

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง โดยสรุปเสนอเนื้อหาตามลำดับหัวข้อ ต่อไปนี้

1. สาระและมาตรฐานการเรียนรู้การงานอาชีพและเทคโนโลยีในหลักสูตรการศึกษา
ขั้นพื้นฐาน พุทธศักราช 2551

2. สาระเกี่ยวกับการวัดผลภาคปฏิบัติ

2.1 ความหมายของการวัดผลภาคปฏิบัติ

2.2 ประเภทของการวัดผลภาคปฏิบัติ

2.3 หลักในการวัดผลภาคปฏิบัติ

2.4 เครื่องมือที่ใช้วัดผลภาคปฏิบัติ

2.5 ความหมายของการสังเกต

2.6 ประเภทของการสังเกต

2.7 การสร้างแบบวัดภาคปฏิบัติ

3. การหาคุณภาพและการตรวจสอบคุณภาพของเครื่องมือ

3.1 ความเที่ยงตรง

3.2 อำนาจจำแนก

3.3 ความเชื่อมั่น

4. การสร้างเกณฑ์ปกติ

5. งานวิจัยที่เกี่ยวข้อง

5.1 งานวิจัยในประเทศ

5.2 งานวิจัยต่างประเทศ

1. สาระและมาตรฐานการเรียนรู้กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี ในหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551

กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี เป็นกลุ่มสาระที่ช่วยพัฒนาให้ผู้เรียน มีความรู้ ความเข้าใจ มีทักษะพื้นฐานที่จำเป็นต่อการดำรงชีวิตและรู้เท่าทันการเปลี่ยนแปลง สามารถนำความรู้เกี่ยวกับการดำรงชีวิต การอาชีพและเทคโนโลยีมาประยุกต์ใช้ในการทำงาน อย่างมีความคิดสร้างสรรค์และแข่งขันในสังคมไทยและสากล เห็นแนวทางในการประกอบอาชีพ รักการทำงานและมีเจตคติที่ดีต่อการทำงาน สามารถดำรงชีวิตอยู่ในสังคมได้อย่างพอเพียงและมีความสุข

สาระและมาตรฐานการเรียนรู้

กลุ่มสาระการเรียนรู้กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี มุ่งพัฒนาผู้เรียนแบบองค์รวมเพื่อให้มีความรู้ความสามารถ มีทักษะในการทำงาน เห็นแนวทางในการประกอบอาชีพและการศึกษาต่อได้อย่างมีประสิทธิภาพ

สาระที่ 1 การดำรงชีวิตและครอบครัว เป็นสาระเกี่ยวกับการทำงานในชีวิตประจำวัน ช่วยเหลือตนเอง ครอบครัวยุคใหม่และสังคมได้ในสภาพเศรษฐกิจที่พอเพียง ไม่ทำลายสิ่งแวดล้อม เน้นการปฏิบัติจริงจนเกิดความมั่นใจและภูมิใจในผลสำเร็จของงานเพื่อให้ค้นพบความสามารถ ความถนัด และความสนใจของตนเอง

มาตรฐาน ง 1.1 เข้าใจการทำงาน มีความคิดสร้างสรรค์ มีทักษะกระบวนการทำงาน ทักษะการจัดการ ทักษะกระบวนการแก้ปัญหา ทักษะการทำงานร่วมกัน และทักษะการแสวงหาความรู้ มีคุณธรรม และลักษณะนิสัยในการทำงาน มีจิตสำนึกในการใช้พลังงาน ทรัพยากรและสิ่งแวดล้อมเพื่อการดำรงชีวิตและครอบครัว

สาระที่ 2 การออกแบบและเทคโนโลยี เป็นสาระเกี่ยวกับการพัฒนาความสามารถของมนุษย์อย่างสร้างสรรค์ โดยนำความรู้มาใช้กับกระบวนการเทคโนโลยี สร้างสิ่งของ เครื่องใช้ วิธีการหรือเพิ่มประสิทธิภาพในการดำรงชีวิต

มาตรฐาน ง 2.1 เข้าใจเทคโนโลยีและกระบวนการเทคโนโลยี ออกแบบและสร้างสิ่งของเครื่องใช้หรือวิธีการ ตามกระบวนการเทคโนโลยีอย่างมีความคิดสร้างสรรค์เลือกใช้เทคโนโลยีในทางสร้างสรรค์ต่อชีวิต สังคม สิ่งแวดล้อมและมีส่วนร่วมในการจัดการเทคโนโลยีที่ยั่งยืน

สาระที่ 3 เทคโนโลยีสารสนเทศและการสื่อสาร เป็นสาระเกี่ยวกับกระบวนการเทคโนโลยีสารสนเทศ การติดต่อสื่อสาร การค้นหาข้อมูล การใช้ข้อมูลและสารสนเทศ การแก้ปัญหาหรือการสร้างงาน คุณค่าและผลกระทบของเทคโนโลยีสารสนเทศและการสื่อสาร

มาตรฐาน ง 3.1 เข้าใจ เห็นคุณค่า และใช้กระบวนการเทคโนโลยีสารสนเทศในการสืบค้นข้อมูลการเรียนรู้ การสื่อสาร การแก้ปัญหา การทำงานและอาชีพอย่างมีประสิทธิภาพ ประสิทธิผลและมีคุณธรรม

สาระที่ 4 การอาชีพ เป็นสาระเกี่ยวกับทักษะที่จำเป็นต่ออาชีพ เห็นความสำคัญของคุณธรรม จริยธรรมและเจตคติที่ดีต่ออาชีพ ใช้เทคโนโลยีได้เหมาะสม เห็นคุณค่าของอาชีพสุจริต และเห็นแนวทางประกอบอาชีพ

มาตรฐาน ง 4.1 เข้าใจ มีทักษะที่จำเป็น มีประสบการณ์ เห็นแนวทางในงานอาชีพ ใช้เทคโนโลยีพัฒนาอาชีพ มีคุณธรรมและเจตคติที่ดีต่ออาชีพ

2. สาระเกี่ยวกับการวัดผลภาคปฏิบัติ (Performance Assessment)

2.1 ความหมายของการวัดผลภาคปฏิบัติ

นักการศึกษาและนักจิตวิทยาได้ให้ความหมายของการวัดผลภาคปฏิบัติไว้หลายท่านดังต่อไปนี้

พงษ์พิมล คำลอย (2545 : 8) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่าการวัดผลภาคปฏิบัติเป็นการวัดความสามารถของบุคคลที่แสดงออกมาเป็นพฤติกรรมภายใต้สถานการณ์และเงื่อนไขที่กำหนด โดยจะวัดทั้งวิธีการและผลงานที่แสดงออกมา

