่า ในระดับอายุที่ห่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จึงสรุปได้ว่าแบบทดสอบทั้ง 3 ฟอร์ม มีความเที่ยงตรงเชิงโครงสร้างในระดับสูง

2. แบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 มีค่าความเชื่อมั่นที่คำนวณด้วยสูตร r_{F-R} เท่ากับ 0.9336, 0.9174 และ .9368 ตามลำดับ และแบบทดสอบฟอร์ม 1, ฟอร์ม 2 และฟอร์ม 3 มีค่าความเชื่อมั่นเมื่อคำนวณด้วยสูตร Ω_w เท่ากับ 0.9034, 0.8525 และ 0.8338 ตามลำดับ ซึ่งทั้ง 3 ฟอร์มมีค่าความเชื่อมั่นสูงกว่า 0.7

3. ค่าความแตกต่างของค่าความเชื่อมั่น ที่คำนวณด้วยสูตรเดียวกันโดยใช้สูตร UX1พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และเมื่อทำการทดสอบรายคู่ ที่คำนวณด้วยสูตร Feldt-Raju ทั้ง 3 ฟอร์ม พบว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และทดสอบรายคู่ที่คำนวณด้วยสูตร Ω_w ทั้ง 3 ฟอร์มพบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ธีระศักดิ์ กองทรัพย์.(2543) การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายเพื่อ 1. การแสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบสำรวจพหุปัญญาสำหรับนักเรียนระดับประถมศึกษา ที่ผู้วิจัยสร้างขึ้นเป็นแบบมาตรวัดประเมินค่า ผู้วิจัยได้ทำการวิเคราะห์หาความเที่ยงตรงเชิงโครงสร้าง 3 วิธี คือวิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี(Multitrait-Multimethod) วิธีหาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบที่คำนวณด้วยสูตรของไฮสและบอร์นสเตดท์ และหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน(Pearson Product Moment Correlation Coefficient) 2. ศึกษาความเชื่อมั่นของแบบสำรวจพหุปัญญา โดยคำนวณด้วยสูตร r_{F-R} และสูตร Ω_w

กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นประถมศึกษาปีที่ 4 ปีการศึกษา 2542 ของโรงเรียนสังกัดเขตภาษีเจริญ กรุงเทพมหานคร จำนวน 646 คน ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน ผลการวิจัยพบว่า

1. หลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบสำรวจพหุปัญญา ที่หาด้วยวิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait-Multimethod) พบว่า สัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงเหมือน มีค่าอยู่ระหว่าง 0.4744 ถึง 0.6902 สัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงจำแนก มีค่าอยู่ระหว่าง 0.0170 ถึง 0.5743

2. ความเที่ยงตรงเชิงโครงสร้างของแบบสำรวจพหุปัญญาฉบับที่ 1 และฉบับที่ 2 ซึ่งหาด้วยสูตรของไฮสและบอร์นสเตดท์(Heise and Bohmstedt) มีค่าเท่ากับ .08038 และ 0.6044 ตามลำดับ

3. ความเที่ยงตรงเชิงโครงสร้างของแบบสำรวจพหุปัญญาทั้ง 2 ฉบับ มีค่าสัมประสิทธิ์สหสัมพันธ์เป็น .06761 มีความสัมพันธ์ทางบวกอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4. หาค่าความเชื่อมั่นของแบบสำรวจพหุปัญญาฉบับที่ 1 แบบตรวจสอบรายการ (Checklist) และฉบับที่ 2 มาตราวัดประเมินค่า(Rating Scale) ที่คำนวณด้วยสูตร Ω_w มีค่าเท่ากับ 0.8900 และ 0.8982 ตามลำดับ ซึ่งทั้ง 2 ฉบับมีค่าความเชื่อมั่นสูงกว่า 0.70

5. ค่าความเชื่อมั่นของแบบสำรวจพหุปัญญาฉบับที่ 1 แบบตรวจสอบรายการ (Checklist) และฉบับที่ 2 แบบมาตราวัดประเมินค่า (Rating Scale) ที่คำนวณด้วยสูตร r_{F-R} มีค่าเท่ากับ 0.7907 และ 0.8614 ตามลำดับ ซึ่งทั้ง 2 ฉบับมีค่าความเชื่อมั่นสูงกว่า 0.70

ธิดาเพ็ญ จรัสศรี. (2543) การวิจัยครั้งนี้มีจุดหมาย ดังนี้) เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบความคิดรวบยอด ที่ทดสอบเด็กเป็นกลุ่ม 2) เพื่อแสดงหลักฐานความเชื่อมั่นของแบบทดสอบความคิดรวบยอด ที่ทดสอบเป็นกลุ่ม โดย คำนวณด้วยสูตร r_B และสูตร KR-20

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นอนุบาล 1-3 ระดับอายุ 4 -6 ปี ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนในสังกัดสังฆมณฑลราชบุรี จังหวัดกาญจนบุรี จำนวน 104 คน ได้มาจากการสุ่มแบบแบ่งชั้น

ผลการวิจัยพบว่า

1. การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้าง โดยใช้หลักพิจารณาตามวิธีการแบบหลายลักษณะ-หลายวิธี (Multitrait-Multimethod) ของแบบทดสอบความคิดรวบยอด มีค่าสัมประสิทธิ์ของความเที่ยงตรงเชิงเหมือน (Convergent Validity) ซึ่งเป็นการวัดในลักษณะเดียวกันด้วยแบบทดสอบต่างชนิดกัน มีค่าตั้งแต่ .7540 ถึง .8539 ซึ่งในแต่ละฉบับมีค่าสัมประสิทธิ์สหสัมพันธ์ สูงกว่า ค่าสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ซึ่งเป็นการวัดลักษณะต่างกันด้วยแบบทดสอบต่างชนิดกัน

2. ความเที่ยงตรงเชิงโครงสร้าง ที่ได้จากค่าสหสัมพันธ์แบบเพียร์สัน ระหว่างแบบทดสอบความคิดรวบยอดที่สร้างขึ้น กับแบบคัดแยกเด็กเป็นหมูน 1 มีค่า .9354

3. ค่าความเชื่อมั่นของแบบทดสอบความคิดรวบยอดจากการหาสัมประสิทธิ์ r_B ของแบบทดสอบย่อยแต่ละฉบับ มีค่าตั้งแต่ .6197 ถึง .9612

4. ค่าความเชื่อมั่นของแบบทดสอบความคิดรวบยอดจากการคำนวณด้วยสูตร KR-20 ของแบบทดสอบย่อยแต่ละฉบับ มีค่าตั้งแต่ .5921 ถึง .9599

สนอง ตรงเที่ยง (2540) การศึกษาครั้งนี้มีจุดมุ่งหมายสำคัญเพื่อตรวจสอบความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบวัดการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปและแบบทดสอบวัดการประเมินค่าทางภาษาแบบการแปลงรูปตามทฤษฎีโครงสร้างทางสถิติปัญญาของกิลฟอร์ด โดยใช้วิธีวิเคราะห์องค์ประกอบเชิงสำรวจ และวิธีวิเคราะห์ยืนยันองค์ประกอบด้วยวิธีความน่าจะเป็นสูงสุดตามวิธีของโจเรสกอก และซอร์บอม(Joreskog and Sorbom. 1989) และคำนวณค่าความเชื่อมั่นด้วยสูตรเฟลด์ต-ราจู(Feldt-Raju)

กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2539 ของโรงเรียนสังกัดกรมสามัญศึกษาในจังหวัดนนทบุรี 10 โรงเรียน จำนวน 610 คน จากการสุ่มแบบหลายขั้นตอน โดยแบ่งเป็น 2 กลุ่ม คือ กลุ่มที่ 1 ใช้วิเคราะห์องค์ประกอบเชิงสำรวจจำนวน 310 คน กลุ่มที่ 2 ใช้การวิเคราะห์องค์ประกอบเชิงยืนยันจำนวน 300 คน

จากการศึกษาความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบที่ได้จากการวิเคราะห์องค์ประกอบเชิงสำรวจ พบว่า แบบทดสอบฉบับการถอดรหัส ฉบับการสลับที่ตัวอักษร ฉบับการตัดสินใจถูกต้องของนิพจน์ วัดองค์ประกอบร่วมกันหนึ่งองค์ประกอบ คือองค์ประกอบ การประเมินค่าทางสัญลักษณ์แบบการแปลงรูป และแบบทดสอบฉบับการตัดสินใจการใช้ประโยชน์ของสิ่งของ ฉบับการเปรียบเทียบประโยคที่ใช้บรรยายภาพ วัดองค์ประกอบร่วมกันหนึ่งองค์ประกอบคือองค์ประกอบ การประเมินค่าทางภาษาแบบการแปลงรูป

ความเที่ยงตรงเชิงโครงสร้างที่ได้จากการวิเคราะห์ยืนยันองค์ประกอบด้วยวิธีความน่าจะเป็นสูงสุดพบว่า แบบทดสอบฉบับการถอดรหัส ฉบับการสลับที่ตัวอักษร ฉบับการตัดสินใจถูกต้องของนิพจน์ วัดองค์ประกอบด้านการประเมินค่าทางสัญลักษณ์แบบการแปลงรูป อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และแบบทดสอบฉบับตัดสินใจการใช้ประโยชน์ของสิ่งของ และฉบับการเปรียบเทียบประโยคที่ใช้บรรยายภาพ วัดองค์ประกอบด้านการประเมินค่าทางภาษาแบบการแปลงรูปอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 แบบทดสอบการประเมินค่าทางสัญลักษณ์แบบการแปลงรูปจำนวน 3 ฉบับมีค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์คือ .740, .749 และ .763 ตามลำดับ แบบทดสอบการประเมินค่าทางภาษาแบบการแปลงรูป จำนวน 2 ฉบับ มีค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์คือ .717 และ .742 ตามลำดับ

งานวิจัยที่เกี่ยวข้องในต่างประเทศ

โดนัลด์ และคนอื่นๆ (Donald and other n.d. : 179 – 182) ได้ศึกษาเปรียบเทียบการกะประมาณเกี่ยวกับความเชื่อมั่นของแบบทดสอบเติมคำ แบบทดสอบจับคู่ และแบบทดสอบเลือกตอบ ทดสอบกับกลุ่มตัวอย่างนักศึกษาระดับปริญญาตรีที่เรียนวิชาจิตวิทยาเบื้องต้น มหาวิทยาลัยคาลิฟอร์เนีย จำนวน 73 คน โดยการสร้างแบบทดสอบเติมคำ แบบทดสอบจับคู่ และแบบทดสอบเลือกตอบ แบบละ 2 ฉบับ (A และ B) รวมเป็นข้อสอบ 6 ฉบับ ผลการวิจัยพบว่า แบบทดสอบเติมคำมีความเชื่อมั่นสูงกว่าแบบทดสอบจับคู่ และแบบทดสอบจับคู่มีความเชื่อมั่นสูงกว่าแบบทดสอบเลือกตอบ

เบนเดอร์ และเดวิส (Bender and Davis. 1949 : 49-60) ได้สำรวจความคิดเห็นของนักเรียนมัธยมเกี่ยวกับข้อสอบที่ครูสร้างขึ้นใช้เองเกี่ยวกับคำถามที่ใช้ในแบบทดสอบนั้น พบว่า นักเรียนชอบคำถามที่วัดความสามารถในการแก้ไขปัญหา 70 % และชอบคำถามที่เกี่ยวกับข้อเท็จจริงและรูปภาพ 30 % นอกจากนี้เขายังกล่าวไว้อีกว่า นักเรียนมีความเชื่อว่าข้อสอบแบบถูกผิด และแบบเลือกตอบนั้นเปิดโอกาสให้เดาถูกมาก

เกี่ยวกับคำถามที่ว่า ถ้าหากไม่เปิดโอกาสให้นักเรียนมีเวลาเตรียมตัวก่อนที่จะสอบ แล้วนักเรียนจะทำแบบทดสอบชนิดใดนั้น พบว่า 49 % ชอบทำข้อสอบเลือกตอบ 23% ชอบทำข้อสอบถูกผิด 15 % ชอบทำข้อสอบจับคู่ 7 % ชอบทำข้อสอบอัตนัย 4 % ชอบทำข้อสอบเติมคำ และ 2% ชอบทำข้อสอบชนิดปัญหา

เกี่ยวกับคำถามที่ว่า แบบทดสอบชนิดใดที่จะทำให้นักเรียนแสดงความรู้ของตนเองได้เป็นอย่างดี โดยที่นักเรียนมีเวลาในการเตรียมตัวสอบในวิชานั้น พบว่า 38% เป็นแบบทดสอบอัตนัย 29% เป็นแบบทดสอบเติมคำ 15% เป็นแบบทดสอบชนิดปัญหา 9% เป็นแบบทดสอบเลือกตอบ 6% เป็นแบบทดสอบถูกผิด 5% เป็นแบบทดสอบจับคู่

เกี่ยวกับคำถามที่ว่าแบบทดสอบชนิดใดที่ทำให้นักเรียนต้องดูหนังสือมากที่สุด นั้น พบว่าแบบทดสอบอัตนัยมาเป็นอันดับที่ 1 และแบบเติมคำมาเป็นอันดับที่ 2 จากการวิเคราะห์ข้อมูล พบว่า 43% เป็นแบบทดสอบอัตนัย 27% เป็นแบบทดสอบเติมคำ

เกี่ยวกับความเห็นของนักเรียนที่ว่า แบบทดสอบชนิดใดที่นักเรียนเห็นว่าคุ้มค่ากับเวลาที่เขาต้องเร่งท่องจำ (Cramming) เพื่อที่จะสอบนั้น พบว่า ในแบบทดสอบถูกผิดนักเรียนเห็นว่าคุ้มค่าเพียง 35% แบบทดสอบเลือกตอบคุ้มค่า 35% แบบทดสอบเติมคำคุ้มค่า 63% แบบทดสอบชนิดปัญหาคุ้มค่า 65% แบบทดสอบจับคู่คุ้มค่า 40% และแบบทดสอบอัตนัยคุ้มค่า 60%

มเยอร์ (Mayer. 1955 : 128) ได้ศึกษาผลการใช้แบบทดสอบอัตนัยแบบถูกผิดแบบเลือกตอบ และแบบเติมคำโดยการศึกษาเกี่ยวกับเด็กนักเรียน 124 คน แบ่งเด็กออกเป็น 4 กลุ่ม

ให้สอบแบบทดสอบ ดังนี้ กลุ่มที่ 1 สอบแบบทดสอบอัตนัย กลุ่มที่ 2 สอบแบบทดสอบถูกผิด กลุ่มที่ 3 สอบแบบทดสอบเลือกตอบ กลุ่มที่ 4 สอบแบบทดสอบเติมคำ

ผลการวิจัยพบว่า เด็กมีความสนใจในการทำแบบทดสอบทั้ง 4 แบบเท่า ๆ กัน ต่อมา อีก 5 สัปดาห์ นำแบบทดสอบชุดเดิมไปทดสอบกับเด็กกลุ่มเดิม ผลการวิจัยพบว่า แบบทดสอบ อัตนัยและเติมคำได้คะแนนสูงกว่าครั้งแรก แบบทดสอบเลือกตอบและถูกผิดไม่ได้คะแนนสูงขึ้น นั่นคือ แบบทดสอบเลือกตอบและถูกผิดมีความเชื่อมั่นมากกว่าแบบทดสอบอัตนัยเติมคำ แต่ผลการวิจัยครั้งนี้ยังเชื่อไม่ได้ทั้งนี้ขึ้นอยู่กับความคุ้นเคยที่เด็กมีอยู่กับการสอบแบบใดแบบหนึ่งมาก่อนแล้ว และแบบของการคิดสำหรับการตอบคำถามแต่ละแบบด้วย

