

---

# การทดสอบอิงปริเขต\*

---

สมหวัง พิธิยานุวัฒน์

---

## บทนำ

การทดสอบอิงปริเขต (Domain-Referenced Testing : DRT) นับว่าเป็นนวัตกรรมทางการทดสอบที่มีพัฒนาการมาได้ประมาณเกือบยี่สิบปี แต่ทว่ายังไม่เป็นที่ทราบกันกว้างขวาง โดยเฉพาะในประเทศไทย ในบทความนี้จึงมุ่งที่จะนำเสนอทิววิเคราะห์เกี่ยวกับสถานการณ์ที่เหมาะสมในการทดสอบอิงปริเขต แนวคิดพื้นฐานของการทดสอบอิงปริเขต ตลอดจนบทบาทของการทดสอบอิงปริเขต ในการพัฒนาแบบทดสอบอันเป็นเทคโนโลยีของการเขียนข้อสอบที่เรารู้จักในชื่อว่า “เทคนิครูปแบบข้อสอบ” ผู้เขียนหวังว่าบทความนี้จะเป็นตัวกระตุ้นให้มีการวิเคราะห์วิจารณ์เกี่ยวกับการทดสอบอิงปริเขตอย่างกว้างขวางต่อไป เพื่อก่อให้เกิดความรู้ความเข้าใจจนสามารถประเมินคุณค่าของเทคโนโลยีทางการทดสอบแนวนี้ได้ถูกต้องก่อนที่จะนำไปเผยแพร่และปฏิบัติจริง ผู้เขียนหวังว่าในอนาคตอันใกล้คงจะมีการพัฒนาโปรแกรมการทดสอบอิงปริเขตในระดับกลุ่มโรงเรียน และในระดับเขตการศึกษาชั้นในประเทศไทย อันจะช่วยให้การสอนอย่างเป็นระบบในประเทศไทยประสบความสำเร็จมากยิ่งขึ้น อันเป็นเป้าหมายสำคัญของการทดสอบที่มุ่งปรับปรุงพัฒนาและยกคุณภาพของการศึกษาโดยส่วนรวมนั่นเอง

ในบทความนี้ ได้นำเสนอเนื้อหาสาระของการทดสอบอิงปริเขตโดยแบ่งเป็น 3 หัวข้อคือ

---

\* ผู้เขียนขอขอบคุณคณะกรรมการโครงการพัฒนาแบบทดสอบที่กรุณาสนับสนุนให้ผู้เขียนได้เขียนบทความนี้ ขอขอบคุณอาจารย์ พิศิษฐ์ ตันทวนิช ที่ช่วยค้นหาเอกสารบางส่วนเพื่อ การเขียนบทความนี้ บทความฉบับนี้เป็นบทความที่ปรับปรุงจากการนำเสนอในการ สัมมนาระดับชาติ ครั้งที่ 4 เรื่อง เทคนิคการพัฒนาเครื่องมือวัดเพื่อระบบการติดตามปรับปรุงคุณภาพทางการศึกษาโรงแรมแกรนด์พาเลซ พัทยา 29 ตุลาคม -2 พฤศจิกายน 2527

ความจำเป็นและความต้องการการทดสอบอิงปริเขต หลักการและแนวคิดพื้นฐานของการทดสอบอิงปริเขต และเทคนิควิธีการพัฒนาข้อสอบอิงปริเขต

สำหรับรายละเอียดของเนื้อหาสาระในแต่ละหัวข้อมีดังนี้

### 1. ความจำเป็นและความต้องการการทดสอบอิงปริเขต

ในปัจจุบันนี้ทุกคนได้ตระหนักอย่างชัดเจนถึงข้อจำกัดของการใช้วัตถุประสงค์เชิงพฤติกรรม (Behavioral Objectives : BO) ในทางการศึกษาก็คือ การเรียน การสอน โดยเฉพาะการประเมินผลการศึกษาที่เน้นเฉพาะตามวัตถุประสงค์เชิงพฤติกรรมอย่างเคร่งครัด มักนำไปสู่ผลการประเมินที่ไม่ครอบคลุมสาระสำคัญของพฤติกรรมทุกระดับ และในทุกประเภทของสัมฤทธิ์ผลทางการเรียน การเรียนการสอนและประเมินผลการศึกษาที่ยึดตามวัตถุประสงค์เชิงพฤติกรรมอย่างเคร่งครัดมักจะประเมินเฉพาะพฤติกรรมการเรียนรู้ด้านพุทธิพิสัยหรือพุทธิปริเขต (Cognitive Domain) เท่านั้น นอกจากนี้นักการศึกษาทั้งหลายก็ได้ตระหนักถึงข้อจำกัดของการใช้ผลการทดสอบจากแบบสอบมาตรฐานซึ่งโดยทั่วไปก็เป็นแบบสอบอิงกลุ่ม (Norm-Referenced Tests) เพื่อการประเมินการสอนหรือช่วยให้นักเรียนแต่ละบุคคลศึกษาแล้วเรียนอย่างมีประสิทธิภาพมากขึ้น ถ้าประสงค์จะแก้ปัญหาในข้อจำกัดดังกล่าว และมุ่งพิจารณาประเด็นหลักในเรื่องของการทดสอบเพื่อการพัฒนาักเรียนแล้ว เราจำเป็นต้องพัฒนาระบบการทดสอบที่สามารถแสดงให้เห็นถึงความก้าวหน้าในการศึกษาแล้วเรียนในแต่ละเนื้อหาสาระหรือทักษะที่มีการระบุอย่างชัดเจน การทดสอบอิงปริเขตและการทดสอบอิงเกณฑ์ (Criterion-Referenced Testing) จึงเป็นระบบการทดสอบที่พัฒนาขึ้นเพื่อตอบสนองความต้องการดังกล่าว

เมื่อพิจารณาในแง่มุมมองของการสอนอย่างเป็นระบบที่เป็นที่นิยมกันในปัจจุบันพบว่า การติดตามกำกับเกี่ยวกับความก้าวหน้าในการเรียน ของนักเรียนอย่างใกล้ชิดและต่อเนื่องเป็น กุญแจดอกสำคัญที่จะทำให้การสอนอย่างเป็นระบบเกิดประสิทธิภาพตามที่มุ่งหวัง โดยทั่วไปแล้วเราใช้แบบสอบในหลายวัตถุประสงค์เช่น เพื่อกระตุ้นนักเรียนให้ศึกษาแล้วเรียนอย่างตั้งใจเพื่อให้ได้ระดับคะแนนที่แสดงถึงระดับผลสัมฤทธิ์ทางการเรียน เพื่อประเมินโปรแกรมการสอน เพื่อประเมินประสิทธิผลการสอน เพื่อจัดกลุ่มนักเรียน และเพื่อคัดเลือกนักเรียนสำหรับโปรแกรมการศึกษาพิเศษต่าง ๆ สำหรับในการสอนอย่างเป็นระบบนั้น มีจุดเน้นในการใช้แบบสอบที่แตกต่างออกไป กล่าวคือ มุ่งเน้นการใช้แบบสอบเพื่อรับรองผลการศึกษาที่ผ่านมาว่าอยู่ในระดับที่น่าพอใจหรือไม่ ถ้าอยู่ในระดับที่น่าพอใจให้ผ่านไป