ธนศักดิ์ เคนบุปผา (2546 : 11) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่าเป็นการตรวจสอบความรู้ความสามารถ ความคิด ทักษะของผู้ที่เคยทดสอบซึ่งมุ่งวัดพฤติกรรมที่แสดงออกภายใต้สถานการณ์ที่กำหนด สามารถวัดได้ทั้งวิธีการและผลงาน

อาทิตย์ ดั่นเจริญ (2548 : 28) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่าคือเครื่องมือที่สร้างขึ้นเพื่อใช้วัดผลการเรียนรู้ของผู้เรียนในด้าน การปฏิบัติงานที่ให้ผู้เรียนได้แสดงการกระทำออกมาในขณะที่ทดสอบในสถานการณ์ที่จัดขึ้นโดยจะวัดทั้งวิธีการ (Process) และผลงาน (Product) ที่ได้จากการปฏิบัติกิจกรรมนั้นๆ

กษามาศ คงตุก (2549 : 19) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่าการวัดผลภาคปฏิบัติ มักจะเกี่ยวกับการตรวจสอบความรู้ ความสามารถ ความคิด ทักษะ

ของนักเรียนที่แสดงออกด้วยการกระทำ ภายใต้สถานการณ์ที่อาจจะอยู่ในรูปของวิธีการหรือผลงาน

สุรชัย สิงห์แผ่น (2550 : 13) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่า หมายถึง การวัดความสามารถของนักเรียนในการทำกิจกรรมอย่างใดอย่างหนึ่ง โดยที่นักเรียนได้ลงมือปฏิบัติตามจุดมุ่งหมายที่กำหนดขึ้นตามสภาพความเป็นจริงมากที่สุด โดยจะวัดตั้งแต่ขั้นเตรียม ขั้นปฏิบัติ ขั้นผลงาน และขั้นกิจนิสัย

อำนวยการ ฟ้าสว่าง (2550 : 23) ได้ให้ความหมายของการวัดผลภาคปฏิบัติว่า การวัดทักษะภาคปฏิบัติเป็นการวัดความสามารถของบุคคลที่แสดงออกมาเป็นพฤติกรรมทางพุทธิพิสัย ทักษะพิสัยและจิตพิสัย ซึ่งการวัดนั้นจะเริ่มต้นจากการสังเกต การเตรียมการปฏิบัติ การลงมือปฏิบัติ ผลของการปฏิบัติ และลักษณะนิสัยที่จะทำหลังจากเสร็จสิ้นการปฏิบัติงานนั้น ๆ ของผู้เรียน

สมนึก ภัททิยธนี (2553 : 50) ได้ให้ความหมายว่า เป็นการวัดผลงานที่ให้นักเรียนลงมือปฏิบัติซึ่งสามารถวัดได้ทั้งกระบวนการและผลงาน ในสภาพตามธรรมชาติ (สถานการณ์จริง) หรือในสภาพที่กำหนดขึ้น(สถานการณ์จำลอง) เหมาะกับวิชาที่เน้นการปฏิบัติมากกว่าภาคทฤษฎีและสามารถวัดควบคู่ไปกับภาคทฤษฎี คือการใช้แบบทดสอบ

ไพศาล วรรคำ (2554 : 236) ได้กล่าวไว้ว่า แบบทดสอบภาคปฏิบัติ เป็นแบบทดสอบที่ใช้วัดทักษะความสามารถในการปฏิบัติงาน โดยการกำหนดภาระงาน เครื่องมือ วัสดุอุปกรณ์ให้ผู้เข้าสอบได้ปฏิบัติตามคำสั่ง หรือสถานการณ์ที่กำหนด ลักษณะการตอบสนองของผู้เข้าสอบจะไม่ได้ตอบคำถามด้วยการเขียนหรือการพูด แต่จะแสดงออกด้วยการปฏิบัติงาน เช่นการสอบทักษะการทดลอง การสอบปฏิบัติการตอเนื่องจากความหมายของการวัดทักษะภาคปฏิบัติที่กล่าวมา พอสรุปได้ว่า การวัดผลภาคปฏิบัติ หมายถึง การวัดความสามารถของนักเรียนจากการลงมือปฏิบัติกิจกรรมอย่างใดอย่างหนึ่ง ตามจุดประสงค์ที่ตั้งไว้ตามความเป็นจริง โดยวัดผลตามลำดับขั้นตอน คือ ขั้นเตรียม ขั้นปฏิบัติ ขั้นผลงาน และกิจนิสัย

2.2 ประเภทของการวัดผลภาคปฏิบัติ

สมนึก ภัททิยธนี (2553: 50-51) ได้กล่าวว่าประเภทของการวัดผลงานภาคปฏิบัติแบ่งได้หลายลักษณะขึ้นอยู่กับเกณฑ์ที่ใช้ในการแบ่ง มีดังนี้

1. แบ่งตามด้านที่ต้องการวัด แบ่งได้เป็น 2 ประเภท คือ

1.1 การวัดกระบวนการ (Process) เป็นการวัดที่พิจารณาเฉพาะวิธีทำ วิธีปฏิบัติในการทำงานหรือทำกิจกรรมให้สำเร็จ เช่น พิจารณาวีธีที่ผู้เรียนทำการทดลองในห้องปฏิบัติการในวิชาวิทยาศาสตร์ การใช้เครื่องมือช่างทำเฟอร์นิเจอร์ การตีเทนนิสแบบลูกหลังมือ (Back Hand) การกล่าวสุนทรพจน์ ฯลฯ

1.2 การวัดผลงาน (Product) เป็นการวัดที่พิจารณาเฉพาะผลงานหรือผลผลิต ซึ่งเป็น ผลที่เกิดขึ้นจากการทำงานหรือกิจกรรม เช่น ตัวเฟอร์นิเจอร์ที่นักเรียนผลิตออกมา ภาพวาด ของนักเรียน ดอกไม้ประดิษฐ์จากฝีมือนักเรียน ฯลฯ

ในบางครั้งจะประเมินทั้งกระบวนการและผลผลิต เช่น การเสียบกิ่งมะม่วง แต่บางครั้งประเมินเฉพาะด้านใดด้านหนึ่ง เช่น ในการวาดภาพมักประเมินเฉพาะผลงานทางเดียว