ฟรีสบี และสวีเนีย (Frisbie and Sweeney. 1982 : 29 – 34) ได้ทำการศึกษา เปรียบเทียบแบบทดสอบวัดผลสัมฤทธิ์แบบเลือกตอบถูกผิดหลายตัวเลือกกับแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียว กลุ่มตัวอย่างเป็นนักศึกษาระดับอุดมศึกษาที่เรียนวิชา ชีววิทยา (ว410) จำนวน 574 คน ในภาคต้นของปีการศึกษา 1980 เครื่องมือที่ใช้เป็นแบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือกและแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียว โดยตรวจให้คะแนนในแต่ละข้อเป็นรายตัวเลือก คือถ้าผู้สอบตอบได้ถูกต้องตรงตามสภาพที่เป็นจริง ทั้งตัวเลือกถูกและตัวเลือกผิด จะได้ตัวเลือกละ 1 คะแนน แต่ถ้าตอบตรงกันข้ามกับสภาพที่เป็นจริงได้ 0 คะแนนผลการวิจัยพบว่า แบบทดสอบเลือกตอบแบบตอบถูกผิดหลายตัวเลือกมีค่าความเชื่อมั่นและค่าความเที่ยงตรงสูงกว่าแบบทดสอบเลือกตอบแบบตอบถูกตัวเดียว

ออสเตอร์ฮอฟ และพามেলা (Oosterhof and Pamela. 1984 287 – 293) ได้ศึกษา เปรียบเทียบคุณภาพของแบบทดสอบเลือกตอบที่มีตัวเลือกแตกต่างกัน 3 ชุด กับแบบทดสอบเติมคำ (Completion) 1 ชุดในวิชาคณิตศาสตร์ที่โจทย์ปัญหา ในการวิจัยครั้งนี้แบบทดสอบมีข้อสอบชุดละ 12 ข้อ โดยศึกษากับกลุ่มตัวอย่างที่เป็นนักศึกษาธุรกิจศึกษา จำนวน 232 คน ผลการวิจัยพบว่า ค่าความยากของแบบทดสอบเลือกตอบทั้ง 3 ชุด แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ส่วนแบบทดสอบเติมคำยากกว่าแบบทดสอบเลือกตอบทั้ง 3 ฉบับ อย่างมีนัยสำคัญทางสถิติที่ระดับ .02 เมื่อพิจารณาถึงค่าความเชื่อมั่นของแบบทดสอบทั้ง 4 ชุดที่คำนวณโดยสูตร KR – 20 พบว่า แบบทดสอบเติมคำมีค่าความเชื่อมั่นสูงสุดที่ส่วนแบบทดสอบเลือกตอบทั้ง 3 ฉบับมีค่าความเชื่อมั่นใกล้เคียงกัน

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่กำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2543 ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 116 โรงเรียน 1,290 ห้องเรียน จำนวนนักเรียนทั้งหมด 53,774 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่เลือกเรียนวิชางานช่างพื้นฐาน และกำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2543 ในจังหวัดกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวนนักเรียน 1,200 คน จากจำนวนห้องเรียน 30 ห้อง โรงเรียนทั้งหมด 12 โรงเรียน ซึ่งเลือกมาโดยการสุ่ม 2 ขั้นตอน (Random Sampling) โดยดำเนินการตามลำดับขั้นดังนี้

ประมาณกลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าด้วยความเชื่อมั่น 95 เปอร์เซนต์ ($\alpha = .05$) เมื่อเทียบจากตารางขนาดกลุ่มตัวอย่างที่ควรเลือกจากประชากรแล้ว ต้องสุ่มตัวอย่างจำนวน 398 คน (ล้วน สายยศ และอังคณา สายยศ 2539 : 297 ; อ้างอิงมาจาก Yamane, 1967) ในการศึกษาครั้งนี้แบบทดสอบที่ใช้ในการวิจัยมีจำนวน 3 ฉบับ ดังนั้นจึงต้องใช้กลุ่มตัวอย่างจำนวน 1194 คน แบ่งออกเป็น 3 กลุ่ม แต่เพื่อความสะดวกในการวิเคราะห์ข้อมูล ผู้วิจัยจึงใช้กลุ่มตัวอย่าง 1200 คน โดยสุ่มมาจาก โรงเรียน ดังนี้

ขั้นที่ 1 คัดเลือกโรงเรียนที่เปิดสอนวิชาเลือกเครื่องช่างพื้นฐาน จำนวน 74 โรงเรียน
จำนวน 188 ห้องเรียน

ขั้นที่ 2 สุ่มห้องเรียนจากโรงเรียนในขั้นที่ 1 ได้ห้องเรียน จำนวน 30 ห้องเรียน
จำนวนนักเรียน 1,200 คน ดังรายละเอียดในตาราง

ตาราง 3 จำนวนกลุ่มตัวอย่างที่ใช้ในการวิจัย

ลำดับที่	ชื่อโรงเรียน	จำนวนห้องเรียน	จำนวนนักเรียน
1	ปทุมคงคา	3	126
2	ทวีธาภิเศก	3	117
3	เทพศิรินทร์	3	123
4	สวนกุหลาบวิทยาลัย	3	127
5	วัดสุทธิวราราม	3	119
6	วัดประดู่ในทรงธรรม	3	114
7	วัดราชบพิธ	2	79
8	โพธิสารพิทยากร	2	83
9	สุวรรณารามวิทยาคม	2	76
10	จันทร์ประดิษฐาราม	2	73
11	มัธยมวัดหนองแขม	2	82
12	ราชวินิตบางแคปานขำ	2	81
	รวม	30	1,200

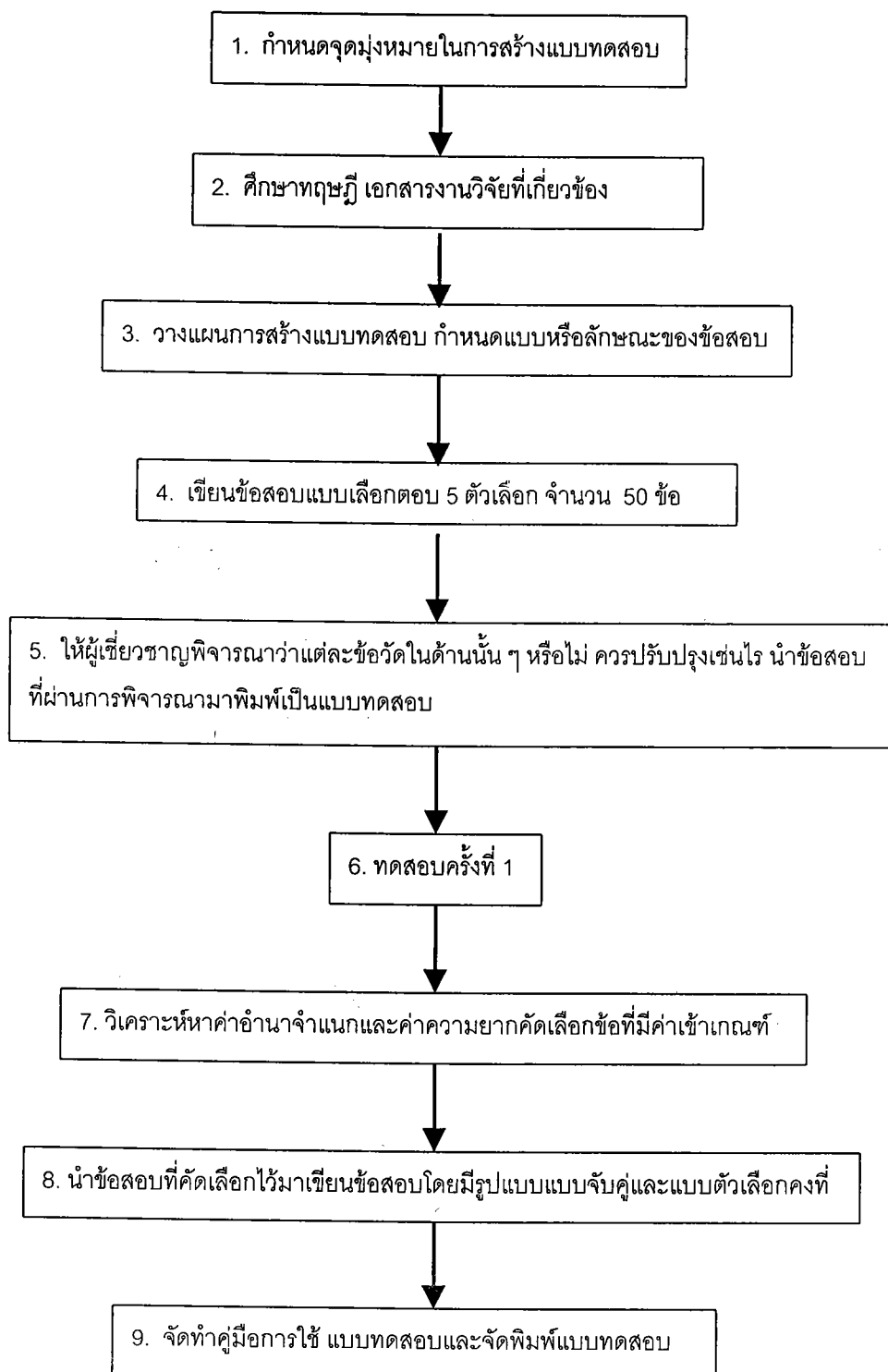
เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบ แบบทดสอบทั้งหมด 3 ฉบับ ดังนี้

1. แบบเลือกตอบ จำนวน 30 ข้อ
2. แบบตัวเลือกคงที่ จำนวน 30 ข้อ
3. แบบจับคู่ จำนวน 30 ข้อ

วิธีดำเนินการสร้างเครื่องมือ

การสร้างเครื่องมือ เพื่อใช้ในการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการสร้างแบบทดสอบ ตามลำดับขั้นตอน ดังนี้



จากภาพประกอบข้างต้น สามารถแสดงรายละเอียดได้ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ

1.1 สร้างแบบทดสอบวัดความรู้ทางเครื่องมือช่าง คือแบบเลือกตอบ, แบบจับคู่ และแบบตัวเลือกคงที่ ตามหลักสูตรของกรมสามัญศึกษา

1.2 เพื่อศึกษาคุณภาพของแบบทดสอบ

2. ศึกษาทฤษฎีเอกสารและงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการสร้างแบบทดสอบ

3. วางแผนการสร้างแบบทดสอบ แบบเลือกตอบ จำนวน 50 ข้อ ประกอบด้วย 2 ส่วน คือรูปภาพที่กำหนดให้เป็นคำถามและส่วนรูปภาพที่เป็นคำตอบให้เลือกซึ่งมี 5 ตัวเลือกโดยกำหนดให้รูปภาพทั้ง 2 ส่วนมีความสัมพันธ์หรือเกี่ยวข้องกัน

4. สร้างข้อคำถามวัดความรู้เกี่ยวกับเครื่องมือ ซึ่งมีรูปแบบของข้อคำถามเป็นแบบเลือกตอบ จำนวน 50 ข้อ

5. นำแบบทดสอบที่สร้างขึ้นในข้อ 4 ไปให้ผู้เชี่ยวชาญ จำนวน 5 ท่าน ตรวจสอบความเที่ยงตรงเชิงพื้นผิว (Face validity) โดยให้ผู้เชี่ยวชาญพิจารณาตรวจสอบแบบทดสอบในแต่ละข้อว่าวัดได้ตรงตามนิยามหรือไม่ จากนั้นนำผลการพิจารณามาหาค่าดัชนีความสอดคล้อง คัดเลือกเฉพาะข้อที่มีค่า IOC มากกว่าหรือเท่ากับ .05

6. ทดลองเครื่องมือ โดยการนำแบบทดสอบที่ผ่านการพิจารณาจากผู้เชี่ยวชาญแล้วไปสอบกับนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 200 คน จากโรงเรียนอัสสัมชัญธนบุรี และกำหนดเวลา ในการให้นักเรียนทำข้อสอบจำนวน 30 ข้อ ใช้เวลา 30 นาที

7. นำผลที่ได้หาค่าคุณภาพ โดยหาค่าความยากง่ายและค่าอำนาจจำแนกแล้วเลือกข้อสอบที่ใช้ได้ตามเกณฑ์ไว้ 30 ข้อ คัดเลือกข้อสอบที่มีค่าความยากในช่วง .20 ถึง .80 และค่าอำนาจจำแนก ตั้งแต่ .20 ขึ้นไป ซึ่งผลจากการทดลองได้ข้อสอบที่มีความยากง่ายในช่วง 0.20 – 0.40 จำนวน 8 ข้อ ช่วง 0.41 – 0.60 จำนวน 8 ข้อ และช่วง 0.61 – 0.80 จำนวน 14 ข้อ

8. นำข้อสอบแบบเลือกตอบที่คัดเลือกไว้ในข้อ 7 มาสร้างแบบทดสอบโดยเปลี่ยนรูปแบบเป็นแบบจับคู่และแบบตัวเลือกคงที่ ฉบับละ 30 ข้อ

9. จัดพิมพ์เป็นรูปเล่มเพื่อใช้ในการวิจัย

ลักษณะของเครื่องมือ

เครื่องมือที่ใช้ในการศึกษาครั้งนี้ เป็นแบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 ฉบับ ดังตัวอย่างต่อไปนี้

ฉบับที่ 1 ลักษณะข้อสอบ เป็นข้อสอบแบบเลือกตอบ ที่ประกอบด้วย 2 ส่วนคือรูปภาพที่กำหนดให้เป็นคำถามและส่วนรูปภาพที่เป็นคำตอบให้เลือก ซึ่งมี 5 ตัวเลือกโดยกำหนดให้รูปภาพทั้ง 2 ส่วนมีความสัมพันธ์หรือเกี่ยวข้องกัน

คำชี้แจง ในแต่ละข้อจะกำหนดรูปภาพมาให้ 1 รูป ทางซ้ายมือแล้วให้หาภาพจากตัวเลือก ก. ข. ค. ง. และ จ. ว่าข้อใดเป็นภาพที่มีความสัมพันธ์เกี่ยวข้องกันกับภาพที่กำหนดให้

ภาพที่กำหนดให้	ก	ข	ค	ง	จ
					

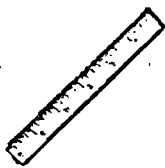
คำตอบที่ถูกต้องคือตัวเลือก ง

ภาพที่กำหนดให้	ก	ข	ค	ง	จ
					

คำตอบที่ถูกต้องคือตัวเลือก ข

ภาพที่กำหนดให้	ก	ข	ค	ง	จ
					

คำตอบที่ถูกต้องคือตัวเลือก ค

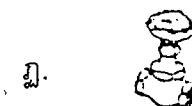
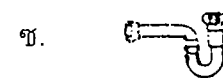
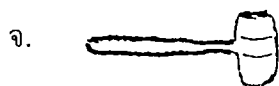
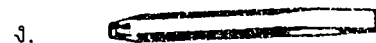
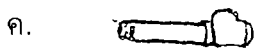
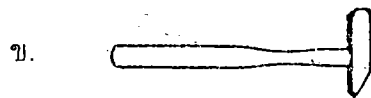
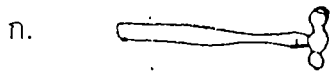
ภาพที่กำหนดให้	ก	ข	ค	ง	จ
					

คำตอบที่ถูกต้องคือตัวเลือก จ

ฉบับที่ 2 ลักษณะข้อสอบ แบบตัวเลือกที่ ประกอบด้วย 2 ส่วนคือส่วนที่เป็นชุดของข้อความเป็นคำถามและส่วนที่เป็นชุดของรูปภาพเป็นตัวเลือก โดยกำหนดให้ชุดของตัวเลือกมีความสัมพันธ์เกี่ยวข้องกัน