เรียนต่อไปได้ แต่ถ้ายังไม่เป็นที่พอใจก็ต้องจัดการสอนซ่อมเสริมให้ด้วยวิธีที่เหมาะสมและมุ่งการทดสอบ ขณะการเรียนการสอนกำลังดำเนินอยู่ เพื่อตรวจสอบว่านักเรียนมีความรอบรู้ในมโนทัศน์ หรือทักษะ สำคัญหรือไม่ นักเรียนบรรลุวัตถุประสงค์ของการสอนหรือไม่ ความต้องการการทดสอบในการสอน อย่างเป็นระบบนี้ ย่อมต้องการการทดสอบที่แตกต่างไปจากเดิม จึงกล่าวได้ว่าการทดสอบอิงปริเขต และการทดสอบอิงเกณฑ์เป็นหัวใจของการสอนอย่างเป็นระบบทีเดียว

ในการใช้ประโยชน์จากผลการทดสอบนั้น จำเป็นต้องมีการแปลความหมายของคะแนนจาก แบบสอบ ในวงการนักวัดและประเมินผลการศึกษา เราทราบดีว่าการแปลความหมายของคะแนนจาก แบบสอบมี 3 ประเภท คือ

1. การแปลความหมายของคะแนนในเชิงอิงกลุ่มซึ่งเป็นการเปรียบเทียบคะแนนของนักเรียน คนใดคนหนึ่งกับคะแนนเฉลี่ยของกลุ่มที่กำหนดให้การแปลความหมายในเชิงนี้เป็นประโยชน์ในการจัด กลุ่มนักเรียน ให้ระดับเกียรตินิยมแก่นักเรียนเก่ง เลือกเด็กเข้าศึกษาในโปรแกรมพิเศษหรือประเมิน โปรแกรมการสอน สถิติที่นิยมใช้แปลงคะแนนเพื่อให้ความหมายในเชิงการเปรียบเทียบกับกลุ่ม หรือ การอิงกลุ่มได้แก่ เปอร์เซนไทล์ และคะแนนมาตรฐานในรูปต่างๆ เช่น Z-score และ T-score เป็นต้น

2. การแปลความหมายคะแนนในเชิงอิงเกณฑ์โดยการกำหนดมาตรฐานสมบูรณ์ (absolute standard) เช่น นักเรียนที่จะสอบผ่านต้องมีคะแนนไม่ต่ำกว่าร้อยละ 80 เป็นต้น นักเรียนที่ได้ คะแนนร้อยละ 55 และ 75 เหมือนกันในประเด็นที่ว่ามิผลสัมฤทธิ์ยังไม่ถึงระดับเกณฑ์มาตรฐาน อย่างไรก็ตามที่ระยะห่างของคะแนนที่สอบได้กับคะแนนที่เป็นเกณฑ์นั้นว่ามีความสำคัญมาก เช่นคนที่ สอบได้ 55% ควรต้องใช้เวลาศึกษาเพิ่มเติมมากกว่าคนที่สอบได้ 75% เป็นต้น Milliman, (1974) เสนอให้กำหนดเกณฑ์เป็นช่วงมากกว่าจะกำหนดเป็นจุด คนที่สอบได้คะแนนอยู่ในช่วงคะแนนเกณฑ์ หรืออยู่ในแดนสนธยา อาจต้องการเพียงการทบทวนโดยย่อก็สามารถทำให้บรรลุเกณฑ์ได้แล้วซึ่งแตกต่าง จากพวกที่ได้คะแนนต่ำกว่าช่วงคะแนนที่เป็นเกณฑ์ที่ต้องศึกษาทบทวนเพิ่มเติมอย่างมาก

3. การแปลความหมายของคะแนนในเชิงอิงปริเขต เป็นการแปลความหมายโดยใช้คะแนน ที่สอบได้เป็นตัวประมาณค่าของระดับผลสัมฤทธิ์ทางการเรียนในห้วงจักรวาลของความรู้และทักษะหรือใน ปริเขตที่กำหนด การแปลความหมายทำนองนี้เกิดขึ้นน้อยมาก แต่นับว่าเป็นสิ่งที่มีความสำคัญยิ่งต่อ การสอนอย่างเป็นระบบ และก็เป็นหัวใจของการทดสอบอิงปริเขต ในปัจจุบันการทดสอบอิงปริเขตมัก จะรู้จักในชื่อของการทดสอบอิงเกณฑ์ไป

## 2. หลักการและแนวคิดพื้นฐานของการทดสอบอิงปริเขต

### 2.1 ประวัติโดยสังเขปของการทดสอบอิงปริเขต

ในปี ค.ศ. 1962 Ebel ได้เขียนบทความเรื่อง Content Standard Test scores (ใน Mehrens & Ebel, 1976, 60-61) โดยให้ความหมายของโดเมนหรือปริเขต ในทิศทางที่ชัดเจน โดยเขียนไว้ว่า

"By the term content standard test score in this discussion we will mean a number that indicates the percent of a systematic sample from a defined domain of tasks which an individual has performed successfully."

บุคคลที่ได้รับการยอมรับว่าเป็นผู้บุกเบิกเรื่องของการทดสอบอิงปริเขตทั้งในเชิงนิยาม และระบบการพัฒนาแบบสอบตามแนวการทดสอบนี้ คือ Wells Hively แห่งมหาวิทยาลัยมิเนโซต้า ปัจจุบันเป็นผู้อำนวยการของ Spaulding Youth Center ที่ New Hampshire Hively ได้เริ่มโครงการ MINNEMAST ซึ่งเป็นโปรแกรมการทดสอบอิงปริเขตในวิชาวิทยาศาสตร์ และคณิตศาสตร์ ในปี ค.ศ. 1966 ได้ผลิตและเผยแพร่เอกสารในเรื่องนี้มากมาย บทความของ Hively ที่มีอิทธิพลมากในเรื่องนี้ก็คือ Domain-Referenced Testing ซึ่งตีพิมพ์ในวารสาร Educational Technology ในปี ค.ศ. 1974 ในวารสารฉบับนี้ทั้งฉบับเป็นเรื่องของการทดสอบอิงปริเขตทั้งสิ้น โดยมี Hively เป็นบรรณาธิการประจำฉบับ ในปีต่อมาคือ ปี ค.ศ. 1975 Hively ร่วมกับ Reynolds ได้เป็นบรรณาธิการตีพิมพ์หนังสือชื่อ "Domain-Referenced Testing in Special Education" ผลงานที่สำคัญของ Hively คือการเสนอแนะระบบการทดสอบใหม่คือการทดสอบอิงปริเขต ทำให้มีโปรแกรมการทดสอบอิงปริเขตอย่างกว้างขวางในมลรัฐมิเนโซต้า ดูเหมือนปัจจุบันแม้จะไม่มีใครได้ยืมการทดสอบอิงปริเขตมากนัก เพราะส่วนใหญ่ถือว่า การทดสอบอิงปริเขตเป็นส่วนหนึ่งของการทดสอบอิงเกณฑ์ไปทั้ง ๆ Hively ได้เขียนไว้อย่างชัดเจนว่าการทดสอบอิงเกณฑ์มีมโนทัศน์ของการเรียนเพื่อรอบรู้ (mastery learning) มาปะปนจนแยกกันไม่ออก ในขณะที่การทดสอบอิงปริเขตมิใช่ได้เฉพาะการเรียนเพื่อรอบรู้เท่านั้น อีกทั้งการทดสอบอิงปริเขตมีคุณลักษณะที่ชัดเจนมีสถิติสนับสนุนอย่างเข้มแข็งอย่างไรก็ดี ในปัจจุบันเราจะได้อินเสมอในเรื่องเทคนิคกำหนดรูปแบบข้อสอบ (Item Form) ซึ่งเป็นเทคโนโลยีทางการเขียนข้อสอบอย่างหนึ่งที่น่าไปใช้ในการทดสอบอิงเกณฑ์ ผู้ที่คิดค้นเทคนิคกำหนดรูปแบบข้อสอบก็คือ Wells Hively