2. แบ่งตามลักษณะสถานการณ์ แบ่งได้เป็น 2 ประเภท คือ

2.1 ใช้สถานการณ์จริง (Real Setting) เป็นการวัดผลงานภาคปฏิบัติโดยใช้สถานการณ์จริง

2.2 สถานการณ์จำลอง (Simulated Setting) การวัดผลภาคปฏิบัติในบางเรื่องต้องใช้สถานการณ์จำลอง เพราะถ้าใช้สถานการณ์จริงจะสิ้นเปลืองมาก มีอันตรายหรือไม่สามารถกระทำได้ เช่น การติดตั้งไฟฟ้าในอาคาร การฝึกนักบินใหม่

3. แบ่งตามสิ่งเร้า แบ่งได้เป็น 2 ประเภท คือ

3.1 ใช้สิ่งเร้าที่เป็นธรรมชาติ (Natural Stimulus) เป็นการวัดผลงานภาคปฏิบัติที่เป็นไปตามธรรมชาติ ผู้วัดไม่ได้เข้าไปยุ่งเกี่ยว เช่น ทักษะทางสังคมของผู้เรียนที่ผู้วัดทำการสังเกตในสภาพที่เป็นไปตามธรรมชาติ ไม่ได้กำหนดให้ปฏิบัติ นิยมใช้วัดคุณลักษณะของบุคลิกภาพ นิสัยการทำงาน ความเต็มใจในการปฏิบัติตามกระบวนการที่กำหนดให้ปฏิบัติ เช่น เกี่ยวกับกฎความปลอดภัย เป็นต้น

3.2 ใช้สิ่งเร้าที่จัดขึ้น (Structured Stimulus) เป็นการวัดโดยจัดสิ่งเร้าที่สามารถแสดงให้เห็นพฤติกรรมที่ต้องการประเมินได้หรือปรากฏให้เห็นเด่นชัด เช่น การให้นักเรียนเตรียมและกล่าวสุนทรพจน์ การให้ทดลองในห้องปฏิบัติการ การอ่านออกเสียง การเล่นเกม ฯลฯ วิธีนี้ จะลดเวลาการสังเกตลง เพราะไม่ต้องรอให้เกิดขึ้นตามธรรมชาติ

กระทรวงศึกษาธิการ (2544 : 6-8) ได้จำแนกประเภทของแบบวัดภาคปฏิบัติตามระดับความเป็นจริงของสถานการณ์ ซึ่งจำแนกได้ 4 ลักษณะ ดังนี้

1. การปฏิบัติงานโดยใช้ข้อเขียน (Paper and Pencil Performance) ลักษณะของเครื่องมือวัดภาคปฏิบัติประเภทนี้แตกต่างจากแบบทดสอบวัดผลสัมฤทธิ์ชนิดเขียนตอบ (Paper and Pencil Performance) โดยที่เครื่องมือวัดประเภทนี้เน้นในการประยุกต์ความรู้และทักษะจากสถานการณ์ที่กำหนดขึ้น คำกริยาแสดงอาการ (Action Verb) ของแบบวัดมักจะมีคำว่าสร้าง (Construct) เช่น โจทย์กำหนดให้นักเรียนสร้างแบบเลื่อนกระดาษ เครื่องมือวัดประเภทนี้อาจจะใช้เป็นการทดสอบขั้นต้นก่อนที่จะให้นักเรียนได้ลงมือปฏิบัติจริง

2. ระบุชื่อและกระบวนการปฏิบัติ (Identification Tests) เป็นรูปแบบที่ให้นักเรียนระบุชื่อ เครื่องมือ หรือชิ้นส่วนของอุปกรณ์ต่างๆ พร้อมทั้งระบุหน้าที่ของสิ่งเหล่านี้ด้วย ถ้าเป็นในระดับที่ซับซ้อน (Complex) อาจจะเป็นรูปแบบของการแสดงชิ้นส่วนของงาน เช่น ส่วนที่เกิดจากไฟฟ้าลัดวงจรแล้วถามนักเรียนถึงกระบวนการปฏิบัติ เพื่อการซ่อมแซม พร้อมทั้งให้ระบุถึงเครื่องมือ วัสดุอุปกรณ์ที่ใช้ในการซ่อมแซมด้วย ถ้าเป็นในระดับที่ซับซ้อนกว่านี้ จะให้นักเรียนฟังเสียงการทำงานของเครื่องกลึง แล้วให้ระบุส่วนที่ชำรุดของเครื่องจักรกลนั้นๆ พร้อมทั้งระบุการซ่อมบำรุงด้วยตัวอย่าง เช่น นักเรียนสามารถบอกชนิดเครื่องดนตรีจากเสียงที่ได้ยินหรือนักเรียนสามารถอธิบายวิธีการขั้นตอนในการเชื่อมโลหะด้วยไฟฟ้าทำต่าง ๆ เป็นต้น

3. การสร้างสถานการณ์จำลอง (Simulated Performance) เนื่องจากผู้สอบไม่สามารถที่จะนำผู้เรียนไปทดสอบภาคปฏิบัติกับสถานการณ์จริงได้ จำต้องกำหนดสถานการณ์ขึ้นมาให้คล้ายคลึงกับสภาพความเป็นจริงมากที่สุด เช่น ให้ผู้สอบขับรถยนต์ที่สนามฝึกขับรถยนต์ ทดสอบการฝึกขับรถของเครื่องบินของบริษัทการบิน ซึ่งมีสถานการณ์จำลองใช้ส่วนของเครื่องบินที่แท้จริง มีสนามบินต่างประเทศกำหนดไว้เป็นสถานการณ์ ซึ่งควบคุมโดยเครื่องคอมพิวเตอร์ การสร้างสถานการณ์จำลองนี้เพื่อเป็นการป้องกันอันตราย และเพื่อเป็นการเสียต่อการชำรุดเสียหายของเครื่องบินที่มีราคาแพง ในระยะเริ่มต้นของการฝึกทักษะสำหรับการประเมินนั้นใช้แบบประเมิน ทั้งวิธีการ (Process) และผลงาน (Product)

4. การกำหนดผลงาน (Work Sample) เครื่องมือวัดภาคปฏิบัติลักษณะนี้สามารถวัดระดับความเป็นจริง (Realism) ของสถานการณ์ได้สูงสุด ผู้ถูกทดสอบจะต้องปฏิบัติงานในสถานการณ์จริง การสร้างสถานการณ์จำลองกับการกำหนดงาน บางครั้งแยกกันได้ยาก แต่ในบางเรื่องก็แยกจากกันได้อย่างชัดเจน เช่น การกำหนดผู้เรียนสร้างตู้ 1 หลัง ผู้เรียนหรือผู้ถูกทดสอบจะต้องลงมือปฏิบัติสร้างตู้โดยใช้วัสดุอุปกรณ์ตามที่กำหนด สร้างตามขนาดที่กำหนดให้ แต่ในสถานการณ์ ของการขับรถยนต์ตามท้องถนน โดยมีเป้าหมายว่าต้องผ่านสิ่ง