คำชี้แจง ใช้ภาพที่กำหนดให้จากภาพ ก ถึง ฎ ตอบคำถามตั้งแต่ข้อ 1 – 4 โดยพิจารณาว่าข้อความแต่ละข้อมีความสัมพันธ์หรือเกี่ยวข้องกับภาพใดมากที่สุด

1. ใช้ในงานประเพณีขึ้นรูปวัสดุงาน เช่น ดอกย่ำ
2. ใช้ต่อกับกับก๊อกน้ำ
3. ใช้เมื่อท่อน้ำเกิดอุดตัน
4. ใช้ขีดขนาดต่าง ๆ ที่ต้องการลงบนชิ้นงาน



คำตอบของข้อ 1 คือ ข้อ ก

คำตอบของข้อ 2 คือ ข้อ ค

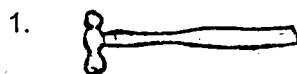
คำตอบของข้อ 3 คือ ข้อ ฎ

คำตอบของข้อ 4 คือ ข้อ ฉ

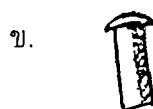
ฉบับที่ 3 ลักษณะข้อสอบ เป็นข้อสอบแบบจับคู่ ประกอบด้วยชุดของรูปภาพ 2 ชุด คือ ชุดของภาพยืนและชุดของภาพเลือกที่มีความสัมพันธ์หรือเกี่ยวข้องกัน

คำชี้แจง ให้นักเรียนดูรูปภาพแต่ละข้อในแนวทางซ้ายมือ แล้วพิจารณาว่าแต่ละข้อ สัมพันธ์เกี่ยวข้องกับข้อใดที่กำหนดไว้จากแนวทางขวามือ เลือกได้ข้อใดนำไปเขียนตอบลงใน กระดาษคำตอบ

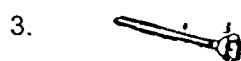
ภาพที่กำหนดให้



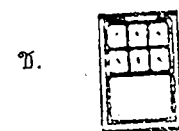
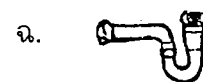
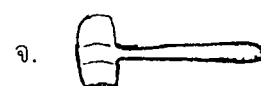
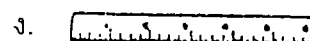
ภาพที่กำหนดให้



ภาพที่กำหนดให้



ภาพที่กำหนดให้



คำตอบข้อ 1 คือ ข้อ ข

คำตอบข้อ 2 คือ ข้อ ก

คำตอบข้อ 3 คือ ข้อ ฉ

คำตอบข้อ 4 คือ ข้อ ง

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ทำการเก็บรวบรวมข้อมูลตามขั้นตอนต่อไปนี้

1. ขอลหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์ในการเก็บรวบรวมข้อมูล
2. ติดต่อโรงเรียนที่เลือกเป็นกลุ่มตัวอย่าง ขออนุญาตผู้บริหารโรงเรียนกำหนดวันเวลาและสถานที่เพื่อนำแบบทดสอบไปทำการทดสอบ
3. นำแบบทดสอบวัดความรู้ทางเครื่องมือช่างทั้ง 3 ฉบับไปทดสอบกับกลุ่มตัวอย่าง โดยอธิบายให้นักเรียนทราบวัตถุประสงค์ของการสอบ และขอความร่วมมือในการทำแบบทดสอบ เพื่อให้ได้ผลตรงกับความเป็นจริงมากที่สุด คือให้นักเรียนห้องหนึ่งทำทั้ง 3 ฉบับ โดยแจกแบบทดสอบสลับกัน หนึ่งคนจะทำ 1 ฉบับ
4. นำกระดาษคำตอบที่ได้มาแยกเป็นแต่ละฉบับ แล้วตรวจให้คะแนนตามเกณฑ์ที่กำหนด ดังนี้ ถ้าตอบถูกให้คะแนน 1 คะแนน ถ้าตอบผิดหรือเว้นว่างให้ 0 คะแนน
4. นำผลที่ได้มาคำนวณหาค่าสถิติ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลครั้งนี้ผู้วิจัยทำการวิเคราะห์ด้วยโปรแกรมคอมพิวเตอร์สำเร็จรูป ได้แก่ โปรแกรมวิเคราะห์ข้อสอบ โปรแกรม SPSS FOR WINDOWS ผู้วิจัยได้ดำเนินการตามลำดับขั้น ดังนี้

1. หาค่าสถิติพื้นฐาน
 - 1.1 หาค่าเฉลี่ย ($\bar{X}$)
 - 1.2 ค่าส่วนเบี่ยงเบนมาตรฐาน (SD)
2. วิเคราะห์ข้อสอบรายข้อ
 - 2.1 การหาค่าความยากง่าย (P) โดยใช้สัดส่วนของจำนวนผู้ตอบถูกในข้อนั้นกับจำนวนผู้ตอบในข้อนั้นทั้งหมด
 - 2.2 การหาค่าอำนาจจำแนก (r) โดยใช้ค่าสหสัมพันธ์แบบพ้อยท์ - ไบซีเรียล (Allen and Yen. 1979 : 122)

$$r_{pbis} = \frac{\bar{X}_i - \bar{X}}{S_x} \cdot \sqrt{\frac{P_i}{1 - P_i}}$$

เมื่อ	r_{pbis}	แทน ค่าอำนาจจำแนก
	$\bar{X}_i$	แทน คะแนนเฉลี่ยของนักเรียนที่ทำข้อที่ i
	$\bar{X}$	แทน คะแนนเฉลี่ยของแบบทดสอบ
	S_x	แทน ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบฉบับนั้น
	P_i	แทน สัดส่วนของคนที่ทำข้อที่ i ถูก

3. หาความเชื่อมั่น โดยใช้สูตรสัมประสิทธิ์ความเชื่อมั่นของ Feldt - Raju (บุญเชิด ภิญโญนนตพงษ์. 2533 : 301)

$$r_{F-R} = \frac{1}{1 - \sum \lambda_i^2} \left[1 - \frac{\sum S_i^2}{S^2} \right]$$

เมื่อ	r_{F-R}	แทน ค่าความเชื่อมั่นของแบบทดสอบ
	S_i^2	แทน ความแปรปรวนของคะแนนแต่ละด้าน
	S^2	แทน ความแปรปรวนของแบบทดสอบทั้งฉบับ
	λ_i^2	แทน ค่าความแปรปรวนของความแปรปรวนรวม ของคะแนนทั้งหมด

4. สถิติทดสอบนัยสำคัญของความแตกต่างของความเชื่อมั่นของแบบทดสอบ โดยใช้สูตร UX_1 (บุญเชิด ภิญโญนนตพงษ์. 2537 : 57 ; อ้างอิงมาจาก Woodruff and Feldt. 1986 : 393-413)วิเคราะห์โดยใช้โปรแกรมสำเร็จรูป ALPHATEST

$$UX_1 = \sum_{i=1}^m \frac{(u_i - \bar{u})^2}{S_u^2 - C_u}$$

$$\bar{u} = \sum_{i=1}^m \frac{u_i}{m}$$

$$u_i = \frac{1}{(1 - r_i)^3}$$

$$S_u^2 = \frac{2}{9m(N_c - 1)} \sum_{i=1}^m u_i^2$$

$$C_u = \frac{4}{9m(m-1)(N_c - 1)} \sum_{i=1}^m \sum_{j=1}^{i-1} r_{ij}^2 u_i u_j$$

$$N_c = N \left(\frac{\bar{n}_h - 1}{n_h + 1} \right)$$

$$\bar{n}_h = \frac{M}{\sum \frac{1}{n_i}}$$

เมื่อ UX_i แทน ค่าสถิติทดสอบที่มีการแจกแจงแบบ χ^2 , $df = m - 1$

r แทน ค่าความเชื่อมั่น

r_{ij} แทน ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร

S_u^2 แทน ค่าความแปรปรวนของ u_i

C_u แทน ค่าความแปรปรวนร่วมของ u_i

m แทน จำนวนค่าสัมประสิทธิ์ความเชื่อมั่น

N แทน จำนวนนักเรียนที่ทดสอบ

$\bar{u}$ แทน ค่าเฉลี่ยของคะแนน

$\bar{n}_h$ แทน จำนวนข้อสอบ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

จากการวิจัยครั้งนี้ เพื่อให้เกิดความเข้าใจตรงกันในการแปลความหมายของผลการวิเคราะห์ข้อมูล ผู้วิจัยจึงได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังนี้

N	แทน	จำนวนนักเรียน
$\bar{X}$	แทน	คะแนนเฉลี่ยของแบบทดสอบ
SD	แทน	ส่วนเบี่ยงเบนมาตรฐาน
CV	แทน	สัมประสิทธิ์การกระจาย
S^2	แทน	ความแปรปรวนของคะแนนทั้งฉบับ
S_i^2	แทน	ความแปรปรวนของคะแนน
λ_i^2	แทน	ความแปรปรวนของความแปรปรวนรวมของคะแนนทั้งหมด
r_{F-R}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
UX_i	แทน	สถิติทดสอบที่มีการแจกแจงแบบ ไค สแควร์ - กำลังสอง

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในการวิจัยครั้งนี้ มีขั้นตอนการดำเนินการดังนี้

1. หาค่าสถิติพื้นฐานของแบบทดสอบ ค่าความเที่ยงตรงที่คำนวณด้วยสูตรสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน และค่าความเชื่อมั่นที่คำนวณด้วยสูตรของเฟลท์ ราซุ
2. ทดสอบความแตกต่างของค่าความเชื่อมั่นของแบบทดสอบที่มีรูปแบบคำถามแตกต่างกันตามสูตร UX_i

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ

ผู้วิจัยนำแบบทดสอบทั้ง 3 ฉบับ ฉบับละ 30 ข้อ ไปทดสอบกับกลุ่มตัวอย่าง แล้วนำกระดาษคำตอบของนักเรียนมาตรวจให้คะแนน แล้วนำคะแนนที่ได้มาคำนวณหาค่าสถิติพื้นฐาน ได้แก่ คะแนนเฉลี่ย คะแนนเบี่ยงเบนมาตรฐาน ดังแสดงในตารางที่ 2 ดังนี้

ตาราง 4 คะแนนเฉลี่ย ค่าเบี่ยงเบนมาตรฐานของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง

แบบทดสอบ	N	K	$\bar{X}$	SD	CV
แบบเลือกตอบ	392	30	16.34	6.56	40.146
แบบตัวเลือกคงที่	401	30	13.79	6.81	49.384
แบบจับคู่	407	30	15.06	6.09	40.438

ผลการวิเคราะห์ตามตาราง 2 พบว่า แบบทดสอบวัดความรู้ทางเครื่องมือช่างที่มีรูปแบบคำถามแบบเลือกตอบ แบบตัวเลือกคงที่ แบบจับคู่มีคะแนนเฉลี่ยเท่ากับ 16.34, 13.79 และ 15.06 คะแนนตามลำดับ โดยแบบทดสอบแบบเลือกตอบง่ายที่สุดรองลงมาคือแบบจับคู่และแบบทดสอบแบบตัวเลือกคงที่ยากที่สุด ส่วนเบี่ยงเบนมาตรฐานของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง มีค่าตั้งแต่ 6.09 ถึง 6.81 โดยแบบทดสอบแบบเลือกตอบมีค่าสูงที่สุดคือ 6.81 และรองลงมาคือแบบทดสอบแบบตัวเลือกคงที่มีคะแนนส่วนเบี่ยงเบนมาตรฐานเท่ากับ 6.56 ส่วนแบบทดสอบแบบจับคู่มีค่าต่ำสุดคือ 6.09 และค่าสัมประสิทธิ์การกระจาย มีค่าเท่ากับ 40.146 ถึง 49.384 โดยทั้งสามฉบับมีค่าการกระจายใกล้เคียงกัน

2. หาค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง

ผู้วิจัยได้นำคะแนนของแบบทดสอบวัดความรู้ทางเครื่องมือช่างของรูปแบบคำถามแบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่ โดยแบ่งแบบทดสอบแต่ละรูปแบบออกเป็น 6 ส่วนๆ ละ 5 ข้อ แล้วนำมาคำนวณค่าความแปรปรวนของคะแนนในแต่ละรูปแบบ (S^2) ผลรวมของค่าความแปรปรวนของคะแนนในแต่ละส่วน ($\sum S_i^2$) และผลรวมของค่าความแปรปรวนของความแปรปรวนรวม ($\sum \lambda_i^2$) แล้วนำค่าที่ได้ไปแทนในสูตร r_{F-R} ของ Feldt – Raju เพื่อหาค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างของแต่ละรูปแบบ ดังแสดงในตาราง 3 ดังนี้

ตาราง 5 ค่าความแปรปรวนของคะแนนในแต่ละรูปแบบ (S^2) ผลรวมของค่าความแปรปรวนของคะแนนในแต่ละส่วน (ΣS_i^2) ผลรวมของค่าความแปรปรวนของความแปรปรวนร่วม ($\Sigma \lambda_i^2$) และค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างของรูปแบบคำถามแบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่

แบบทดสอบ	S^2	ΣS_i^2	$\Sigma \lambda_i^2$	r_{F-R}
แบบเลือกตอบ	42.972	12.7599	.066693	.7533
แบบตัวเลือกคงที่	46.381	12.1473	.066654	.6994
แบบจับคู่	37.053	10.9874	.078184	.7021

ผลการวิเคราะห์ตามตาราง 3 พบว่าค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างคำนวณด้วยสูตร Feldt – Raju แบบเลือกตอบมีค่า เท่ากับ .7533 แบบตัวเลือกคงที่มีค่าเท่ากับ .6994 และแบบจับคู่มีค่าเท่ากับ .702 แสดงว่ามีค่าความเชื่อมั่นอยู่ในระดับค่อนข้างสูง

๑ (2) การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างของรูปแบบคำถามแบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่

การวิเคราะห์ตอนนี้ได้นำค่าความเชื่อมั่น ของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 รูปแบบ ที่คำนวณด้วยสูตร r_{F-R} ของ Feldt – Raju มาทดสอบความแตกต่างโดยใช้สูตร UX_i ปรากฏผลดังแสดงในตาราง 4 ดังนี้

ตาราง 6 เปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 รูปแบบที่คำนวณด้วยสูตร r_{F-R} ของ Feldt – Raju

แบบทดสอบ	N	UX_i	P
แบบเลือกตอบ	392		
แบบตัวเลือกคงที่	401	4.7335	.09163
แบบจับคู่	407		

ผลการวิเคราะห์ตามตาราง 4 พบว่าค่าความเชื่อมั่นที่คำนวณด้วยสูตร r_{F-R} ของ Feldt – Raju มีค่าแตกต่างกันไม่อย่างมีนัยสำคัญทางสถิติ

บทที่ 5

สรุป อภิปรายผลและข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายสำคัญเพื่อศึกษารูปแบบคำถามของแบบทดสอบวัดความรู้ทางเครื่องมือช่างมีผลต่อค่าความเชื่อมั่นมากน้อยเพียงใด โดยกำหนดเป็นจุดมุ่งหมายเฉพาะของการวิจัยครั้งนี้

1. เพื่อหาค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง แบบเลือกตอบ, แบบตัวเลือกคงที่ และแบบจับคู่
2. เพื่อศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบความรู้ทางเครื่องมือช่าง แบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าในครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่กำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2543 ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 116 โรงเรียน จำนวนนักเรียน 53,774 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่เรียนวิชาเลือกเรียนวิชางานช่างพื้นฐาน กำลังศึกษาในภาคเรียนที่ 2 ปีการศึกษา 2543 ในกรุงเทพมหานคร สังกัดกรมสามัญศึกษา จำนวน 1,200 คน ซึ่งได้มาจากการสุ่ม 2 ขั้นตอน (Random Sampling)