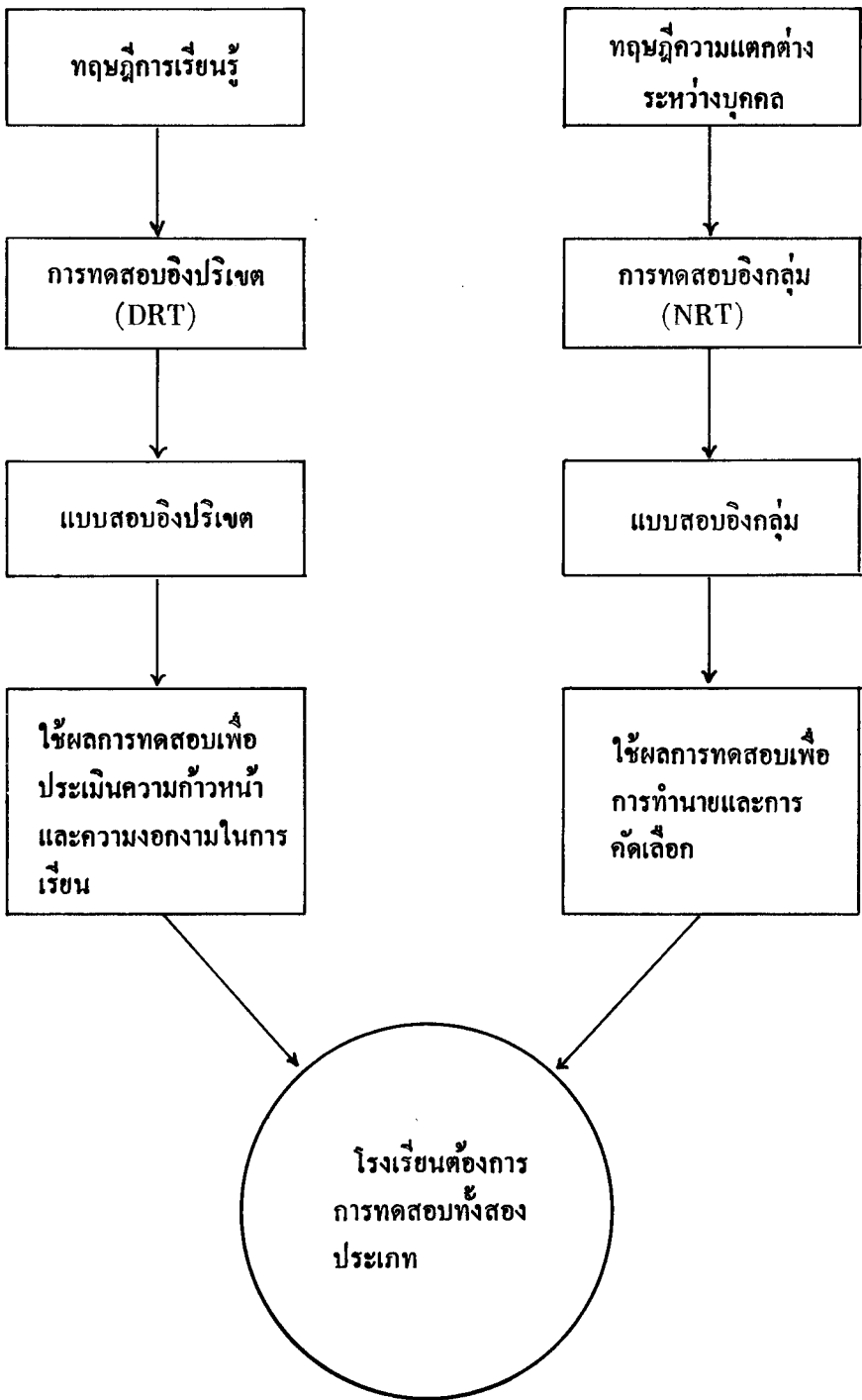
โดยคิดขึ้นในกรอบความคิดของการทดสอบอิงปริเขต Roid และ Haladyna (1982) ได้วิจารณ์ว่า ในการกำหนดคะแนนเกณฑ์ในการทดสอบอิงเกณฑ์เป็นสิ่งที่ปราศจากความหมาย ถ้าไม่มีการกำหนดการแปลความหมายของคะแนนในเชิงอิงปริเขต โดยนัยแล้วการทดสอบอิงเกณฑ์ที่นิยมกันอย่างกว้างขวางในปัจจุบันควรจะเปลี่ยนชื่อเป็นการทดสอบอิงปริเขตมากกว่า อย่างไรก็ตามความเป็นไปได้คงจะต่ำเพราะการทดสอบอิงเกณฑ์ได้มีประวัติศาสตร์อันยาวนานและหยั่งรากลึกอย่างฝังแน่นเสียแล้วในวงการทดสอบนี้ทั้งของประเทศไทยและต่างประเทศโดยเฉพาะประเทศสหรัฐอเมริกา

## 2.2 ความแตกต่างระหว่างการทดสอบอิงปริเขตกับการทดสอบอิงกลุ่ม

ดังที่ได้กล่าวมาแล้วว่าในการประเมินผลสัมฤทธิ์ทางการเรียนไม่ว่าในด้านพุทธิปริเขต หรือด้านอื่น ๆ ก็ตาม ครูมีความสนใจว่านักเรียนมีความรู้ความสามารถทำอะไรได้บ้าง ทั้งในสถานการณ์ในโรงเรียน และในสถานการณ์ชีวิตจริง อะไรเป็นสิ่งที่นักเรียนยังไม่มีความรู้ความสามารถ ครูสนใจข่าวสารเช่นนั้นมากกว่าสนใจว่าใครเก่งกว่าใคร กล่าวคือข่าวสารทำนองนี้ ซึ่งได้รับจากการทดสอบด้วยแบบสอบอิงปริเขตจะช่วยให้ทั้งครูและนักเรียนปรับปรุงและยกระดับคุณภาพของการเรียนการสอน ส่วนข่าวสารว่าใครเก่งกว่าใครนั้นก็จะเป็นข่าวสารที่มีประโยชน์ในเชิงการตัดสินใจรวมสรุปทั้งหลายจะได้การทดสอบด้วยแบบสอบอิงกลุ่ม