ใดบ้าง เช่น ทางแยก วงเวียน สะพาน ลักษณะนี้ก็เป็นกำหนัดงาน ในการประเมินผลนั้นก็ ใช้แบบประเมิน เพื่อประเมินทั้งวิธีการ (Process) และผลงาน (Product)

บุญชม ศรีสะอาด (2540 : 55-56) ได้จำแนกแบบทดสอบภาคปฏิบัติตาม ลักษณะของงานที่กำหนดให้เป็น 3 ประเภท ดังนี้

1. แบบจำแนก (Recognition or Identification) เป็นแบบที่วัดความสามารถ ในการจำลักษณะที่จำเป็นของการกระทำหรือผลงานหรือจำแนกสิ่งของ ลักษณะของการวัด เช่น จะทำเครื่องมือมีความบกพร่องนั้นคิดพินิจไปจากสภาพปกติ (อาจปรับไว้ไม่ดีหรือนำ บางชิ้นออก) ให้ผู้สอบจำแนกจุดบกพร่องนั้น เป็นการวัดการรู้จักถึงความผิดพลาดความ ถูกต้องของเครื่องมือ กระบวนการ หรือผลผลิต ลักษณะของการวัดอีกลักษณะหนึ่ง คือ ให้ ผู้สอบพิจารณาตัดสินเลือกผลงานที่ดีและที่ด้อย อาจเป็นผลงานทางศิลปะ ผลงานทาง การเขียนเรียงความ

2. แบบการใช้สถานการณ์จำลอง (Simulated Situation) เป็นการวัดที่ไม่ใช่ สถานการณ์จริง แต่จำลองสถานการณ์หรือการปฏิบัติจริงที่มุ่งวัด เช่น วัดความสามารถใน การบังคับทิศทางพวงมาลัย และปฏิกริยาในการหยุดรถ โดยใช้เครื่องจำลองไม่ได้ซึ่งรถยนต์ จริงตามถนน แม้ว่าไม่ใช่สถานการณ์ แต่ก็มีข้อดีหลายประการ เช่น มีความประหยัด สะดวก และปลอดภัยกว่าสถานการณ์จริงมาก แบบทดสอบประเภทนี้บางครั้งเรียกว่า Miniature Test

3. แบบใช้ตัวอย่างงาน (Work Sample) เป็นการให้ปฏิบัติตามภาวะปกติของ การปฏิบัติงานประเภทนั้นๆ อาจกำหนดให้ปฏิบัติตามลำดับที่สมบูรณ์ของพฤติกรรมทำงานก็ได้ เนื่องจากงานบางอย่างจะต้องใช้เวลาและมีปัญหาเรื่องค่าใช้จ่าย โดยทั่วไปจึงนิยมเลือก ตัวอย่างของการปฏิบัติซึ่งสามารถพยากรณ์พฤติกรรมทั้งหมดได้อย่างเพียงพอ งานบางอย่างจะ มีความ ถูกผิดอย่างชัดเจน เช่น การปาเป้า การสอบพิมพ์ดีด การให้คะแนนการปฏิบัติ งาน ประเภทนี้ จะมีความเป็นปรนัย แต่งานบางอย่างให้คะแนนยาก ขึ้นกับการพิจารณาของ ผู้ประเมิน เช่น การเล่นเกมหรือคุณภาพของงาน การปฏิบัติที่สะท้อนจากผลงานที่ปรากฏ เช่น การวาดภาพ

จากประเภทของการวัดที่กล่าวมา ผู้วิจัยได้เลือกใช้หลักการวัดผลงานภาคปฏิบัติ ตามลักษณะการใช้สถานการณ์จริง โดยให้ผู้เรียนลงมือปฏิบัติจริงตามขั้นตอน ซึ่งจะสามารถ วัดพฤติกรรมที่แท้จริงของผู้ปฏิบัติได้

2.3 หลักในการวัดผลภาคปฏิบัติ

สมนึก ภักดิ์ทิษณีย์ (2553 : 50) ได้กล่าวถึงการวัดผลภาคปฏิบัติว่าเป็นสิ่งที่ควรคำนึงถึงในการสอบวัดผลภาคปฏิบัติคือ

1. ชั้นเตรียมงาน
2. ชั้นปฏิบัติงาน
3. เวลาที่ใช้ในการทำงาน
4. ผลงาน

สุรชัย สิงห์แผ่นดิน (2550 : 17) ได้กล่าวว่าหลักในการวัดผลงานภาคปฏิบัติจะขึ้นอยู่กับลักษณะของงานที่สอดคล้องกับสภาพจริง ความมุ่งหมายและวัตถุประสงค์ในการสอนเรื่องนั้น ๆ ซึ่งได้ใช้หลักการในการวัดผลทักษะภาคปฏิบัติคือ

1. ชั้นเตรียมงาน
2. ชั้นปฏิบัติงาน
3. ชั้นผลงาน
4. ชั้นกึ่งนิสัย

หลักในการวัดผลงานภาคปฏิบัติที่ดี มีดังนี้

1. ต้องกำหนดจุดประสงค์ของการวัดทักษะให้ชัดเจน ซึ่งอาจวัดที่กระบวนการหรือผลงานหรือทั้งสองอย่าง
2. เนื้อหาสาระของงานที่ให้นักเรียนปฏิบัติมีลักษณะสอดคล้องกับสภาพจริง (Authentic)
3. คุณภาพของสิ่งที่วัด ในครั้งหนึ่งๆมีจำนวนเพียงพอและสามารถวัดได้ตรง
4. กำหนดเงื่อนไขในการวัดให้ชัดเจน
5. ในการวัดโดยใช้สิ่งเร้าที่จัดขึ้น (Structured Stimulus) ต้องเขียนคำชี้แจงอย่างกระชับ ชัดเจนและสมบูรณ์
6. แบบฟอร์มที่ใช้วัดมักจะเป็นแบบตรวจสอบรายการ (Checklist) หรือมาตราส่วนประมาณค่า (Rating Scale) ซึ่งสามารถให้คะแนนได้สะดวก แต่ต้องระบุเกณฑ์การให้คะแนนอย่างชัดเจน