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ คือแบบทดสอบวัดความรู้ทางเครื่องมือช่างวิชางานช่างพื้นฐาน (ง 013) เนื้อหาบทที่ 1-6 ที่ผู้วิจัยสร้างขึ้นเองจำนวน 3 ฉบับ ฉบับละ 30 ข้อ แต่ละฉบับมีรูปแบบคำถามแตกต่างกันดังนี้

1. แบบเลือกตอบ
2. แบบตัวเลือกคงที่
3. แบบจับคู่

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยนำแบบทดสอบวัดความรู้ทางเครื่องมือช่างทั้ง 3 รูปแบบ ไปทดสอบกับนักเรียนที่เป็นกลุ่มตัวอย่างโดยวิธีการเก็บรวบรวมข้อมูลดังนี้

1. ติดต่อโรงเรียนที่เลือกเป็นกลุ่มตัวอย่าง ขออนุญาตผู้บริหารโรงเรียนกำหนดวันเวลาและสถานที่
2. วางแผนดำเนินการสอบ โดยผู้วิจัยติดต่อกับอาจารย์ที่สอนวิชางานช่างพื้นฐานเพื่อให้อาจารย์กรุณาเก็บให้
3. จัดเตรียมแบบทดสอบและกระดาษคำตอบให้เพียงพอกับจำนวนกลุ่มตัวอย่างที่จะทดสอบแต่ละโรงเรียน และเรียงลำดับแบบทดสอบเป็นชุดๆ ตั้งแต่รูปแบบที่ 1 ถึงรูปแบบที่ 3 หมุนเวียนกันไปอย่างเป็นระบบ จัดส่งให้อาจารย์ที่ติดต่อไว้ในแต่ละโรงเรียน
4. ทดสอบนักเรียนโดยให้ปฏิบัติดังนี้
 - 4.1 จัดให้นักเรียนนั่งเป็นแถว
 - 4.2 ชี้แจงให้นักเรียนทราบวัตถุประสงค์ของการสอบ
 - 4.3 แจกแบบทดสอบให้นักเรียนตามลำดับของแบบทดสอบที่จัดไว้ คือ รูปแบบ 1 แจกให้คนที่ 1 และรูปแบบที่ 2 – 3 แจกให้คนที่ 2 – 3 ตามลำดับ แจก รูปแบบที่ 1 ให้คนที่ 4 และรูปแบบที่ 2 - 3 ให้คนที่ 5 – 6 ตามลำดับ และแจกแบบทดสอบให้คนต่อๆ ไปตามระบบที่จัดไว้
 - 4.4 ชี้แจงวิธีตอบแบบทดสอบ
5. หลังดำเนินการสอบนักเรียนที่เป็นกลุ่มตัวอย่างครบทุกโรงเรียนแล้ว รวบรวมกระดาษคำตอบของแบบทดสอบแต่ละรูปแบบ
6. นำกระดาษคำตอบที่ได้จากข้อ 5 มาตรวจให้คะแนน ด้วยวิธี 0 – 1 เพื่อนำผลมาวิเคราะห์ทางสถิติและทดสอบสมมติฐาน

สรุปผลการวิเคราะห์ข้อมูล

การศึกษาครั้งนี้มุ่งที่จะศึกษารูปแบบคำถามของแบบทดสอบวัดความรู้ทางเครื่องมือช่างมีผลต่อค่าความเชื่อมั่นมากน้อยเพียงใด โดยศึกษาค่าความเชื่อมั่นของแบบทดสอบวัด

ความรู้ทางเครื่องมือช่างแบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่ จากสูตร Feldt – Raju และการเปรียบเทียบค่าความเชื่อมั่นที่คำนวณได้ โดยการทดสอบความแตกต่างความเชื่อมั่นด้วยสูตร UX_1 ผลการวิจัยปรากฏผลดังนี้

ค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างที่มีรูปแบบคำถามแบบเลือกตอบ มีค่าเท่ากับ 0.7533 แบบตัวเลือกคงที่ มีค่าเท่ากับ 0.6994 และแบบจับคู่ มีค่าเท่ากับ 0.7021 เมื่อนำค่าความเชื่อมั่นที่คำนวณได้ไปทดสอบความแตกต่างความเชื่อมั่นด้วยสูตร UX_1 พบว่าค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ($UX_1 = 4.7335$ $P < 0.09163$)

อภิปรายผล

จากผลการวิเคราะห์พบว่า

ค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่าง ที่มีรูปแบบคำถามแบบเลือกตอบ แบบตัวเลือกคงที่ แบบจับคู่ มีค่าเท่ากับ 0.7533 , 0.6994 และ 0.7021 ตามลำดับ กล่าวได้ว่าแบบทดสอบวัดความรู้ทางเครื่องมือช่างทั้ง 3 รูปแบบ มีค่าความเชื่อมั่นค่อนข้างสูง (ล้วน สายยศ และอังคณา สายยศ. 2539 : 244) หมายความว่าแบบทดสอบนั้นมีผลการวัดที่มีความคงที่แน่นอนเป็นที่เชื่อถือได้ และเมื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ทางเครื่องมือช่างที่คำนวณด้วยสูตร Feldt – Raju ของแบบทดสอบทั้ง 3 รูปแบบ พบว่ามีค่าความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติซึ่งไม่สอดคล้องกับสมมุติฐานการวิจัยที่ผู้วิจัยคาดหมายไว้ แต่สอดคล้องกับงานวิจัยของไพบุลย์ จิตรโรต(2514) และ ออสเตอร์ฮอฟและพามาเลา(1984) ที่กล่าวว่าแบบทดสอบแบบเลือกตอบชนิดต่างๆ และแบบเติมคำมีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญ แบบทดสอบวัดความรู้ทางเครื่องมือช่างทั้ง 3 รูปแบบนั้นเป็นแบบทดสอบที่ใช้ภาพเป็นทั้งคำถามและคำตอบเป็นการสร้างแรงจูงใจไม่ให้เกิดความเบื่อหน่ายมีความเพลิดเพลินเป็นการท้าทายความสามารถของนักเรียนโดยไม่รู้ตัว และส่วนหนึ่งอาจมาจากแบบทดสอบแบบจับคู่ นั้นไม่เป็นไปตามหลักการของการสร้างแบบทดสอบคู่ขนาน โดยผู้วิจัยสร้างส่วนที่เป็นคำถามของแบบจับคู่โดยนำตัวคำตอบของแบบเลือกตอบ แบบตัวเลือกคงที่ มาสร้าง ซึ่งที่ถูกต้องๆ นำคำถามของแบบทดสอบทั้ง 2 รูปแบบมาสร้างเป็นคำถามของแบบจับคู่ด้วย

ข้อเสนอแนะในการวิจัย

1. การสร้างแบบทดสอบแบบจับคู่ ชุดที่เป็นชุดคำถามควรจะเป็นชุดเดียวกับคำถามในแบบเลือกตอบและแบบตัวเลือกคงที่

2. การทำวิจัยครั้งต่อไปไม่ควรทำในเวลาที่นักเรียนกำลังจะสอบภาคปลายเพราะจะทำให้
ให้นักเรียนไม่ตั้งใจทำเท่าที่ควร

3. ควรศึกษารูปแบบข้อคำถามในทำนองเดียวกันนี้กับวิชาอื่นๆ เพื่อประโยชน์ทาง
การศึกษาต่อไป

ข้อเสนอแนะในการนำไปใช้

จากการวิจัย พบว่าแบบทดสอบวัดความรู้ทางเครื่องมือช่างทั้ง 3 รูปแบบคือ
แบบเลือกตอบ แบบตัวเลือกคงที่ และแบบจับคู่ มีค่าความเชื่อมั่นไม่แตกต่างกัน ฉะนั้น การ
เลือกแบบทดสอบจึงสามารถเลือกที่จะนำไปใช้ได้ทั้ง 3 รูปแบบ ทั้งนี้ควรคำนึงความสะดวก
และองค์ประกอบอื่น ๆ ด้วย

บรรณานุกรม

บรรณานุกรม

- กรมวิชาการ กระทรวงศึกษาธิการ. (2534). คู่มือการเรียนการสอนการงาน งานช่างพื้นฐาน. โรงพิมพ์คุรุสภาลาดพร้าว. กรุงเทพมหานคร.
- :กระทรวงศึกษาธิการ. (2534). หลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). โรงพิมพ์คุรุสภา. กรุงเทพมหานคร.
- กรองทอง เทพศิริอำนวย. (2525). การเปรียบเทียบคุณภาพของแบบสอบวิชาคณิตศาสตร์ชนิดเลือกตอบและชนิดเติมคำ. วิทยานิพนธ์ ค.ม.(วัดผลและประเมินผล).กรุงเทพฯ: บัณฑิตวิทยาลัย.จุฬาลงกรณ์มหาวิทยาลัย. อุดສຳເນາ.
- จักรกฤษณ์ สำราญใจ. (2531). ประสิทธิภาพเชิงสัมพันธ์ของข้อสอบเลือกตอบชนิดตัดสินคำตอบทุกตัวเลือกเทียบกับข้อสอบเลือกตอบชนิดแบบฉบับในระบบสอบผลสัมฤทธิ์. วิทยานิพนธ์ดุขฎิบัณฑิต.(วัดผลและประเมินผล). กรุงเทพฯ : บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย. อุดສຳເນາ.
- จินดา โตอนันต์. (2527). การเปรียบเทียบคุณลักษณะของแบบสอบสัมฤทธิ์ผล แบบสอบเลือกตอบถูกผิดทุกตัวเลือกและแบบสอบเลือกตอบตัวเลือกเดียว. วิทยานิพนธ์ ค.ม.(วัดผลและประเมินผล).กรุงเทพฯ : บัณฑิตวิทยาลัย. จุฬาลงกรณ์มหาวิทยาลัย. อุดສຳເນາ.
- จุฑารัตน์ ดันเจริญ. (2538). การเปรียบเทียบคุณภาพของแบบทดสอบเรขาคณิต แบบเติมคำแบบเลือกตอบซ้อนและแบบเลือกตอบถูกผิดซ้อน. ปริญญาโท กศ.ม.(วัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. อุดສຳເນາ.
- ชูศักดิ์ เปลี่ยนภู.(2539). การประเมินผลการเรียนรู้ภาคปฏิบัติ. เอกสารประกอบการสอนวิชา Laboratory and Work shop Instruction ภาควิชาครุศาสตร์ ไฟฟ้า สถาบันเทคโนโลยีพระจอมเกล้าธนบุรี.
- ชวาล แพรัดกุล. (2516) "รูปแบบของคำถามชนิดเลือกตอบ," พัฒนาวัดผล 9. กรุงเทพฯ : กรุงเทพมหานครการพิมพ์.
- :(2518). เทคนิคการวัดผล. พิมพ์ครั้งที่ 6 กรุงเทพฯ: ไทยวัฒนาพานิช.
- เชาวนา ขวลิตธำรง.(2537). การวัดผลการศึกษาปฏิบัติงาน. มหาวิทยาลัยศรีนครินทรวิโรฒ ปทุมวัน.
- นภาพร อมรเลิศสินไทย. (2514). การศึกษาผลการใช้แบบทดสอบแบบเลือกตอบและแบบอัตนัยตอบสั้นๆ วัดระดับความรู้ชั้นต่างๆในวิชาสังคมศึกษา ชั้นประถมศึกษาปีที่ 7 จังหวัดพระนครศรีอยุธยา. ปริญญาโท กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : วิทยาลัยวิชาการศึกษา ประสานมิตร. อุดສຳເນາ.

- นพมพร พรหมมา. (2536). การเปรียบเทียบคุณภาพของแบบทดสอบปรนัยที่มีรูปแบบตัวเลือกต่างกัน. ปรินทิพนิพนธ์. (วัดผลการศึกษา). มหาวิทยาลัยนเรศวร. อุดรธานี.
- บุญเชิด ภิญโญอนันตพงษ์. (2523). การวัดผลและประเมินผลการศึกษา ทฤษฎีและการประยุกต์. กรุงเทพฯ : อักษรเจริญทัศน์.
- บุญเชิด ภิญโญอนันตพงษ์. (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงสัมพันธ์. ปรินทิพนิพนธ์ กศ.ด. (การทดสอบและวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ. อุดรธานี.
- บังอร กมลวัฒนา. (2542). การเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาที่มีรูปแบบการตอบและการจัดเรียงปัญหาต่างกัน. ปรินทิพนิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. อุดรธานี.
- พวงแก้ว ปุณณกนก และ สุวิมล ว่องวานิช. (2524). การวัดภาคปฏิบัติ. จุฬาลงกรณ์มหาวิทยาลัย.
- ไพบุลย์ จิตรโศ. (2514). การใช้แบบทดสอบแบบเลือกตอบชนิดต่าง ๆ และแบบเติมคำ วัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ของนักเรียนชั้นประถมศึกษาปีที่ 4 และประถมศึกษาปีที่ 7 ของโรงเรียนในสังกัดเทศบาลนครธนบุรี. ปรินทิพนิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. วิทยาลัยวิชาการศึกษา ประสานมิตร. อุดรธานี.
- ไพศาล หวังพานิช. (2526). การวัดผลการศึกษา. ไทยวัฒนาพานิช.
- ล้วน สายยศ และอังคณา สายยศ. (2539). เทคนิคการวัดผลการเรียนรู้. กรุงเทพมหานคร : ชมรมเด็ก.
- วรรณพิศ นามเย็น. (2532). ผลของรูปแบบตัวเลือกของแบบทดสอบเลือกตอบที่มีต่อค่าความยาก ค่าอำนาจจำแนกและค่าความเชื่อมั่นของแบบทดสอบในวิชาคณิตศาสตร์. ปรินทิพนิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. อุดรธานี.
- สมบุญ ชิตพงษ์. (2522). การวัดผลสัมฤทธิ์ทางการเรียน. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- สนั่น สิทธิวัง. (2512). การเปรียบเทียบคุณภาพของข้อสอบปรนัยแบบเลือกตอบ ที่ครูคิดสร้างตัวเลือกขึ้นเอง กับข้อสอบที่ใช้ตัวเลือกจากคำตอบของ นักเรียนในวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 4 ประถมปีที่ 7 และชั้นมัธยมศึกษาปีที่ 3. ปรินทิพนิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : วิทยาลัยวิชาการศึกษา ประสานมิตร. อุดรธานี.
- สวัสดี ประทุมราช และคณะ. (2524). การสร้างและพัฒนาแบบทดสอบภาษาอังกฤษ เพื่อใช้สอบคัดเลือกนิสิตระดับบัณฑิตศึกษา. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

- สุธรรม์ จันทรหอม. (2513). การศึกษาผลของการใช้แบบทดสอบแบบถูกผิด แบบเลือกตอบ และแบบเติมคำ วัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นประถมศึกษาปีที่ 7. ปรินูญานิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : วิทยาลัยวิชาการศึกษา ประสานมิตร. อัดสำเนา.
- สุมาลี จันทรชลอ. (2542). การวัดและประเมินผล. ศูนย์สื่อเสริมกรุงเทพ. มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
- สมพงษ์ อิศวศุภฤกษ์. (2542). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจและความยาวของแบบทดสอบต่างกัน. ปรินูญานิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.อัดสำเนา.
- อนันต์ ศรีโสภ. (2520). การวัดและการประเมินผลการศึกษา. กรุงเทพฯ : ไทยวัฒนาพานิช.
- อรอนงค์ บำรุง. (2542). แสดงหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์. ปรินูญานิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- อัมพร วิชัยศรี. (2541) การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างและการประมาณค่าความเชื่อมั่นรวมของแบบทดสอบความสามารถทางสมองตามแนวโครงสร้างของ โอติส – เลนนอน. ปรินูญานิพนธ์ กศ.ม. (วัดผลการศึกษา). กรุงเทพฯ: มหาวิทยาลัยศรีนครินทรวิโรฒ. อัดสำเนา.
- Anastasi, Anne. (1968). *Psychological Testing*. New York : Mc Graw – Hill inc.
- Bingham, Watter Van Dyke. (1937). *Aptitude and Aptitude Testing*. Harper and Brother.
- Baron, Dennis and Harold W Bernard. (1958). *Education Techniques for Classroom Teachers*. New York: McGraw- Hill.
- Binder, William J.R. and Robert A. Davis.(1949.September) "What High School Students Think About Teacher-Made Examinations," *The Journal of Educational Research*. 43 : 58-65 p.
- Donald, Zimmerman W. and Richard H. Williams. *Journal of Experimental Education*. n.d. 179 – 182 p.
- Cronbach, L.J., , N.Rajaratnam and G. C.Gleser. (1963). "Theory of Generalizability : A Liberalization of Reliability Theory," *The British Journal of Statistical Psychology*.
- Cronbach, Lee J. (1970). *Essentials of Psychological Testing*. New York: Harper International Edition Harper & Row.
- Feldt, L.S. (1965). "The Approximate Sampling Distribution of Kuder-Richardson Reliability Coefficient Twenty," *Psychometrika*.30 : 357-370.