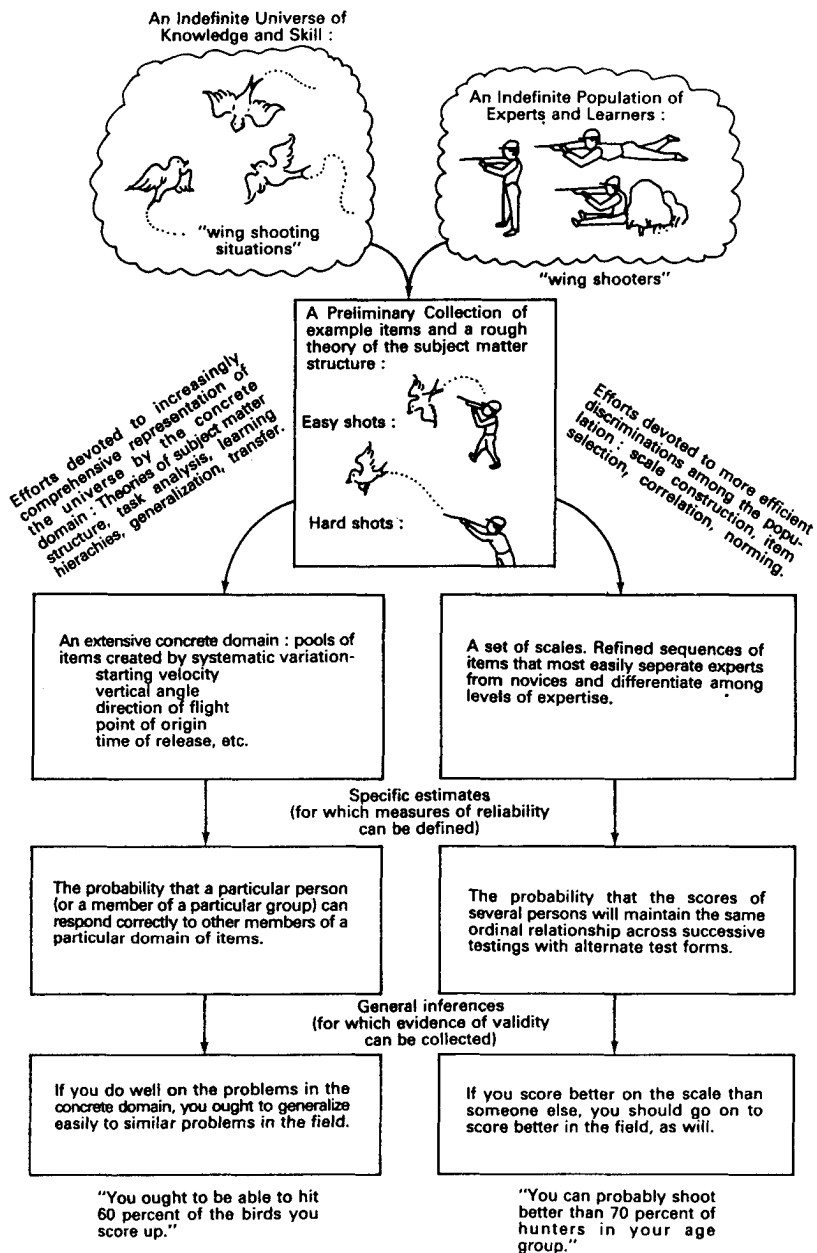
เมื่อวิเคราะห์ถึงความแตกต่างของการทดสอบอิงปริเขตกับการทดสอบอิงกลุ่มจะพบความแตกต่างที่ชัดเจนมาก กล่าวคือทฤษฎีเบื้องหลังของการทดสอบอิงปริเขต ทฤษฎีการเรียนรู้ ในขณะที่ทฤษฎีเกี่ยวกับความแตกต่างระหว่างบุคคลเป็นทฤษฎีที่อยู่เบื้องหลังของการทดสอบอิงกลุ่ม นอกจากนี้ผลการทดสอบอิงปริเขตมุ่งใช้เพื่อประเมินความก้าวหน้าในการเรียน ตลอดจนความอ่องงามในการเรียนในขณะที่ผลการทดสอบจากการทดสอบอิงกลุ่มมุ่งใช้ผลเพื่อการทำนายและการคัดเลือกเป็นสำคัญ อย่างไรก็ตามโรงเรียนมีความต้องการโปรแกรมการทดสอบทั้งสองประเภท ปัญหามีอยู่เพียงว่าจะใช้โปรแกรมการทดสอบนั้นได้อย่างถูกต้องเหมาะสมหรือไม่ โดยสรุปความแตกต่างในระดับมหภาคของการทดสอบอิงปริเขต และการทดสอบอิงกลุ่ม แสดงดังแผนภูมิที่ 1

แผนภูมิที่ 1 ความแตกต่างบางประการระหว่างการทดสอบอิงปริเขต (DRT) กับการทดสอบอิงกลุ่ม (NRT)



สำหรับความแตกต่างระหว่างการทดสอบอิงปริเซตและการทดสอบอิงกลุ่มในเชิงการพัฒนา  
แบบสอบแสดงในแผนภูมิที่ 2

แผนภูมิที่ 2 ความแตกต่างในการพัฒนาแบบสอบใน DRT และ NRT (W. Hively 1972, p.7)



จากตัวอย่างในแผนภูมิที่ 2 ซึ่งคัดลอกมาจากบทความของ Wells Hively ที่แสดงถึงความคล้ายคลึงและความแตกต่างของการพัฒนาแบบสอบอิงปริเขตและสอบอิงกลุ่ม ตลอดจนวิธีตรวจสอบคุณสมบัติในเชิงคุณภาพของแบบสอบทั้งสองประเภท โดยสรุปการสร้างแบบสอบทั้งสองแนวเริ่มต้นด้วยวัตถุประสงค์เหมือนกันคือ คำอธิบายเกี่ยวกับจักรวาลของความรู้ที่ค่อนข้างจะคลุมเครือ และค่านิยมคร่าว ๆ เกี่ยวกับประชากรของผู้เชี่ยวชาญกับผู้เรียน การเก็บข้อมูลเบื้องต้นเพื่อให้ได้มาซึ่งทฤษฎีหรือโครงสร้างคร่าว ๆ ของเนื้อหาสาระ ตลอดจนกำหนดหัวข้อตามโครงสร้างของเนื้อหาสาระที่กำหนด จากจุดนี้ไปการพัฒนาแบบสอบจะแตกต่างกันอย่างชัดเจน กล่าวคือ นักพัฒนาแบบสอบอิงปริเขตก็จะทุ่มความพยายามเพื่อกำหนดปริเขตของเนื้อหาสาระให้มีความชัดเจน และมีความเป็นตัวแทนของจักรวาลของความรู้ให้มากที่สุด โดยอาศัยทฤษฎี เช่น โครงสร้างของเนื้อหาวิชา การวิเคราะห์งาน ทฤษฎีลำดับขั้นของการเรียนรู้ ทฤษฎีการถ่ายโยงความรู้ และทฤษฎีว่าด้วยความเป็นทั่วไป (generalization) ส่วนนักพัฒนาแบบสอบอิงกลุ่มจะทุ่มความพยายามไปในการจำแนกระหว่างผู้เรียนเป็นสำคัญโดยการกำหนดการสร้างมาตรฐาน การเลือกข้อสอบ ศึกษาสหพันธ์ และการกำหนดปกติวิสัยในการแปลความหมายคะแนน