จากหลักการวัดผลภาคปฏิบัติดังกล่าวข้างต้น จะเห็นได้ว่าหลักในการวัดผลภาคปฏิบัติได้แบ่งขั้นตอนในการวัดที่คล้าย ๆ กัน คือจะต้องมีชั้นเตรียมงาน ชั้นปฏิบัติงาน ชั้นผลงาน รวมทั้งเวลาที่ใช้ในการปฏิบัติงานและสุดท้ายจะต้องมีการวัดคุณลักษณะนิสัยในการปฏิบัติงาน

ผู้วิจัยจึงยึดหลักในการวัดผลภาคปฏิบัติตามหลักการของ สมนึก กัททิยธนี และสุรชัย สิงห์แผ่น โดยผู้วิจัยได้นำขั้นเวลาที่ใช้ในการปฏิบัติงาน ไปรวมอยู่ในหัวข้อย่อยของขั้นผลงาน ดังนั้นหลักในการวัดผลทักษะภาคปฏิบัติที่ผู้วิจัยนำมาใช้วัดพฤติกรรมของผู้ปฏิบัติ คือ ขั้นเตรียมงาน ขั้นปฏิบัติงาน ขั้นผลงานและกิจนิสัย ซึ่งจะสามารถวัดพฤติกรรมของผู้ปฏิบัติ ได้อย่างครอบคลุม

2.4 เครื่องมือการวัดผลภาคปฏิบัติ

บุญชม ศรีสะอาด (2540 : 55) ได้กล่าวถึงเครื่องมือวัดผลภาคปฏิบัติว่า เป็นเครื่องมือที่วัดเกี่ยวกับทักษะทางกาย เช่น การสร้างแบบทดสอบวัดสมรรถภาพทางกลไก การสร้างแบบทดสอบวัดภาคปฏิบัติในวิชาต่าง ๆ ที่สร้างขึ้นเพื่อวัดความสามารถในการปฏิบัติหรือการกระทำของผู้เรียน ซึ่งอาจจำแนกแบบทดสอบวัดภาคปฏิบัติตามลักษณะของงานที่กำหนดให้ทำเป็น 3 ชนิดดังนี้

1. แบบการจำแนก (Recognition or Identification) เป็นแบบวัดที่วัดความสามารถในการจำลักษณะที่จำเป็นของการกระทำหรือผลงานหรือจำแนกสิ่งของ
2. แบบใช้สถานการณ์จำลอง (Simulated Situation) เป็นการวัดที่ไม่ใช้สถานการณ์จริง แต่จะจำลองสถานการณ์หรือปฏิบัติจริงที่มุ่งวัด เช่น วัดความสามารถในการบังคับทิศทางด้วยพวงมาลัยและปฏิกริยาในการหยุดรถ โดยใช้เครื่องจำลองไม่ได้ขับรถยนต์จริงตามท้องถนน

3. แบบใช้ตัวอย่าง (Work Sample) เป็นการให้ปฏิบัติตามภาวะปกติของการปฏิบัติงานประเภทนั้น ๆ อาจกำหนดให้ปฏิบัติตามลำดับที่สมบูรณ์ของพฤติกรรมหรืออาจเลือกเพียงตัวอย่างของพฤติกรรมการทำงานก็ได้ เนื่องจากงานบางอย่างจะต้องใช้เวลาและปัญหาค่าใช้จ่าย โดยทั่วไปจึงนิยมเลือกตัวอย่างของการปฏิบัติซึ่งสามารถพยากรณ์พฤติกรรมทั้งหมดได้อย่างเพียงพอ งานบางอย่างจะมีความถูกต้องอย่างชัดเจน เช่น การปาเป้า การสอบพิมพ์ดีด การให้คะแนนการปฏิบัติงานประเภทนี้จะมีความเป็นปรนัย แต่งานบางอย่างให้คะแนนยาก ขึ้นกับการพิจารณา ของผู้ประเมิน เช่น การเล่นเกม การวาดภาพ ฯลฯ

ซึ่งแบบใช้ตัวอย่างงาน (Work Sample) นับว่าเป็นแบบที่มีความเที่ยงตรงและความเชื่อมั่นมากกว่าแบบอื่น ๆ ในการประเมินผลภาคปฏิบัติผู้ประเมินผลจะอาศัยการสังเกตวิธีปฏิบัติหรือผลผลิต โดยใช้แบบทดสอบที่มีหลายลักษณะ ได้แก่ การเรียงอันดับ การประเมินระดับค่า การตรวจสอบรายการ การเปรียบเทียบระดับผลผลิต และการใช้ระเบียบพฤติกรรม

สุรชัย สิงห์แผ่น (2550 : 20) ได้กล่าวว่า ลักษณะของเครื่องมือที่ใช้ประกอบการวัดผลภาคปฏิบัติมีลักษณะดังนี้

1. การสังเกต (Observation) การสังเกตที่ดีต้องปล่อยให้ผู้ถูกสังเกตอยู่ในสภาพที่ปกติเพื่อจะได้ข้อมูลตามความเป็นจริง การสังเกตอาจทำได้โดยผู้สังเกตเข้าไปอยู่ในกลุ่มด้วยเปรียบเสมือนเป็นสมาชิกผู้หนึ่งของกลุ่ม หรือผู้สังเกตจะแอบอยู่ที่อื่น โดยไม่ให้ผู้ถูกสังเกตรู้ตัวได้ ในการสังเกตจะต้องมีการวางแผนเสียก่อนว่าสังเกตเมื่อไร ะไรบ้าง ตั้งจุดมุ่งหมายของการสังเกต แต่ละครั้งนอกจากนั้นจะต้องเตรียมบันทึกข้อมูลโดยใช้เครื่องต่าง ๆ เช่น มาตรการส่วนประมาณค่า การบันทึกต่างๆ แบบสำรวจพฤติกรรม

2. การเรียงอันดับ (Ranking) การจัดอันดับเป็นวิธีที่จะเรียงลำดับนักเรียนในคุณสมบัติหนึ่งๆ ตามที่กำหนดให้ซึ่งสามารถใช้ในการวัดวิธีการหรือผลงานก็ได้ แต่ส่วนใหญ่จะใช้ การวัดผลงานมากกว่า การใช้จัดอันดับจะมีความเชื่อมั่นสูงขึ้น ถ้าจัดอันดับคุณสมบัติหนึ่งๆที่จำเพาะและมีคำจำกัดความของคุณสมบัตินั้นชัดเจน แต่ถ้าจัดอันดับหลายอย่างในคราวเดียวกัน จะทำให้ ค่าความเชื่อมั่นต่ำลง ตัวอย่าง เช่น การเรียนขับรถ ครูจัดอันดับเกี่ยวกับความสามารถ ในการหยุดรถอย่างไม่ใช้อันดับความสามารถในการใช้รถ (ซึ่งรวมทាំង การออกรถ การขับพวงมาลัย การจอด เป็นต้น)