- Frisbie, David A., and Swiiney, Daryl C. (1982). "The relative Merits of Multiple True-False Achievement Test," *Journal of Educational Measurement*, 19(1) : 29-35.
- Ferhuson, G.A.(1989). *Statistical Analysis Psychology and Education* 6 th ed Singapore: McGraw Hill.
- Levy, Joan U. and M.R., Novick.(1992). *Mechanical Aptitudes and Spatial Relation*. 2ad ed., New York : Schuster.
- Mehrens, William A and Leehman, Irvin J. Lehmann. (1973). *Measurement and Evaluation in Education and Psychology*. NewYork : Holt Rinehart and Winston.
- Myers, Charl T. "A Note on A Spatial Relations Pretest and Posttest," (1958). *Educational and Psychological Measurement*. 13 : 596 - 600
- Oosterhof, Albert C. and Pameta K. Coats. (1984.Summer). "Comparison of Difficulties and Reliabilities of Quantitative Word Problems in Completion and Multiple-Choice Item Formats," *Applied Psychological Measurement*. 8 : 287 – 293.
- Remmers, H. H. and N. L. Gage. (1955). *Educational Measurement and Evaluation*. Rev. ed., New York, Harper and Brothers.
- Ross, C.C. (1955). *Measurement in Today School*. Prentice-Hall Inc.
- Stanley, Lulian C. and Kenneth D. Hopkins. (1972). *Educational and Psychological Measurement and Evaluation*. 5 th ed. New Jersey : Prentice-Hall,Inc., Englewood Cikffs.

ภาคผนวก

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญประเมินความสอดคล้องของข้อคำถามกับนิยามศัพท์เฉพาะ

รายชื่อผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญประเมินความสอดคล้องของข้อคำถามกับนิยามศัพท์เฉพาะของ
แบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 5 ท่าน ดังนี้

รองศาสตราจารย์นิภา ศรีไพโรจน์	ภาควิชาการวัดผลและวิจัยทางการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์ชวลิต รวยอาจิณ	ภาควิชาการวัดผลและวิจัยทางการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ
อาจารย์โอฬาร กาลวิบูลย์	โรงเรียนฝึกอาชีพกรุงเทพมหานคร (หลวงพ่อดำศักดิ์ ๕)
อาจารย์ชัชชัย มณีโชติ	โรงเรียนประสาธน์รัฐประชากิจ
อาจารย์บุญร่วม เหล็กสูงเนิน	โรงเรียนฝึกอาชีพกรุงเทพมหานคร (ม้วน บำรุงศิลป์)

ภาคผนวก ข

ค่าความยาก(p) อำนาจจำแนก(r) ของแบบทดสอบวัดความรู้ทางเครื่องมือช่างแบบเลือก
ตอบจำนวน 50 ข้อ

ตาราง ค่าความยาก(p) อำนาจจำแนก(r) ของแบบทดสอบวัดความรู้ทางเครื่องมือช่างแบบ
เลือกตอบจำนวน 50 ข้อ

ข้อที่	P	R	ผลการวิเคราะห์
1	.99	-.05	จำแนกไม่ได้ ง่ายมาก คัดออก
2	.42	.30	จำแนกดี ยากง่าย ปานกลาง คัดเลือกไว้
3	.71	.22	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
4	.33	.27	จำแนกปานกลาง ค่อนข้างยาก คัดเลือกไว้
5	.67	.24	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
6	.15	.04	จำแนกต้องปรับปรุง ยากมาก คัดออก
7	.25	-.07	จำแนกไม่ได้ ค่อนข้างยาก คัดออก
8	.76	.59	จำแนกดีมาก ค่อนข้างง่าย คัดเลือกไว้
9	.72	.65	จำแนกดีมาก ค่อนข้างง่าย คัดเลือกไว้
10	.77	.27	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
11	.66	.28	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
12	.45	.25	จำแนกปานกลาง ยากง่ายปานกลาง คัดเลือกไว้
13	.50	.24	จำแนกปานกลาง ยากง่ายปานกลาง คัดเลือกไว้
14	.34	.34	จำแนกดี ค่อนข้างยาก คัดเลือกไว้
15	.29	.35	จำแนกดี ค่อนข้างยาก คัดเลือกไว้
16	.21	.31	จำแนกดี ค่อนข้างยาก คัดเลือกไว้
17	.18	.16	จำแนกต้องปรับปรุง ยากมาก คัดออก
18	.96	.11	จำแนกต้องปรับปรุง ง่ายมาก คัดออก
19	.49	.21	จำแนกปานกลาง ยากง่ายปานกลาง คัดเลือกไว้
20	.27	.17	จำแนกต้องปรับปรุง ค่อนข้างยาก คัดออก
21	.77	.23	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
22	.31	.00	จำแนกไม่ได้ ค่อนข้างยาก คัดออก
23	.63	.18	จำแนกต้องปรับปรุง ค่อนข้างง่าย คัดออก
24	.17	.05	จำแนกต้องปรับปรุง ยากมาก คัดออก
25	.79	.62	จำแนกดีมาก ค่อนข้างง่าย คัดเลือกไว้
26	.51	.20	จำแนกปานกลาง ยากง่ายปานกลาง คัดเลือกไว้
27	.64	.28	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
28	.57	.19	จำแนกต้องปรับปรุง ยากง่ายปานกลาง คัดออก
29	.56	.33	จำแนกดี ยากง่ายปานกลาง คัดเลือกไว้

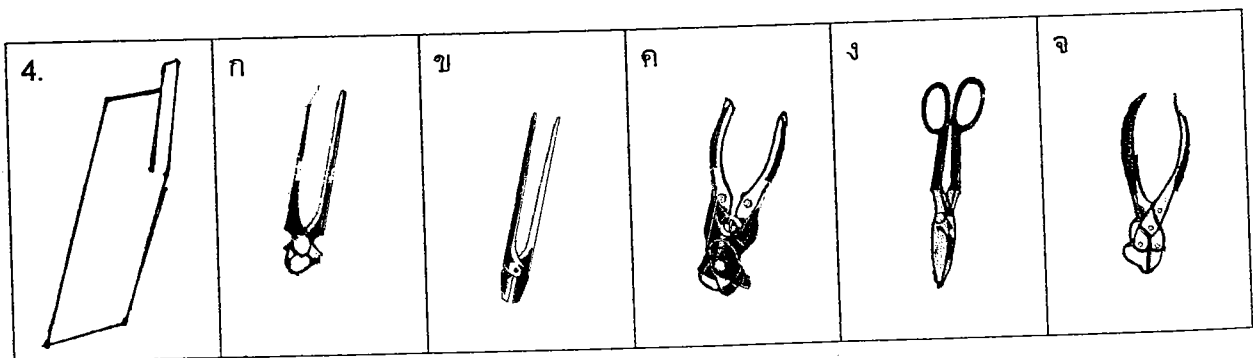
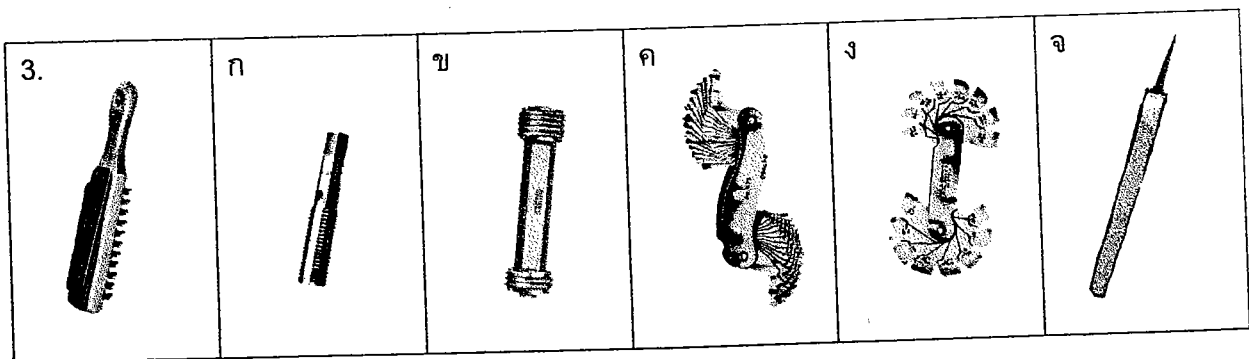
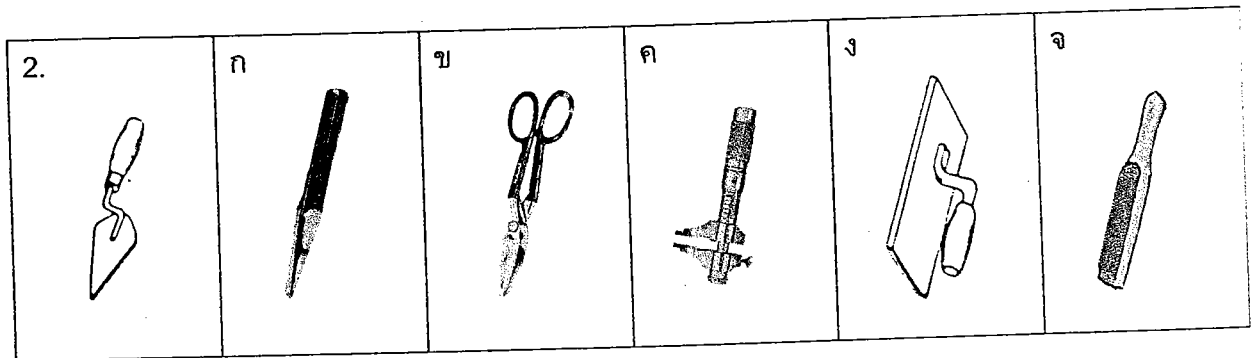
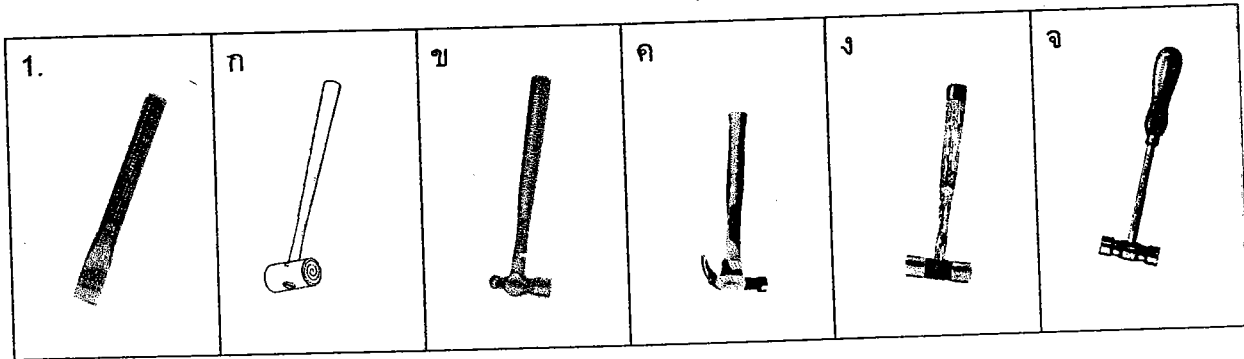
ข้อที่	P	R	ผลการวิเคราะห์
30	.21	.04	จำแนกต้องปรับปรุง ค่อนข้างยาก คัดออก
31	.19	.11	จำแนกต้องปรับปรุง ยากมาก คัดออก
32	.69	.03	จำแนกต้องปรับปรุง ค่อนข้างง่าย คัดออก
33	.13	.12	จำแนกต้องปรับปรุง ยากมาก คัดออก
34	.35	.14	จำแนกต้องปรับปรุง ค่อนข้างยาก คัดออก
35	.75	.55	จำแนกดีมาก ค่อนข้างง่าย คัดเลือกไว้
36	.80	.22	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
37	.62	.25	จำแนกปานกลาง ค่อนข้างง่าย คัดเลือกไว้
38	.12	.24	จำแนกปานกลาง ยากมาก คัดออก
39	.32	.23	จำแนกปานกลาง ค่อนข้างยาก คัดเลือกไว้
40	.31	.22	จำแนกปานกลาง ค่อนข้างยาก คัดเลือกไว้
41	.64	.32	จำแนกดี ค่อนข้างง่าย คัดเลือกไว้
42	.36	.19	จำแนกต้องปรับปรุง ค่อนข้างยาก คัดออก
43	.21	.22	จำแนกปานกลาง ค่อนข้างยาก คัดเลือกไว้
44	.79	.51	จำแนกดีมาก ค่อนข้างง่าย คัดเลือกไว้
45	.44	.44	จำแนกดีมาก ยากง่ายปานกลาง คัดเลือกไว้
46	.26	.10	จำแนกต้องปรับปรุง ค่อนข้างยาก คัดออก
47	.43	.36	จำแนกดี ยากง่ายปานกลาง คัดเลือกไว้
48	.21	.24	จำแนกปานกลาง ค่อนข้างยาก คัดเลือกไว้
49	.08	-.05	จำแนกไม่ได้ ยากมาก คัดออก
50	.03	.17	จำแนกไม่ได้ ยากมาก คัดออก