## 2.8 นิยามสำคัญในการทดสอบอิงปริเขต

2.3.1 นิยามของ “ข้อสอบ” ข้อสอบในการวัดอิงปริเขต หมายถึงเซตของคำชี้แจงเพื่อที่ทำให้พฤติกรรมเกิดขึ้น สืบสอบและวัดพฤติกรรมที่เกิดขึ้น ข้อสอบตามนิยามนี้ประกอบด้วยคำชี้แจงเกี่ยวกับการนำเสนอสิ่งเร้า การบันทึก การตอบสนองและการตัดสินว่าการตอบสนองหรือคำถามนั้นเหมาะสมหรือไม่ (W. Hively, 1974 p.6) เป็นที่น่าสังเกตว่าคำว่า “ข้อสอบ” ในที่นี้มีได้หมายเฉพาะข้อสอบข้อเขียนเท่านั้น แต่หมายถึงตัววัดพฤติกรรมการแสดงออกของมนุษย์ในทุกรูปแบบ

2.3.2 นิยามของปริเขต ปริเขตหรือโดเมน ประกอบด้วยเซตระบุอย่างชัดเจนของข้อสอบต่าง ๆ (W. Hively, 1974, p.8)

2.3.3 นิยามของรูปแบบข้อสอบ หมายถึงรายการของกฎเกณฑ์เพื่อทำให้เกิดเซตของข้อสอบที่สัมพันธ์กัน (W. Hively, 1974, p. 8)

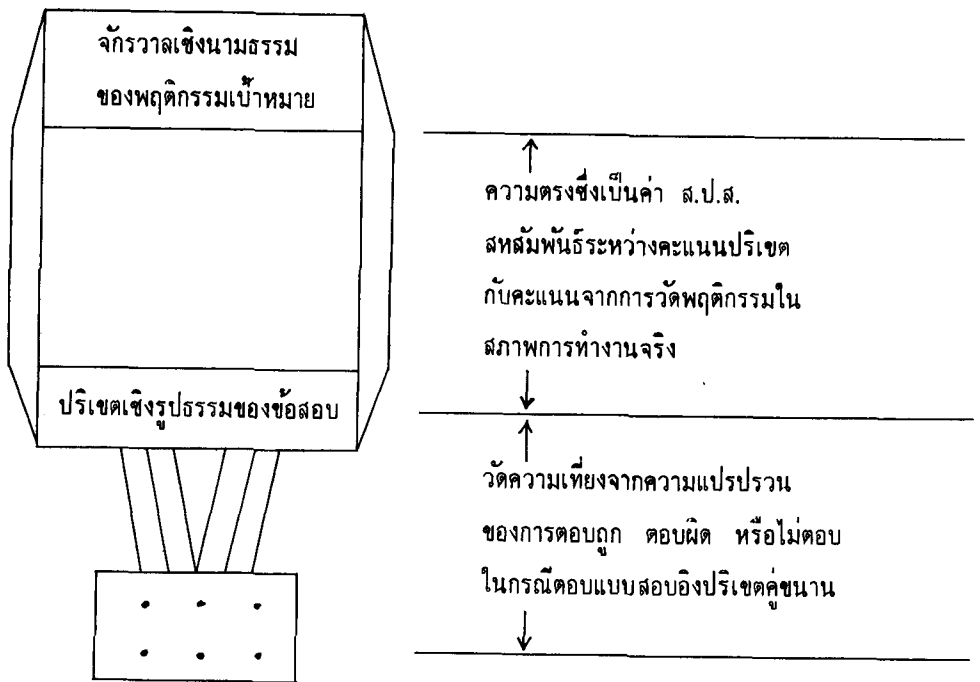
2.3.4 ความเที่ยงของแบบสอบอิงปริเขตเกี่ยวข้องกับความต้องการที่บุคคลใดบุคคลหนึ่งประมาณค่าความน่าจะเป็นที่จะตอบข้อสอบภายในปริเขตที่กำหนดให้ได้อย่างถูกต้อง (W. Hively, 1974, p.8) ซึ่งคิดคำนวณได้จากสหสัมพันธ์ระหว่างคะแนนจากแบบสอบอิงปริเขตคู่ขนาน



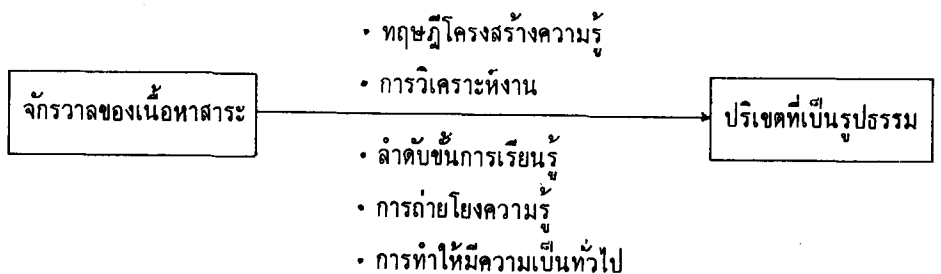
2.3.5 ความตรงของแบบสอบอิงปริเขตเกี่ยวข้องกับความสำเร็จในการสรุปความเป็นสากลของพฤติกรรมในปริเขตกับพฤติกรรมในจักรวาลของความรู้ซึ่งเป็นประชากรของปริเขต (W. Hively, 1974, p.8) ซึ่งเรารู้จักกันดีในเชิงความตรงเชิงเกณฑ์สัมพัทธ์

นิยามของความเที่ยงและความตรงแสดงในแผนภูมิที่ 3

แผนภูมิที่ 3 แสดงนิยามของความเที่ยงและความตรงของแบบสอบอิงปริเขต



ตัวอย่างข้อสอบ  
 ที่นำเสนอต่อผู้รับการทดสอบ  
 คนใดคนหนึ่ง ณ ระยะเวลาใด เวลาหนึ่ง