3. มาตรการส่วนประมาณค่า (Rating Scale) มาตรการส่วนประมาณค่าเป็นเครื่องมือที่ใช้มากในการประเมินค่าการปฏิบัติ มาตรการส่วนประมาณค่ามีหลายรูปแบบ แต่ที่นิยมใช้มากที่สุดที่จัดคุณลักษณะต่อเนื่อง ซึ่งจะแบ่งคุณลักษณะนั้นตามระดับความสูง-ต่ำ โดยกำหนดเป็น 2 ระดับขึ้นไป จนถึงประมาณ 10 ระดับ ความเชื่อถือได้ของข้อมูล จากมาตรการส่วนประมาณค่านอกจาก จะขึ้นอยู่กับความสามารถของผู้ถูกสังเกต ความไม่มีอคติ แล้วยังขึ้นอยู่กับกำหนัดความหมายคุณสมบัติที่กำลังจะพิจารณาด้วยนั้นจะต้องกำหนดคุณสมบัติให้ชัดเจน และเป็นคุณสมบัตีย่อยเช่นเดียวกับการจัดอันดับ

4. แบบสำรวจพฤติกรรม (Checklist) แบบสำรวจพฤติกรรมของนักเรียนรายการของพฤติกรรมให้ผู้สังเกตบันทึกว่า พฤติกรรมนั้นๆ เกิดขึ้นหรือไม่ โดยส่วนมากมักจะบอกว่ามีหรือไม่ ในการสังเกตการณ์ปฏิบัติงาน บางครั้งอาจให้ผู้สังเกตบันทึกลำดับที่ของการปฏิบัติหรือพฤติกรรมตามลำดับตั้งแต่ 1 เป็นต้นไปก็ได้ ซึ่งในลักษณะนี้ก็จะทำให้มองเห็นภาพรวมของการปฏิบัติอีกครั้ง

5. การบันทึกต่างๆ (Anecdotal Record) การบันทึกในกลุ่มนี้มักจะเป็นวิธีการที่ไม่ได้กำหนดรูปแบบไว้อย่างชัดเจนเหมือนวิธีอื่น ๆ ผู้บันทึกค่อนข้างจะอิสระในการบันทึก

ลงไปมากกว่าเครื่องมือชนิดอื่น ๆ การบันทึกเพียงครั้งเดียว อาจไม่สามารถให้ข้อมูลที่มีความหมายมากนัก แต่การบันทึกอย่างต่อเนื่องหลายครั้งจะให้ข้อมูลที่ชัดเจนขึ้น ใน การบันทึกผู้สังเกตจะเขียนถึงพฤติกรรมหรือเหตุการณ์ที่เกิดขึ้นเท่านั้น ไม่ใส่ความคิดเห็นลงไป ยกเว้นกรณีที่ต้องการให้ใส่ความคิดเห็นก็แยกในส่วนที่แสดงความเห็นอย่างชัดเจน

พิชิต ฤทธิจรูญ (2552 : 77) ได้กล่าวถึงเครื่องมือที่ใช้ผลภาคปฏิบัติซึ่งเป็นการวัด ด้านทักษะพิสัยว่า เครื่องมือวัดพฤติกรรมด้านทักษะพิสัยมีหลายประเภทได้แก่ การทดสอบ ภาคปฏิบัติ การสังเกต แบบตรวจสอบรายการ มาตรฐานส่วนประมาณค่า เพิ่มสะสมงาน และการประเมินตามสภาพจริง ซึ่งแต่ละประเภทต่างมีความเหมาะสมกับงานแตกต่างกัน การจะ เลือกใช้เครื่องมือประเภทใดนั้นขึ้นอยู่กับลักษณะของงานและวัตถุประสงค์ในครั้งนั้น ๆ สำหรับวัตถุประสงค์ครั้งหนึ่ง ๆ อาจใช้เครื่องมือหรือวิธีการวัดหลาย ๆ วิธีก็ได้ หากการวัด โดยเครื่องมือหรือวิธีการเดียว ไม่สามารถให้ข้อมูลที่ชัดเจนเพียงพอ

โดยธรรมชาติสาระการเรียนรู้กลุ่มสาระการอาชีพและเทคโนโลยี (งานงานเกษตร) ผู้เรียนจะต้องเรียนรู้ทั้งภาคทฤษฎีและภาคปฏิบัติควบคู่ไปด้วยกัน ผู้วิจัยเห็นว่าเครื่องมือที่ใช้นำมาวัดการทดสอบภาคปฏิบัติ โดยวัดผลจากการลงมือปฏิบัติจริงของผู้เรียนในสถานการณ์จริงเพื่อตรวจสอบความสามารถของผู้เรียนในด้านต่าง ๆ จากการสังเกตโดยใช้เครื่องมือ ประกอบการสังเกตคือการบันทึกข้อมูลลงในแบบตรวจสอบรายการ

2.5 ความหมายของการสังเกต (Observation)

สุรวาท ทองบุ (2550 : 52) ได้ให้ความหมายของการสังเกตว่า การสังเกตเป็น เครื่องมือเก็บรวบรวมข้อมูลที่ต้องอาศัยประสาทสัมผัสหลายอย่าง โดยเฉพาะอย่างยิ่งประสาท สัมผัสทางตาและหูเป็นสำคัญ การสังเกตใช้ได้ดีสำหรับการศึกษาลักษณะและพฤติกรรม ของบุคคล รวมถึงปรากฏการณ์และพิธีการต่าง ๆ ใช้เป็นเครื่องมือรวบรวมข้อมูลได้ทั้งด้าน วิทยาศาสตร์ กายภาพและทางด้านพฤติกรรมศาสตร์และสังคมศาสตร์

พิชิต ฤทธิจรูญ (2552 : 72) ได้ให้ความหมายของการสังเกตว่า การสังเกตเป็น การเก็บรวบรวมข้อมูลโดยใช้ประสาทสัมผัสทั้ง 5 ในลักษณะของการเฝ้าดู ศึกษาเหตุการณ์ ปรากฏการณ์ เพื่อให้เข้าใจธรรมชาติของสิ่งที่สังเกต หรือพฤติกรรมของสิ่งที่เราต้องศึกษา อาจเป็นลักษณะบุคลิกภาพ การใช้คำพูด ภาษาท่าทาง กิจกรรม ทักษะและความสามารถ รวมทั้งสภาพแวดล้อมต่าง ๆ