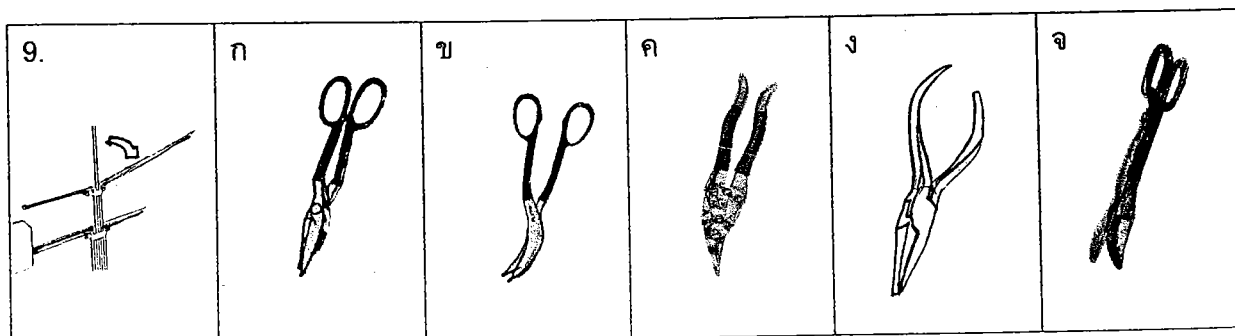
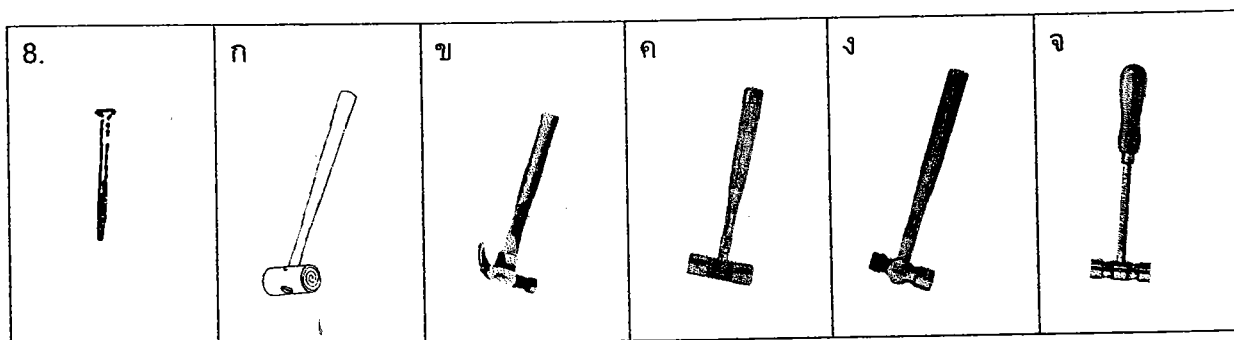
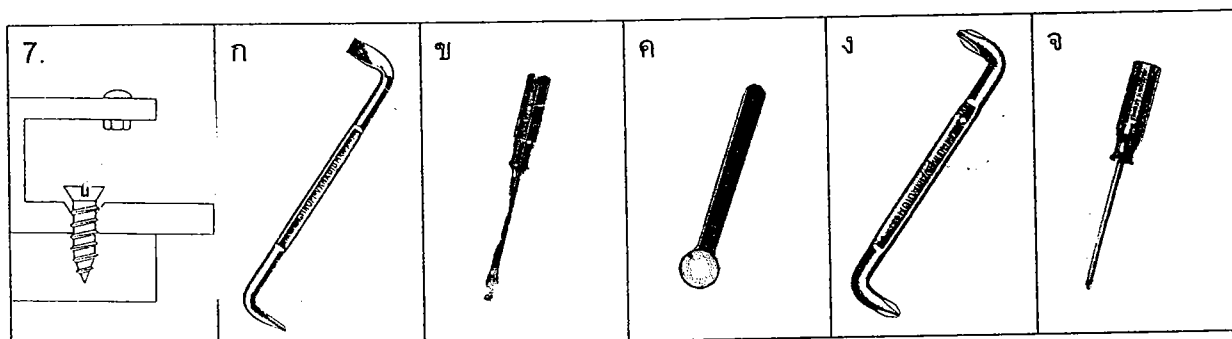
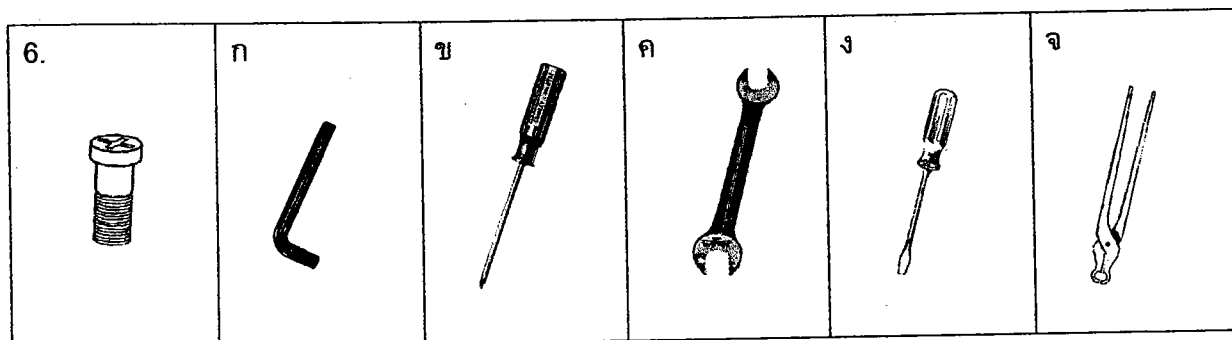
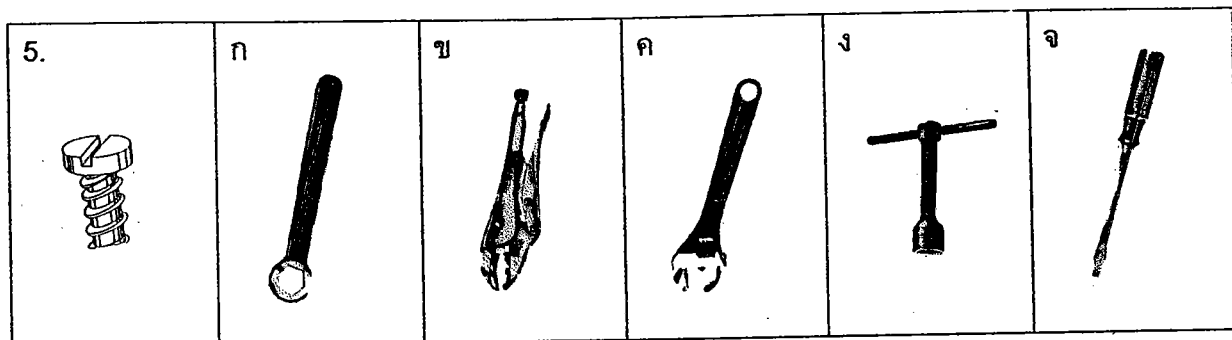
ภาคผนวก ค
แบบทดสอบวัดความรู้ทางเครื่องมือช่าง จำนวน 3 รูปแบบ

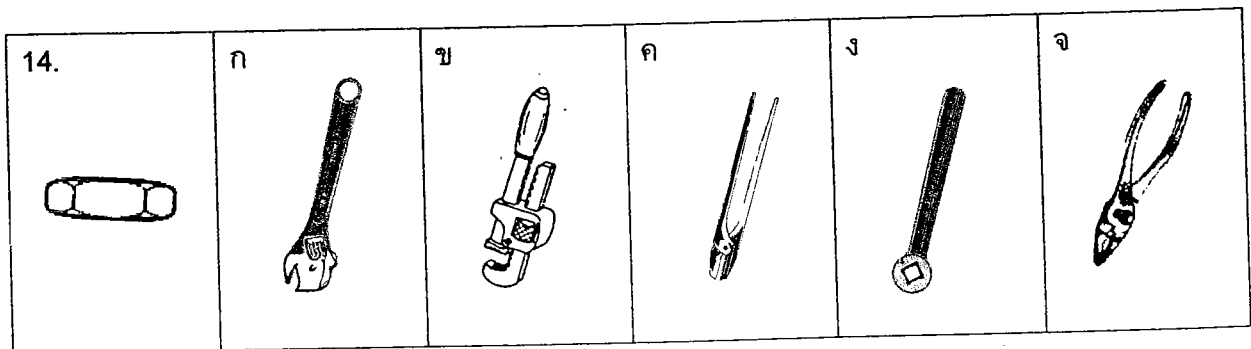
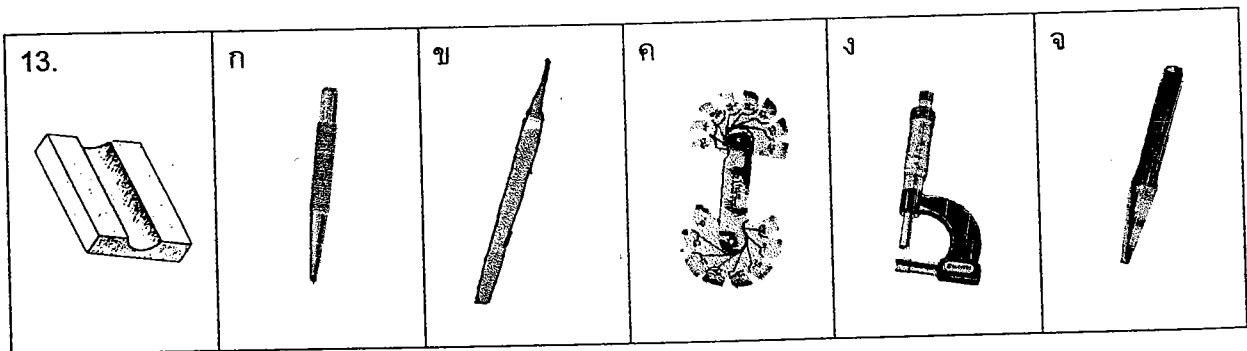
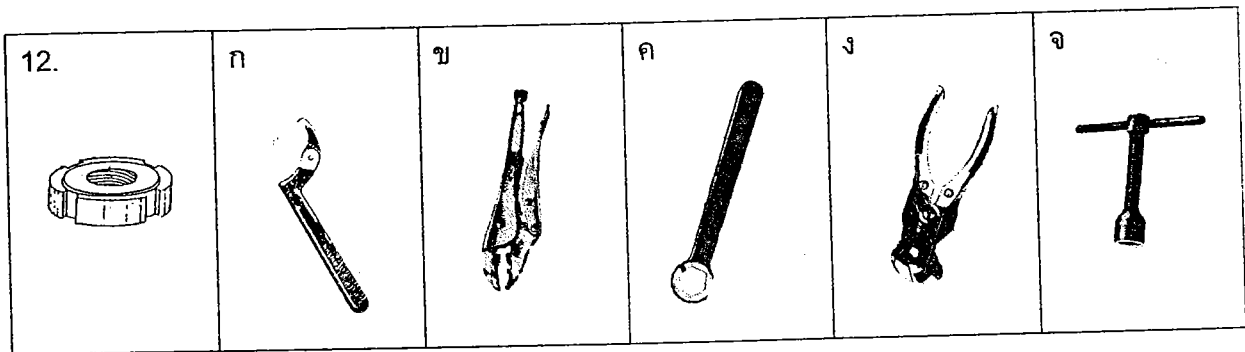
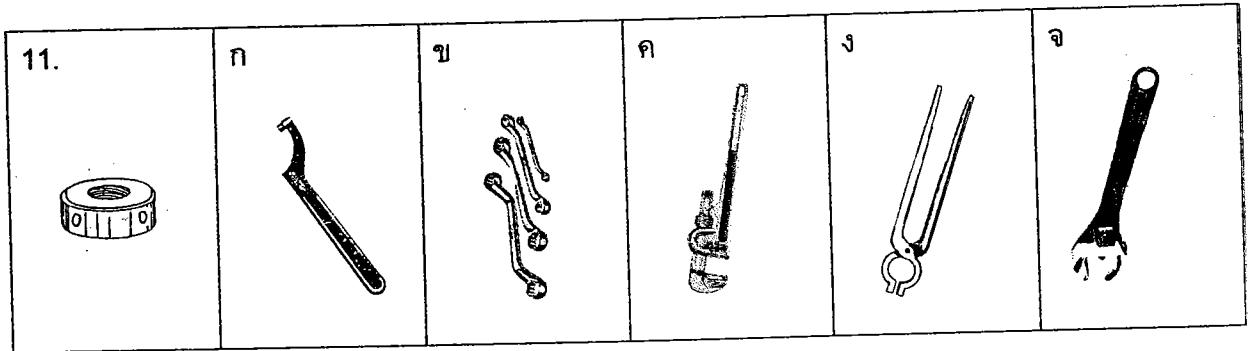
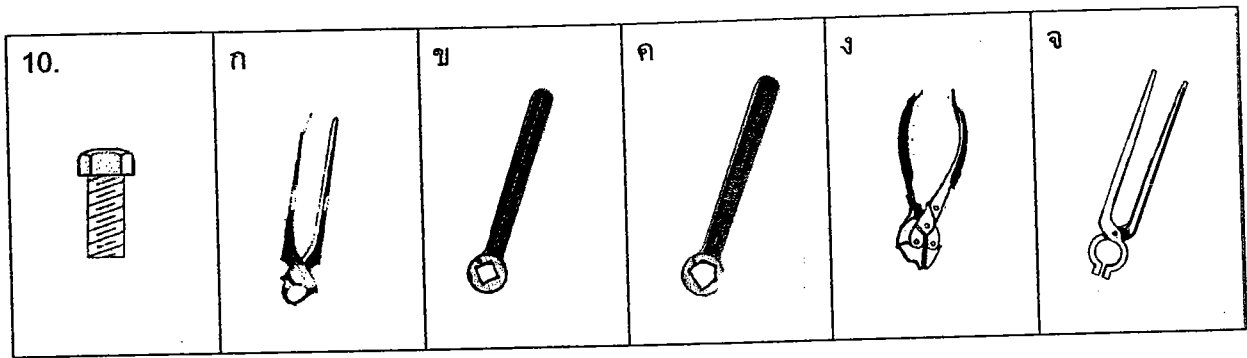
แบบทดสอบฉบับที่ 1