### 3. เทคนิควิธีการพัฒนาแบบสอบอิงปริเขต

ในแผนภูมิที่ 3 นอกจากแสดงถึงนิยามของความเที่ยง และความตรงของแบบสอบอิงปริเขตแล้ว ยังแสดงถึงขั้นตอนสำคัญของการพัฒนาแบบสอบอิงปริเขตด้วย โดยทั่วไปในการพัฒนาแบบสอบอิงปริเขตมีขั้นตอนที่สำคัญ คือ

1. การกำหนดวัตถุประสงค์ของการสอน
2. กำหนดปริเขตที่ชัดเจน ถ้าปริเขตกว้างเกินไป ก็ควรกำหนดเป็นปริเขตย่อย เพื่อให้ปริเขตมีความเป็นตัวแทนของจักรวาลความรู้
3. พัฒนาข้อสอบโดยกำหนดรูปแบบข้อสอบ
4. บรรณาธิกรณข้อสอบทั้งในเชิงหลักเหตุผล และหลักประจักษ์
5. พัฒนาแบบสอบทั้งฉบับพร้อมนำไปทดสอบหาคุณภาพ และจัดทำให้อยู่ในรูปมาตรฐานเพื่อนำไปใช้ต่อไป

ในการพัฒนาแบบสอบอิงปริเขตความสำคัญเน้นไปที่การนิยามจักรวาลของเนื้อหาสาระกับการพัฒนาข้อสอบเป็นสำคัญ ในบทความนี้จึงขอเน้นเฉพาะสองเรื่องนี้

อันที่จริงการกำหนดจักรวาลของเนื้อหาสาระเป็นที่สนใจของการพัฒนาแบบสอบทุกประเภท ดังที่ R. Tyler (1950) ได้สร้างผังการสร้างข้อสอบ หรือตารางการวิเคราะห์หลักสูตร ซึ่งเป็นตารางสองทางซึ่งแสดงมิติของเนื้อหาสาระระดับพฤติกรรมหรือระดับผลสัมฤทธิ์ที่มุ่งหวังตั้งระบุไว้ในวัตถุประสงค์ของการสอนนั่นเอง วิธีการนี้ยังเป็นที่นิยมในการพัฒนาแบบสอบอิงกลุ่ม ต่อมาในปี ค.ศ. 1969 Lord และ Novick ได้คิดโมเดลลักษณะแฝง (latent trait model) ขึ้นก็เพื่อกำหนด a latent achievement continuum ที่อยู่ภายใต้เซตของข้อสอบต่างๆ ที่สร้างขึ้น วิธีการนิยามจักรวาลของเนื้อหาสาระเช่นนี้เป็นรากฐานสำคัญของ Item Response Theory ในปีเดียวกัน Hively, Patterson และ Page (1968) และ Osburn (1968) ได้นิยามจักรวาลของความรู้ที่มีเอกลักษณ์พิเศษโดยถือว่า เป็นกลุ่มของข้อสอบที่เป็นไปได้ทั้งหมดที่เป็นตัวแทนของปริเขตการสอนที่แสดงในรูปของเนื้อหาสาระ

Johnson, T.J. (1974, pp. 43--48) ได้อธิบายถึงประเภทของปริเขตว่ามี 5 ประเภทด้วยกัน คือ

1. ปริเขตของวัตถุประสงค์ทางการศึกษา เช่น พุทธิปริเขต เป็นต้น
2. ปริเขตของสารสนเทศ เช่น เซทของความรู้ เป็นต้น

3. ประเภทของเทคโนโลยี เช่น เศรษฐของผลิตผลทางการศึกษา เป็นต้น
4. ประเภทของการสอน เช่น เศรษฐของกิจกรรมทางการศึกษา เป็นต้น
5. ประเภทของการปฏิบัติ หรือส่วนประกอบการ (Performance) เช่น เศรษฐของข้อสอบ เป็นต้น

สำหรับความสัมพันธ์ระหว่างประเภท และสารสนเทศที่ได้แสดงในแผนภูมิที่ 4

แผนภูมิที่ 4 ความสัมพันธ์ของประเภทกับสารสนเทศ

ประเภท	หน้าที่ของสารสนเทศที่ได้
1. วัตถุประสงค์	1. เราต้องการทำอะไร
2. ความรู้	2. เรารู้อะไร
3. ผลิตผล	3. สิ่งที่เราสามารถจะทำได้ หรือสำเร็จให้เห็นได้โดยใช้สิ่งที่มียอยู่แล้ว
4. การสอน	4. สิ่งที่เราทำจริงด้วยสิ่งที่มียอยู่
5. ข้อสอบ	5. สิ่งที่เราวัด

ในการกำหนดประเภทสามารถใช้วิธีอุปนัย (induction) ตามวิธีของ Hively หรือ วิธีของ Baker ซึ่งเป็นแบบนิรนัยโดยเริ่มจากวัตถุประสงค์เชิงพฤติกรรม (BO) แล้วสร้างโตเมนหรือประเภทมาแทนกลุ่ม BO ส่วนวิธีของ Hively เป็นการกำหนดข้อสอบหรืองานต่างๆ แล้ว จึงจัดกลุ่มข้อสอบเป็นประเภท สิ่งที่สำคัญคือ ความเป็นตัวแทนของประเภท

ในการกำหนดประเภทซึ่งมีความเป็นตัวแทนที่ดีและวิธีการพัฒนาข้อสอบด้วยนั้น Hively ได้สร้างเทคนิครูปแบบข้อสอบ (Item Form) ขึ้น รูปแบบข้อสอบก็คือ กฎเกณฑ์ในการก่อให้เกิด เศรษฐของข้อสอบที่เท่าเทียมกันซึ่งล้วนเป็นตัวแทนที่ดีของกลุ่มมโนทัศน์ และทักษะที่เกี่ยวข้องสัมพันธ์กัน กลุ่มตัวอย่างของข้อสอบก็จะเป็นแบบสอบอิงประเภท ซึ่งสามารถนำไปใช้ประเมินความรู้และสัมฤทธิ์ผล

ของนักเรียนได้ เพื่อให้คะแนนจากแบบสอบอิงปริเขต มีคุณลักษณะของตัวคาดคะเนระดับผลสัมฤทธิ์ทางการเรียนที่ปราศจากอคติ และเป็นตัวคาดคะเนที่พอเพียงที่จะให้ข้อมูลเพื่อลงข้อสรุปเกี่ยวกับระดับผลสัมฤทธิ์ของการเรียนนั้น ทั้งนี้แบบสอบอิงปริเขตจะต้องเป็นไปตามเงื่อนไข 2 ประการคือ ต้องมีการระบุเขตของข้อสอบที่เป็นไปได้ภายในปริเขตของความรู้อะพสุติกรรมไว้ล่วงหน้า และกลุ่มตัวอย่างข้อสอบในแบบสอบอิงปริเขตต้องเป็นตัวอย่างแบบสุ่มของข้อสอบทั้งหมดในปริเขตนั้น

ตัวอย่างรูปแบบข้อสอบที่ง่ายที่สุดโดยพัฒนาจากข้อสอบที่มีอยู่แล้ว มีดังนี้ (Hivaly, et. al. 1973)