บุญชม ศรีสะอาด (2553 : 84) ได้ให้ความหมายของการสังเกตว่า การสังเกต เป็นเทคนิคการรวบรวมข้อมูล ที่ผู้สังเกตใช้สายตาเฝ้าดูหรือศึกษาเหตุการณ์ ปรากฏการณ์

ต่าง ๆ เพื่อให้เข้าใจลักษณะธรรมชาติและความเกี่ยวข้องกันระหว่างองค์ประกอบต่าง ๆ ของเหตุการณ์หรือปรากฏการณ์นั้นๆ

สมนึก กัททิษณี (2553 : 32) ได้ให้ความหมายของการสังเกตว่า การสังเกตคือการพิจารณาปรากฏการณ์ต่าง ๆ ที่เกิดขึ้น เพื่อค้นหาความจริงบางประการ โดยอาศัยประสาทสัมผัสของผู้สังเกตโดยตรง ทำให้ได้ข้อมูลแบบปฐมภูมิ (Primary Data) ซึ่งเป็นข้อมูลที่นำเชื่อถือ

ไพศาล วรคำ (2554 : 210) ได้ให้ความหมายของการสังเกตว่า การสังเกตหมายถึง การใช้ประสาทสัมผัสทั้งห้าในการรับรู้ข้อเท็จจริงจากปรากฏการณ์ใด ๆ โดยผู้วิจัยซึ่งถือว่าเป็นเครื่องมือที่สำคัญที่สุดในการสังเกต เพื่อให้ได้ข้อมูลเชิงลึกเกี่ยวกับปรากฏการณ์ที่ไม่สามารถวัดได้โดยตรงและมีความซับซ้อน

จากความหมายของการสังเกตที่กล่าวมา สรุปได้ว่าการสังเกตหมายถึงการใช้ประสาทสัมผัสหลายอย่างโดยเฉพาะตาและหูเป็นสำคัญ เพื่อพิจารณาปรากฏการณ์ที่เกิดขึ้น ทำให้ได้ข้อมูลแบบปฐมภูมิ (Primary Data)

2.6 ประเภทของการสังเกต

ประเภทของการสังเกต มีนักวิชาการหลายท่านได้แบ่งประเภทของการสังเกตไว้หลายแบบ ตามแต่ทัศนะในการแบ่ง ดังนี้

สุรวาท ทองบุ (2550 : 52-53) แบ่งประเภทของการสังเกตออกเป็น 2 ประเภทใหญ่ ๆ คือ

1. การสังเกตทางตรง (Direct Observation) เป็นการสังเกตที่ผู้สังเกตต้องเฝ้าดูเหตุการณ์หรือพฤติกรรมที่เกิดขึ้นด้วยตนเอง โดยอาศัยประสาทสัมผัสทางตาเป็นส่วนใหญ่ รองลงมาคือหู

การสังเกตทางตรงสามารถแบ่งย่อยได้เป็น 2 ลักษณะคือ

1.1 การสังเกตแบบมีส่วนร่วม (Participant Observation) เป็นการสังเกตที่ผู้สังเกตเข้าไปมีส่วนร่วมในเหตุการณ์หรือกิจกรรมนั้น ๆ ด้วย ผู้ถูกสังเกตอาจจะรู้ตัวหรือไม่รู้ตัวก็ได้ แต่ในกรณีที่ต้องการพฤติกรรมที่เป็นไปตามธรรมชาติแลความเป็นจริง ก็ไม่ควรให้ผู้ถูกสังเกตรู้ตัว

1.2 การสังเกตแบบไม่มีส่วนร่วม (Nonparticipant Observation) เป็นการสังเกตอยู่นอกวง กระทำตนเป็นผู้ดูอย่างเดียว ผู้ถูกสังเกตไม่รู้ตัวว่ามีคนคอยสังเกตอยู่ การสังเกตแบบนี้ มีข้อเสียคือไม่สามารถติดตามดูพฤติกรรมทุกอย่างได้

2. การสังเกตทางอ้อม (Indirect Observation) เป็นการสังเกตที่ผู้สังเกตไม่ได้เห็นเหตุการณ์หรือพฤติกรรมที่เกิดด้วยตนเอง แต่อาศัยการถ่ายทอดด้วยเครื่องมืออย่างใดอย่างหนึ่ง เช่น การสังเกตการณ์ถ่ายทำภาพยนตร์ ถ่ายภาพไว้ เป็นต้น

บุญชม ศรีสะอาด (2553 : 84-85) ประเภทของการสังเกตการณ์แบ่งได้หลายแบบ ขึ้นกับเกณฑ์และทักษะในการแบ่ง โดยทั่วไปแบ่งได้เป็น 2 ประเภทคือ

1. การสังเกตทางตรง (Direct Observation) เป็นการสังเกตการณ์ที่ผู้สังเกตการณ์สัมผัสกับเหตุการณ์ที่เกิดขึ้น โดยตรงด้วยตนเอง

2. การสังเกตทางอ้อม (Indirect Observation) เป็นการสังเกตการณ์ที่ผู้สังเกตการณ์ไม่ได้เฝ้าดูหรือศึกษาเหตุการณ์หรือปรากฏการณ์นั้น ๆ โดยตรง แต่จะดูหรือศึกษาจากบันทึกมา เช่น จากภาพยนตร์ โทรทัศน์ เทปบันทึกภาพ เป็นต้น

ไพศาล วรคำ (2554 : 210-211) จำแนกประเภทของการสังเกตออกเป็น 2 ลักษณะดังนี้

1. การสังเกตแบบมีส่วนร่วม (Participatory Observation) เป็นการสังเกตที่ผู้สังเกตเข้าไปมีบทบาทเป็นส่วนหนึ่งของปรากฏการณ์นั้นเช่นเดียวกับบุคคลอื่น ๆ ที่อยู่ในปรากฏการณ์หรือแสดงบทบาทเป็น “คนใน” ของปรากฏการณ์นั้น เช่นผู้วิจัยแสดงบทบาทเป็นครูผู้ช่วยเพื่อสังเกตพฤติกรรมการเรียนการสอนของนักเรียน หรือผู้วิจัยเข้าไปเป็นสมาชิกคนหนึ่งของกลุ่มคนในสลัม เพื่อสังเกตและเรียนรู้วิถีชีวิตของคนในชุมชน เป็นต้น วัตถุประสงค์ของการสังเกตแบบมีส่วนร่วม ก็คือเพื่อให้สามารถรวบรวมข้อมูลที่แท้จริงจากปรากฏการณ์ โดยผู้วิจัยได้รับประสบการณ์จากปรากฏการณ์โดยตรงและบุคคลที่อยู่ในปรากฏการณ์นั้น ไม่รู้ตัวในกรณีนี้ผู้วิจัยจึงไม่จำเป็นต้องแสดงตนว่าเป็นผู้วิจัย