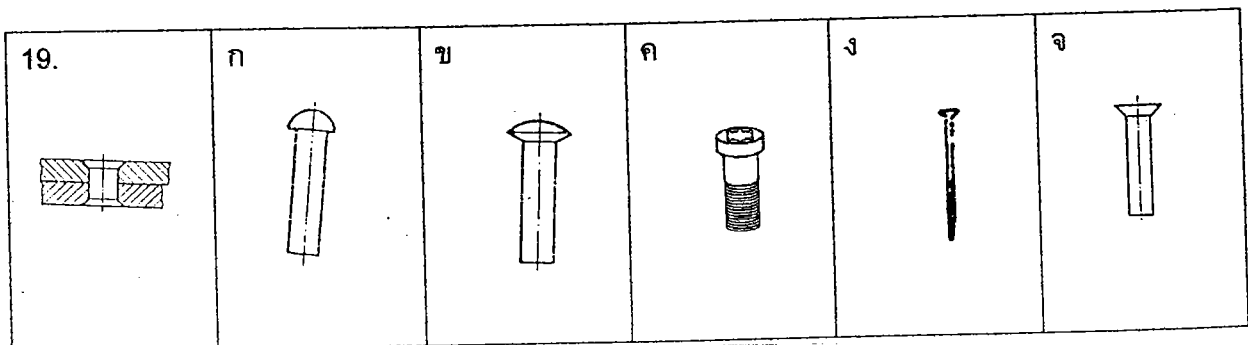
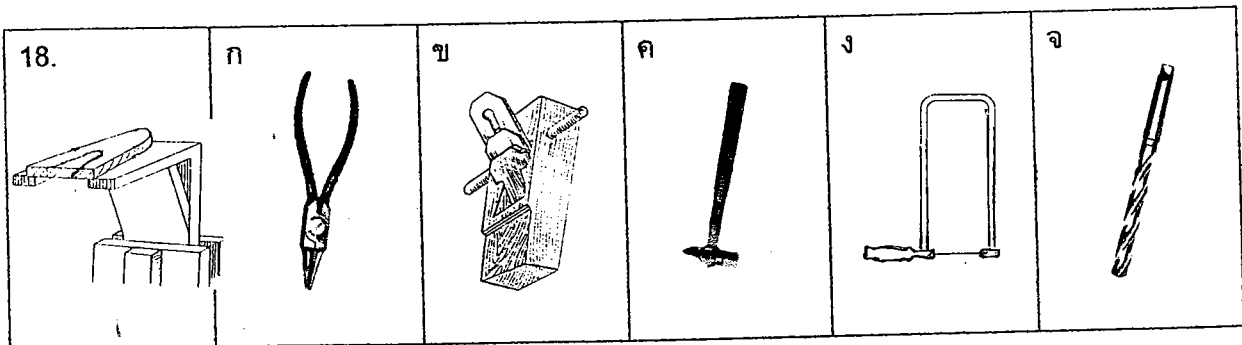
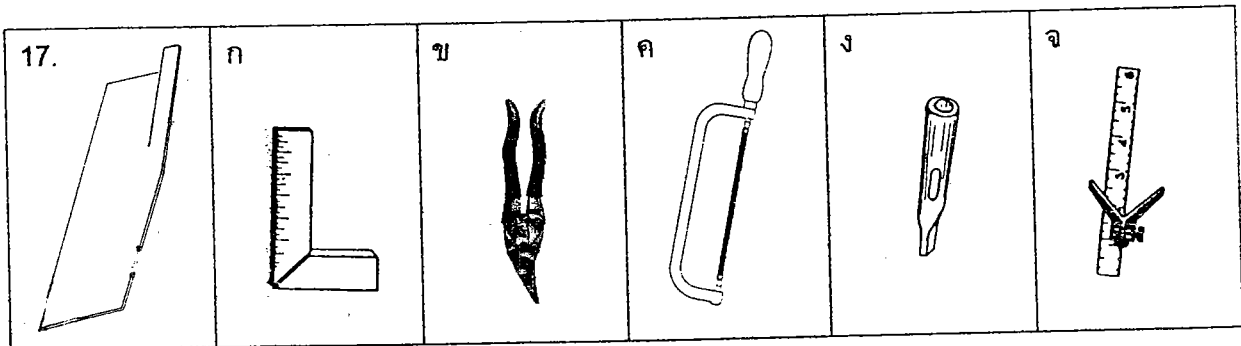
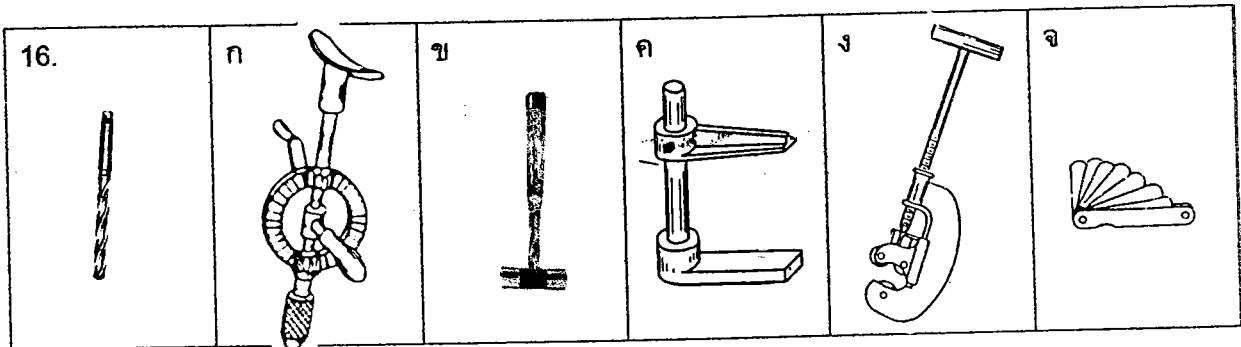
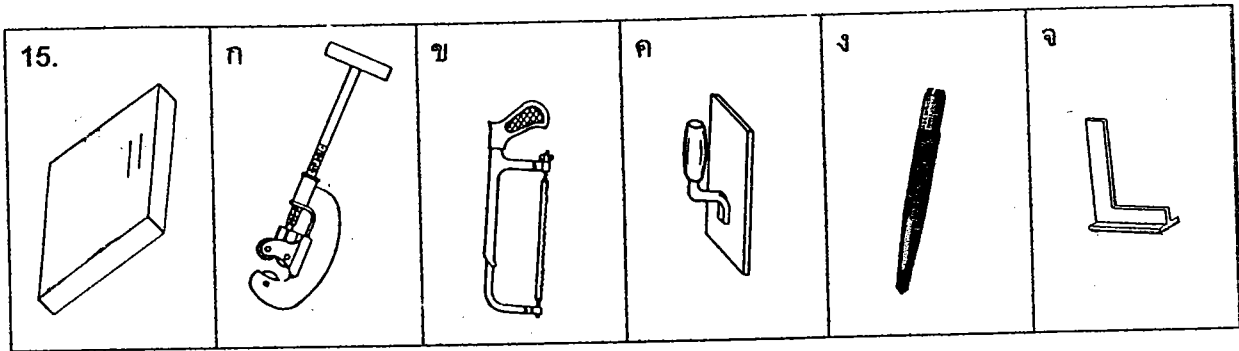
แบบทดสอบฉบับนี้มีทั้งหมด 30 ข้อ มีเวลาทำ 30 นาที

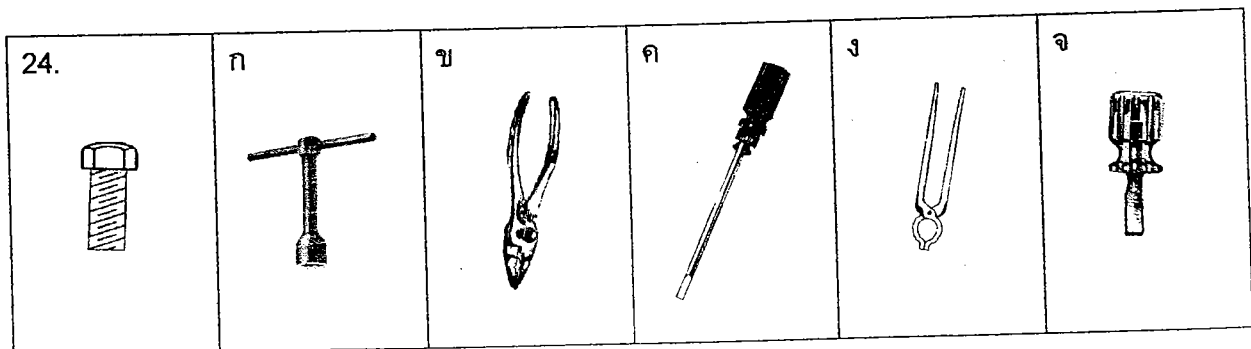
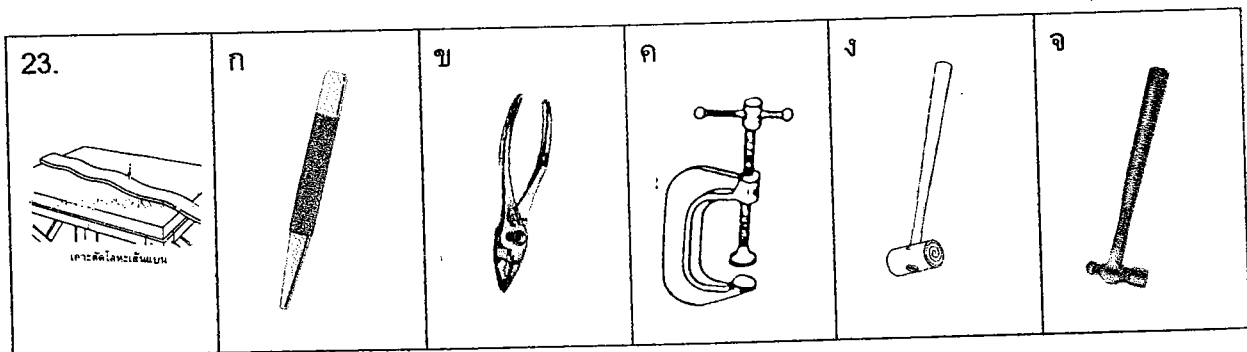
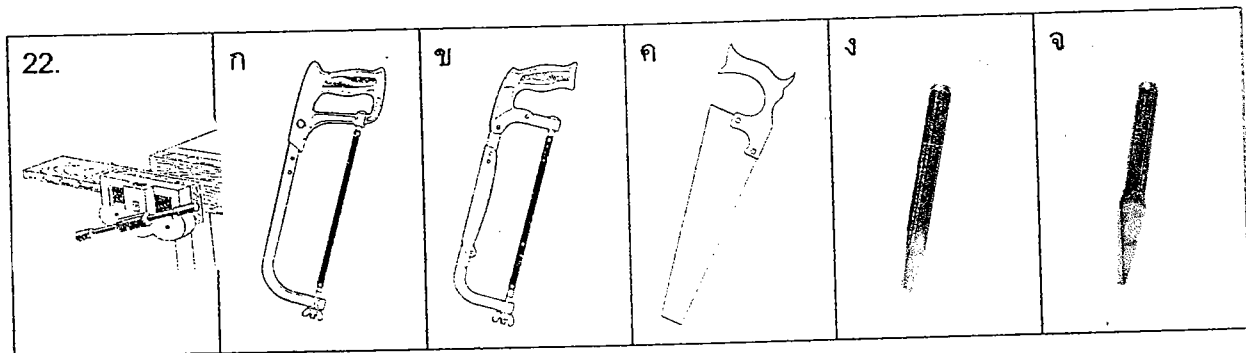
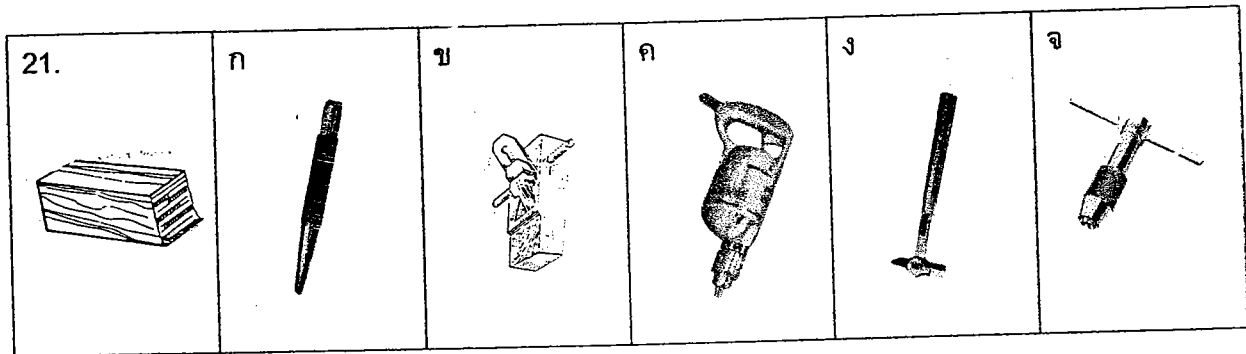
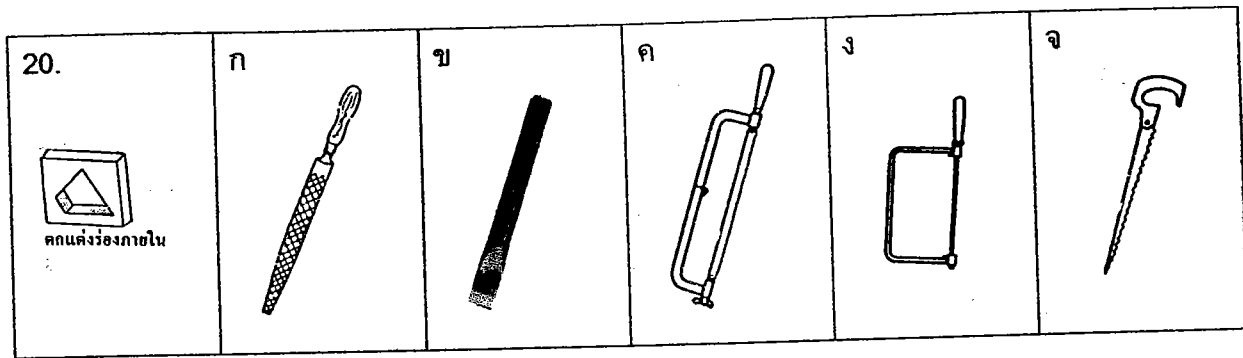
คำชี้แจง ในแต่ละข้อจะกำหนดรูปภาพมาให้ 1 รูป ทางซ้ายมือแล้วให้หาภาพจากตัวเลือก ก. ข. ค. ง. และ จ. ว่าข้อใดเป็นภาพที่มีความสัมพันธ์เกี่ยวข้องกันกับภาพที่กำหนดให้

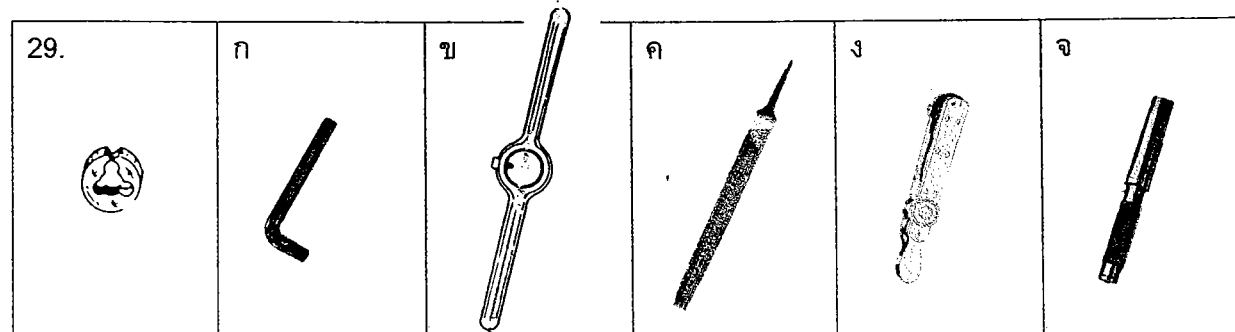
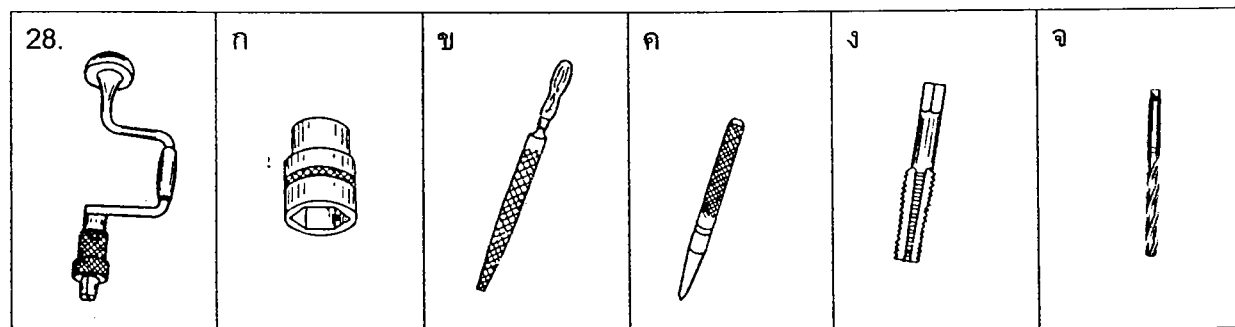
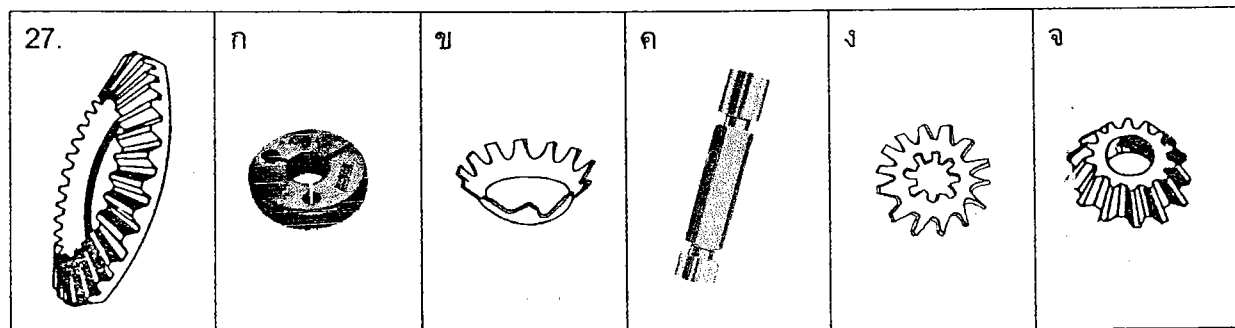
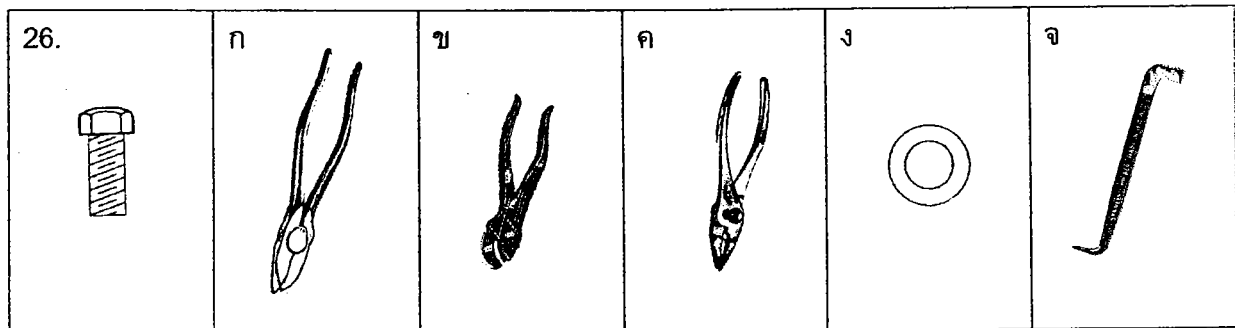
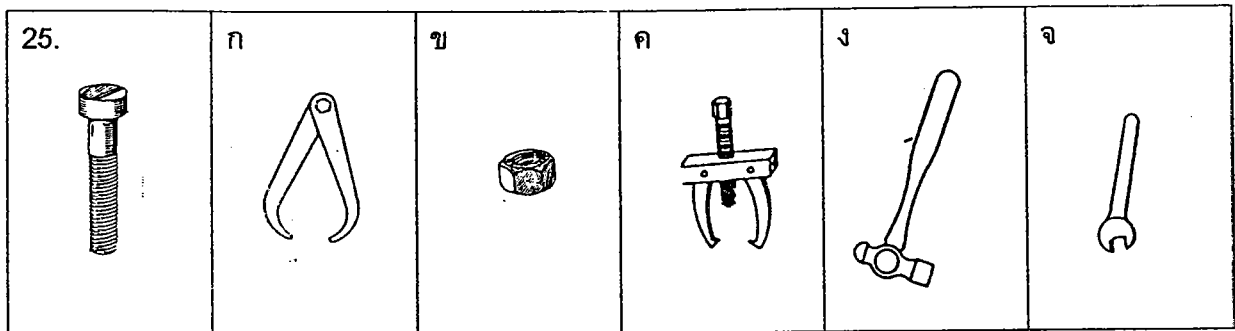