### ข้อสอบที่มีอยู่แล้ว

กลุ่มตัวอย่างแบบสุ่มซึ่งเป็นน้ำหนักของรถบรรทุก 100 คันที่แล่นผ่านมายังจุดตรวจข้างหน้าหน้า มีค่าเฉลี่ยเท่ากับ 40,250 ปอนด์ และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2500 จงหาช่วงความเชื่อมั่นร้อยละ 95 สำหรับน้ำหนักที่แท้จริงของรถบรรทุกที่แล่นผ่านมายังจุดตรวจข้างหน้าหน้า โดยให้ถูกต้องอย่างน้อยทศนิยมหนึ่งตำแหน่ง

### รูปแบบข้อสอบอย่างง่าย

กำหนดให้มีกลุ่มตัวอย่างแบบสุ่มของ (N) (สิ่ง) ซึ่งมีค่าเฉลี่ยของ (มิติที่มุ่งวัด) เท่ากับ (M) และมีส่วนเบี่ยงเบนมาตรฐานเป็น (SD) จงหาช่วงความเชื่อมั่นที่ (95%, 99%) ของค่าเฉลี่ยที่แท้จริงของ (มิติที่มุ่งวัด) ของ (สิ่งใดสิ่งหนึ่ง) โดยคิดคำนวณให้ถูกต้องอย่างน้อย ทศนิยมหนึ่งตำแหน่ง

### เซทของการแทนค่า

N เป็นจำนวนของสิ่งหนึ่งเป็นเลขจำนวนเต็มมากกว่า 30 (เราต้องการคาดคะเนช่วงความเชื่อมั่นของกลุ่มตัวอย่าง ขนาดใหญ่, Freud, 1960, p.217)

ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานเป็นเลขจำนวนเต็มหรือทศนิยมก็ได้ (ให้ถูกต้องอย่างน้อย 1 ตำแหน่ง)

สิ่งของหรือมิติที่มุ่งวัดสามารถเป็นอะไรก็ได้ เช่น น้ำหนัก ความสูง หรือคุณลักษณะอื่น ๆ

### การให้คะแนน

คาดว่านัสนิตจะใช้สูตรที่จะได้ช่วงความเชื่อมั่น เท่ากับ  $\bar{X} \pm (1.96 \text{ หรือ } 2.58) \frac{SD}{\sqrt{N}}$   
 อีกตัวอย่างหนึ่งของรูปแบบข้อสอบแสดงในแผนภูมิที่ 5

แผนภูมิที่ 5 แสดงรูปแบบข้อสอบ (W. Hively, 1974, p.8)

#### ITEM FORM SHELL

<p>DIRECTIONS :</p> <p>Read Script to Child :</p> <p>.....</p> <p>Write down Child's exact word.</p>	<p>SCRIPT :</p> <p>Tell me a number that is (a) .....</p>
--	---

#### CELL MATRIX

	CELL NUMBER			
	1	2	3	4
Script (a)	"greater than (b <sub>1</sub> )"	"less than (b <sub>1</sub> )"	"greater than (b <sub>1</sub> ) but less than (b <sub>2</sub> )"	
Numberals (b)	$0 \leq b_1 \leq 19$	$1 \leq b_1 \leq 20$	$0 \leq b_1 \leq 18$ and $(b_1 + 2) \leq b_2 \leq 20$	$0 \leq b_1 \leq 19$ and $b_2 = b_1 + 1$

การวิเคราะห์คุณภาพของแบบสอบอิงปริเขตดำเนินการได้ดังนี้

#### 1. การวิเคราะห์รายข้อ

##### 1.1 คำนวณหาดัชนีความง่าย

ดัชนีความง่าย = สัดส่วนของนักเรียนที่ตอบถูกในแต่ละข้อ (หลังจากที่ได้รับการสอนแล้ว)

##### 1.2 คำนวณหาดัชนีความไว (Sensitivity Index)

$$S_i = \frac{R_{\text{pos}_i} - R_{\text{pre}_i}}{T}$$

เมื่อ  $S_i$  = ดัชนีความไวของข้อสอบข้อที่  $i$

$R_{\text{pos}_i}$  = จำนวนผู้ตอบข้อสอบข้อ  $i$  ถูกหลังเรียน

$R_{\text{pre}_i}$  = จำนวนผู้ตอบข้อสอบข้อ  $i$  ถูกก่อนเรียน

$T$  = จำนวนผู้สอบทั้งหมด

ค่าของ  $S_i$  จะอยู่ระหว่าง  $-1$  และ  $+1$  ข้อสอบที่ดีควรมีค่า  $S_i$  เป็น  $+$  ยิ่งใกล้  $1$  ยิ่งเป็นข้อสอบที่มีคุณภาพดี

## 2. การวิเคราะห์คุณภาพทงฉบับ

### 2.1 คำนวณหาสัมประสิทธิ์ความตรงของแบบสอบอิงเกณฑ์ โดยสูตรของ

$$\text{ส.ป.ส. ความตรง} = \frac{a + c}{a + b + c + d}$$

เมื่อ  $a$  = จำนวนคนที่ผ่านเกณฑ์หลังจากได้รับการสอน

$b$  = จำนวนคนที่ผ่านเกณฑ์เมื่อยังไม่ได้รับการสอน

$c$  = จำนวนคนที่ไม่ผ่านเกณฑ์เมื่อยังไม่ได้รับการสอน

$d$  = จำนวนคนที่ไม่ผ่านเกณฑ์ แม้ได้รับการสอนแล้ว

ส.ป.ส. ความตรง จะมีค่าระหว่าง  $0$  กับ  $1$  ยังมีค่าใกล้  $1$  แสดงว่าเป็นแบบสอบที่วัดความตรงตามที่ต้องการมากที่สุด

### 2.2 คำนวณหาความเที่ยงของแบบทดสอบโดยสูตรของ SWAMINATHAN, HAMBLETON และ ALGINA

$$K = \frac{P_o - P_e}{1 - P_e}$$

เมื่อ  $K$  = ส.ป.ส. ความเที่ยงของแบบสอบ

$P_o$  = สัดส่วนที่เกิดขึ้น

$P_e$  = สัดส่วนที่คาดหวัง

ตัวอย่าง ในการทดสอบ 2 ครั้ง หลังจากนักเรียนได้รับการสอนแล้ว โดยที่ช่วงห่างของการทดสอบ 2 สัปดาห์ ปรากฏสัดส่วนของนักเรียนที่สอบผ่านตามเกณฑ์ (80%) ดังแสดงในตารางต่อไปนี้

สอบครั้งที่ 2 \n สอบครั้งที่ 1	สอบครั้งที่ 1		รวม
	ผ่าน	ไม่ผ่าน	
ผ่าน	.66	.14	.80
ไม่ผ่าน	.04	.16	.20
รวม	.70	.30	1.00

จงคำนวณหาความเที่ยงของแบบสอบฉบับนี้  
ในกรณีนี้

$$P_o = .66 + .16 = .82$$

$$P_e = (.70)(.80) + (.30)(.20) = .62$$

$$K = \frac{.82 - .62}{1 - .62} = .53$$

ซึ่งนับว่าเป็นแบบสอบที่มีความเที่ยงค่อนข้างต่ำ

สำหรับการประมาณค่าสัมประสิทธิ์ความเที่ยงโดยการทดสอบเพียงครั้งเดียวสามารถคำนวณได้โดยสูตร signal/noise ratios ที่เสนอโดย R. Brennan and M.Kane (1977)

เมื่อได้พัฒนาแบบสอบอิงปริเขตจนมีคุณภาพแล้ว เราก็สามารถนำไปทดสอบนักเรียนคะแนนที่ได้ จากแบบสอบอิงปริเขตสามารถประมาณค่าโดยสูตรต่อไปนี้

$$\begin{aligned} \hat{P} &= \frac{N_1}{N} P_1 + \frac{N_2}{N} P_2 + \dots + \frac{N_k}{N} P_k \\ &= \sum_{i=1}^k \frac{N_i}{N} P_i \end{aligned}$$

เมื่อ  $\hat{P}$  = ค่าประมาณของคะแนนปริเขต

$N_i$  = จำนวนข้อสอบในปริเขตย่อยที่  $i$

$N = \sum N_i =$  จำนวนข้อสอบทั้งหมดในปริเขต

$P_i =$  % ของข้อสอบในปริเขตย่อยที่  $i$  ที่ผู้สอบสอบผ่าน

ตัวอย่างเช่น ในปริเขตการบวกเลขแบ่งออกเป็นปริเขตย่อย สิบเซท คือ ปริเขตย่อยที่ 1 มี 25 ข้อ และปริเขตย่อยที่ 2 มี 75 ข้อ ได้จัดทำแบบสอบชุดหนึ่ง จำนวน 40 ข้อ โดยสุ่ม 10 ข้อ จากปริเขตย่อยที่หนึ่ง และสุ่มจำนวน 30 ข้อ จากปริเขตย่อยที่สอง ปรากฏว่า เด็กชาย ก. ตอบข้อสอบในปริเขตย่อยที่หนึ่งได้ถูกต้องทุกข้อและตอบข้อสอบในปริเขตย่อยที่สองถูก 25 ข้อ จึงประมาณค่าคะแนนปริเขตการบวกของเด็กชาย ก.

$$\text{จากสูตร } \hat{P} = \frac{N_1}{N} P_1 + \frac{N_2}{N} P_2 + \dots + \frac{N_k}{N} P_k$$

$$\hat{P} = \frac{25}{100} (100) + \frac{75}{100} (83.3)$$

$$= 87.50$$

ถ้าสมมติว่า คะแนนผ่านขั้นต่ำของปริเขตการบวก คือ 85 ก็จึงสรุปได้ว่า เด็กชาย ก. สอบได้ กล่าวคือ เป็นผู้มีความรู้รอบในปริเขตการบวก

ค่าคะแนนปริเขต ( $\hat{P}$ ) จะเป็นตัวคาดคะเนระดับผลสัมฤทธิ์ทางการเรียนที่ปราศจากอคติ

(unbiased estimator) และมีความพอเพียง (sufficient statistic) ก็ต่อเมื่อต้องมีการระบุเซทของข้อสอบที่เป็นไปได้ในแต่ละปริเขตไว้ล่วงหน้า และข้อสอบในแบบสอบอิงปริเขตต้องเป็น กลุ่มตัวอย่างแบบสุ่มของข้อสอบในปริเขตนั้น เป็นที่น่าสังเกตว่า คะแนนรวมใช้เป็นตัวคาดคะเนที่ดีได้ ก็ต่อเมื่อข้อสอบแต่ละข้อล้วนวัดในสิ่งเดียวกัน หรือแบบสอบนั้นมีลักษณะเอกพันธ์ (Homogeneous) แบบสอบอิงปริเขตเป็นแบบสอบที่มีคุณสมบัติในเชิงเอกพันธ์สูง ซึ่งสามารถใช้คะแนนรวมเป็นตัวประมาณค่าผลสัมฤทธิ์ทางการเรียน หรือลักษณะ (trait) ที่มุ่งวัดได้

Millman (1974) ได้เสนอว่า ในการเปรียบเทียบคะแนนปริเขตกับคะแนนเกณฑ์นั้นควรจะมีการกำหนดคะแนนเกณฑ์เป็นช่วงมากกว่าเป็นจุด กล่าวคือ Millman ได้เสนอให้ใช้หลักสูตรคาดคะเนช่วงคะแนนเกณฑ์ ดังนี้

$$UB = \sqrt{\left[ \frac{N-n}{N-1} \right] \frac{P_s (1-P_s)}{n}}$$

เมื่อ UB = ช่วงคะแนนเกณฑ์ (%)



$N$  = จำนวนข้อสอบในปริเซต

$n$  = จำนวนข้อสอบในแบบสอบอิงปริเซต

$P_s$  = เกณฑ์ผ่าน (%)

ตัวอย่างเช่น  $N = 100$   $n = 40$   $P_s = 85\%$  คำนวณได้  $UB = 9\%$  จะได้คะแนนที่อยู่ในช่วงแดนสนธยา คือ ระหว่าง 76% ถึง 94% ใครได้ต่ำกว่า 76% ถือว่าตก และใครได้เกิน 94% ถือว่าได้เป็นต้น

### บรรณานุกรม

- Brennan, R.L., and Kane, M.T. "Signal/Noise Ratios for Domain Referenced Tests" *Psychometrika* 42, 4, December, 1977, 609-625
- Hively, W.A. Test - item pool for MINNEMAST science unit 2.1 : Measuring weight. Unpublished paper. MINNEMAST Project University of Minnesota, 1966.
- Hively, W., and Reynolds, M.C. (Editors). *Domain - Referenced Testing in Special Education*. The Leadership Training Institute/Special Education, University of Minnesota, 1975.
- Hively, W. Domain-Referenced Testing. Working Paper, MINNEMAST Project, University of Minnesota, Fall, 1972 (mimeo)
- Hively, W. "Introduction to Domain-Referenced Testing." *Educational Technology* vol XIV, No.6, 1974, pp.5 - 10.
- Hively, W. Patterson, H.L. & Page, S.A. "'universe defined' system of arithmetic achievement test." *Journal of Educational Measurement* .5. 1968, 275-290.
- Johnson, T.J. "Program and Product Evaluation from a Domain-Referenced Viewpoint." *Educational Technology*. vol XIV, No.6, 1974, pp. 43-48.
- Millman, J. "Sampling Plans for Domain-Referenced Tests." *Educational Technology*. vol. XIV, No.6, 1974. pp.17-21.

- Millman, J. "Sampling plans for Domain-Referenced tests." **Educational Technology**, 14, 1974, 17-21.
- Lord, F.M. & Novick, M.R. **Statistical theories of mental test scores** Reading, Massachusetts : Addison-Wiley, 1968.
- Roid, G.H., and Haladyna, T.M. **A Technology for Test-Item Writing**. New York : Academic Press, 1982.