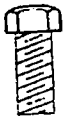
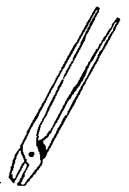










30.	ก	ข	ค	ง	จ
					

แบบทดสอบฉบับที่ 2

แบบทดสอบฉบับนี้มีทั้งหมด 6 ตอน ๆ ละ 5 ข้อ รวมทั้งหมด 30 ข้อ มีเวลาทำ 30 นาที

คำชี้แจง ใช้ภาพที่กำหนดให้ตอบคำถามโดยพิจารณาว่าข้อความแต่ละข้อมีความสัมพันธ์หรือเกี่ยวข้องกับภาพใดมากที่สุด

ตอนที่ 1 จงตอบคำถามข้อ 1 – 5

1. ใช้ตอกสกัดเพื่อขึ้นรูปงาน
2. ถ้าต้องการทำให้ผิวหน้าของปูนที่ฉาบเรียบ
3. ใช้สำหรับตะไบโลหะให้เรียบพอสมควร
4. ถ้าต้องการตัดโลหะแผ่นให้เป็นรูปร่าง ๆ
5. ใช้สำหรับขันหรือคลายสกรูที่มีหัวผ่าตา

ก.



ข.



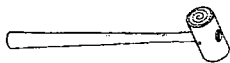
ค.



ง.



จ.



ฉ.



ช.



ซ.



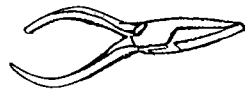
ตอนที่ 2 จงตอบคำถามข้อ 6 – 10

6. ใช้สำหรับขันหรือคลายสกรูที่มีหัวเป็นแฉก
7. ใช้สำหรับขันหรือคลายสกรูที่อยู่ในช่องหรือซอกที่มีหัวผ่าตรง
8. ใช้ตอกงัดกับงานที่มีความแข็งแรงน้อยเช่นใช้ถอนตะปู
9. ถ้าต้องการจับโลหะแบนหรือตัดสายไฟ
10. ใช้ในการขันเฉพาะสกรูและนอตหัวหกเหลี่ยมข้อดีคือใช้ในการขันที่ออกแรงมาก ๆ

ก.



ข.



ค.



ง.



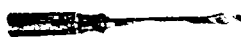
จ.



ฉ.



ช.



ซ.



ตอนที่ 3 จงตอบคำถามข้อ 11 – 15

11. ใช้ในการขันนอตปรับหรือนอตล็อกซึ่งนอตนี้จะมีรูด้านข้าง
12. ใช้ขันนอตขนาดเล็กและขนาดใหญ่โดยไม่ต้องใช้ประแจหลายตัวด้านข้างของนอตจะต้องทำร่องไว้
13. ถ้าต้องการตรวจสอบร่องได้ว่ารัศมี ใช้ตามที่ต้องการหรือไม่
14. ใช้งานคล้ายประแจปากตายแต่ดีกว่าตรงที่สามารถปรับใช้กับหัวสกรูได้หลายขนาด
15. ถ้าต้องการตรวจสอบความถูกต้องของมุมชิ้นงานเพื่อให้ได้งานมุมฉากตามต้องการ

ก.



ข.



ค.



ง.



จ.



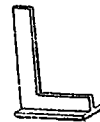
ฉ.



ช.



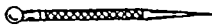
ซ.



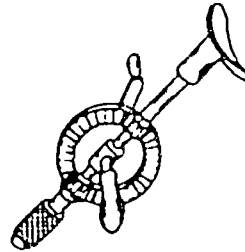
ตอนที่ 4 จงตอบคำถามข้อ 16 – 20

16. หมุนเจาะได้ด้วยแขนหมุน ที่มีการผ่อนแรงเจาะรูได้โครงไม่เกิน 5 ม.ม. ใช้ในการทำกรอบพระพลาสติก
17. เลื่อยที่ใช้สำหรับตัดโลหะแผ่น
18. ถ้าต้องการตัดไม้เป็นรูปต่าง ๆ ที่มีลายเอียงมาก
19. หมุดหัวฝังที่มีลักษณะย้าหัวหมุดเรียบเท่าความหนาของแผ่นงาน
20. ถ้าต้องการตกแต่งร่องภายใน

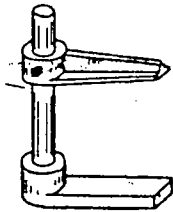
ก.



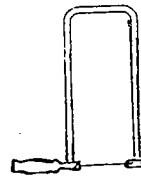
ข.



ค.



ง.



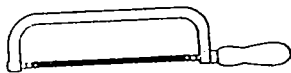
จ.



ฉ.



ช.



ซ.



ตอนที่ 5 จงตอบคำถามข้อ 21 – 25

21. ถ้าต้องการทำให้ผิวหน้าไม้ที่ขรุขระนั้นเรียบ
22. ต้องการตัดแผ่นไม้ออกเป็นชิ้น ๆ
23. ถ้าต้องการเคาะตัดโลหะเส้นแบนให้ตรง
24. ใช้สำหรับขันหรือคลายนอตและสลักเกลียวขนาดโต ๆ
25. ใช้ขันยึดสลักเกลียวกับชิ้นงานมีหัวหกเหลี่ยม

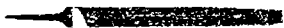
ก.



ข.



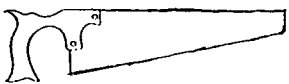
ค.



ง.



จ.



ฉ.



ช.



ซ.



ตอนที่ 6 จงตอบคำถามข้อ 26 – 30

26. ทำหน้าที่รองรับอยู่ระหว่างนอตกับชิ้นงานซึ่งมีพื้นอยู่ด้านนอก

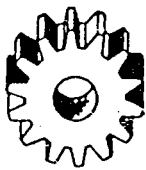
27. เป็นอุปกรณ์เพื่อส่งถ่ายกำลัง

28. ปากจับดอกสว่านชนิดใช้ประจำ

29. เมื่อต้องการคลายหรือขันนอตให้แน่น

30. เน้นเคื่องมือที่ทำเกี่ยวนวด

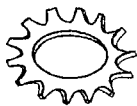
ก.



ข.



ค.



ง.



จ.



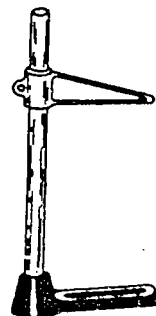
ฉ.



ช.



ซ.



แบบทดสอบฉบับที่ 3

แบบทดสอบฉบับนี้มีทั้งหมด 6 ตอน ๆ ละ 5 ข้อ รวมทั้งหมด 30 ข้อ มีเวลาทำ 30 นาที

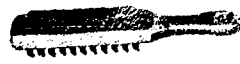
คำชี้แจง ให้นักเรียนดูภาพที่อยู่ทางซ้ายมือแล้วพิจารณาว่าแต่ละข้อสัมพันธ์กับข้อใดที่กำหนดไว้ทางขวามือ เลือกได้ข้อใดนำไปตอบลงในกระดาษคำตอบที่กำหนดให้

ตอนที่ 1 จงตอบคำถามข้อ 1 – 5 ต่อไปนี้

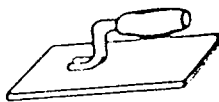
1.



ก.



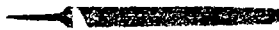
2.



ข.



3.



ค.



4.



ง.



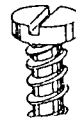
5.



จ.



ฉ.



ช.



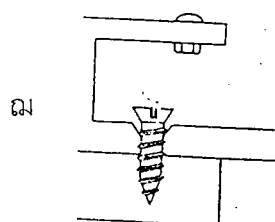
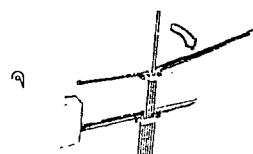
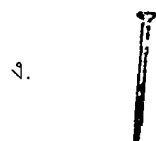
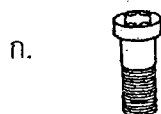
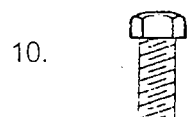
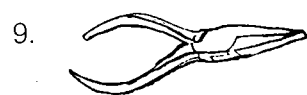
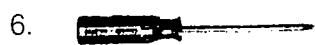
ซ.



ณ.



ตอนที่ 2 จงตอบคำถามข้อ 6 – 10 ต่อไปนี้



ตอนที่ 3 จงตอบคำถามข้อ 11 – 15 ต่อไปนี้

11.



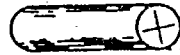
ก.



12.



ข.



13.



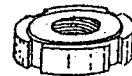
ค.



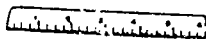
14.



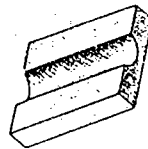
ง.



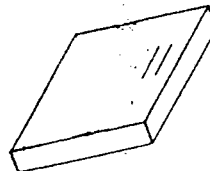
15.



จ.



ฉ.



ช.



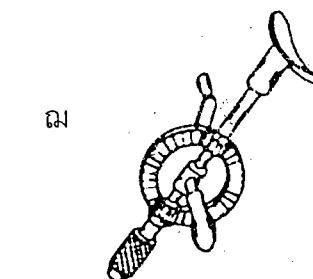
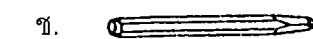
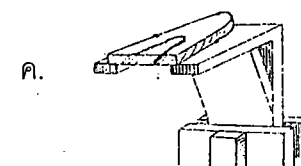
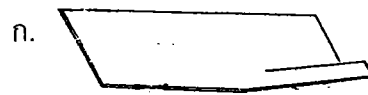
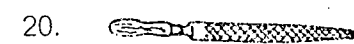
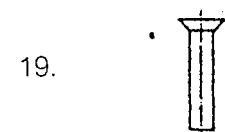
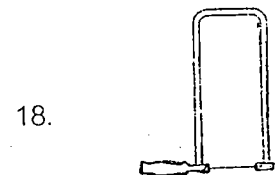
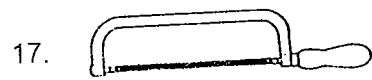
ฐ.



ฎ.



ตอนที่ 4 จงตอบคำถามข้อ 16 – 20 ต่อไปนี้



ตอนที่ 5 จงตอบคำถามข้อ 21 – 25 ต่อไปนี้

21.



ก.



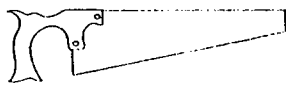
22.



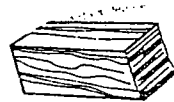
ข.



23.



ค.



24.



ง.



25.



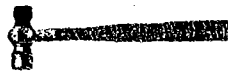
จ.



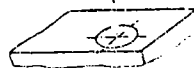
ฉ.



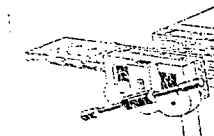
ช.



ซ.



ณ.



ตอนที่ 6 จงตอบคำถามข้อ 26 – 30 ต่อไปนี้

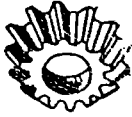
26.



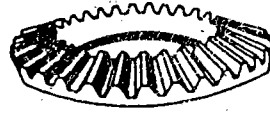
ก.



27.



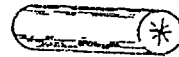
ข.



28.



ค.



29.



ง.



30.



จ.



ฉ.



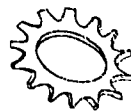
ช.



ซ.



ณ.



ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ นางสาวสุรี ใจกล้า

เกิดวันที่ 4 เดือนมกราคม

สถานที่เกิด

สถานที่อยู่ปัจจุบัน

พุทธศักราช 2514

อำเภอบางเลน จังหวัดนครปฐม

3/1 หมู่ 6 ตำบลบางระกำ

อำเภอบางเลน จังหวัดนครปฐม 73130

ตำแหน่งหน้าที่การงานปัจจุบัน

สถานที่ทำงานปัจจุบัน

นักวิชาการศึกษา 4

ฝ่ายการศึกษาอาชีพ กองส่งเสริมอาชีพ

สำนักพัฒนาชุมชน กรุงเทพมหานคร

44 อาคารศรีจุลทรัพย์ ชั้น 15 ถนนพระราม 1

แขวงรองเมือง เขตปทุมวัน

กรุงเทพมหานคร 10330

ประวัติการศึกษา

พ.ศ. 2526

ประถมศึกษา

โรงเรียนวัดสุขวัฒนาราม

พ.ศ. 2532

มัธยมศึกษาตอนต้น-ปลาย

โรงเรียนราชินี บุรีณะ

พ.ศ.2537

ครุศาสตรบัณฑิต สาขาวัดผลการศึกษา

สถาบันราชภัฏบ้านสมเด็จเจ้าพระยา

พ.ศ.2544

การศึกษามหาบัณฑิต สาขาวัดผลการศึกษา

มหาวิทยาลัยศรีนครินทรวิโรฒ