2. การสังเกตแบบไม่มีส่วนร่วม (Non-participatory Observation) เป็นการสังเกตที่ผู้วิจัยหรือผู้สังเกตไม่ได้มีส่วนร่วมกับกิจกรรมหรือปรากฏการณ์ที่เกิดขึ้น ผู้วิจัยมีบทบาทเป็นเพียง “คนนอก” ที่เฝ้าสังเกตพฤติกรรมหรือปรากฏการณ์ทางสังคม จึงเปรียบเสมือนบุคคลแปลกหน้า ในปรากฏการณ์นั้น โดยทั่วไปการสังเกตแบบไม่มีส่วนร่วม ผู้วิจัยจะแสดงสถานภาพที่แท้จริง ให้บุคคลในปรากฏการณ์นั้นทราบเพื่อให้ผู้ที่เกี่ยวข้องในปรากฏการณ์มีความเข้าใจในวัตถุประสงค์ของผู้วิจัยและให้ความร่วมมือที่จะช่วยให้ได้ข้อมูลที่ถูกต้องและเป็นจริง

ในการสังเกตผู้วิจัยจะต้องกำหนดจุดมุ่งหมายหรือประเด็นที่จะทำการสังเกตในแต่ละครั้งให้ชัดเจน ตลอดจนการกำหนดเวลาหรือช่วงเวลาในการสังเกต โดยอาจใช้เครื่องมือ

ช่วยในการบันทึกผลการสังเกต เช่น แบบบันทึกการสังเกต แบบตรวจสอบรายการ หรือแบบประมาณค่า นอกจากนี้อาจใช้การบันทึกภาพ หรือบันทึกวีดิทัศน์ช่วยในการสังเกตด้วยก็ได้

จากการจัดแบ่งประเภทของการสังเกตดังกล่าวข้างต้น ผู้วิจัยใช้การสังเกตโดยผู้สังเกตต้องเฝ้าดูเหตุการณ์หรือพฤติกรรมที่เกิดขึ้นด้วยตนเอง โดยอาศัยประสาทสัมผัสทางตาเป็นส่วนใหญ่ และได้กำหนดเรื่องที่จะสังเกตไว้ล่วงหน้าแน่นอนแล้ว

2.7 การสร้างเครื่องมือวัดผลภาคปฏิบัติ

บุญชม ศรีสะอาด (2540 : 57-58) ได้เสนอวิธีการสร้างแบบทดสอบภาคปฏิบัติไว้ว่าผู้วิจัยจะต้องเลือกเรื่องใดเรื่องหนึ่งในวิชาใดวิชาหนึ่ง ซึ่งเป็นเรื่องที่ทำกรเรียนการสอนโดยวิธีการปฏิบัติและสมควรวัด โดยให้นักเรียนปฏิบัติด้วย วิธีดำเนินการสร้างแบบทดสอบภาคปฏิบัติ มีขั้นตอน ดังนี้

1. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง
2. วิเคราะห์งานและเลือกงานที่เป็นตัวแทน
3. ให้ผู้เชี่ยวชาญตรวจสอบความครอบคลุมของงานที่เป็นตัวแทนหลักสูตร
4. เขียนจุดประสงค์เชิงพฤติกรรม
5. ให้ผู้เชี่ยวชาญตรวจสอบความสอดคล้องระหว่างงานกับจุดประสงค์
6. เขียนแบบทดสอบภาคปฏิบัติ
7. กำหนดสัดส่วนของคะแนนและเกณฑ์การให้คะแนนในการปฏิบัติ
8. ให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรง สัดส่วนของคะแนนและความเป็นปรนัย ของเกณฑ์ในการให้คะแนน
9. จัดพิมพ์แบบทดสอบ
10. นำแบบทดสอบไปทดสอบครั้งที่ 1
11. วิเคราะห์แบบทดสอบเพื่อหาคุณภาพ
12. แก้ไขปรับปรุงและนำไปทดสอบครั้งที่ 2
13. วิเคราะห์แบบทดสอบเพื่อหาคุณภาพ
14. จัดทำคู่มือการใช้แบบทดสอบและจัดพิมพ์แบบทดสอบเป็นรูปเล่ม

ธนศักดิ์ เคนบุปผา (2546 : 25) ได้สรุปขั้นตอนการสร้างและพัฒนาแบบวัดภาคปฏิบัติ ดังนี้

1. การวิเคราะห์งาน ได้แก่วิเคราะห์วัตถุประสงค์ วิเคราะห์ขั้นตอนการทำงาน ตรวจสอบความเที่ยงตรงเชิงเนื้อหา หากความสัมพันธ์ระหว่างพฤติกรรมที่ต้องการวัดในแต่ละขั้นตอน

2. ขึ้นสร้างเครื่องมือ
3. ขึ้นการทดลองใช้เครื่องมือ
4. ขึ้นวิเคราะห์หาคุณภาพ
5. ขึ้นนำเครื่องมือไปใช้

สมนึก ภักดิ์ทิษณี (2553 : 51-52) ได้กล่าวถึงขั้นตอนการสร้างแบบวัดผลงานภาคปฏิบัติ ไว้ดังนี้

1.วิเคราะห์งานและเขียนรายการ ประกอบด้วยขั้นตอนดังนี้

1.1 วิเคราะห์งานหรือเลือกงานที่เป็นตัวแทน โดยวิเคราะห์จุดมุ่งหมายที่ระบุไว้ในหลักสูตร (ถ้ามี) และรายละเอียดของงานที่มุ่งให้ผู้เรียนฝึก เพื่อค้นหาทักษะและความสามารถที่เกี่ยวข้องในกิจกรรมนั้น และเนื่องจากการวัดผลงานภาคปฏิบัติต้องใช้ในการสังเกต ดังนั้นทักษะ ที่มุ่งวัดควรเป็นสิ่งมองเห็นได้ในขณะสอบวัด และควรเป็นทักษะที่ยากกว่าทักษะที่ปฏิบัติ เป็นกิจกรรมที่นี้ควรคำนึงถึงข้อจำกัดเรื่องเวลาและวัสดุอุปกรณ์ที่ใช้ในการดำเนินการสอบวัดด้วย

1.2 กำหนดขั้นตอนการปฏิบัติงานที่จะวัด โดยทั่วไปจะประกอบด้วยขั้นเตรียมงาน ขั้นปฏิบัติงาน ขั้นผลงาน และ/หรือวัดเกี่ยวกับเวลาที่ใช้ในการปฏิบัติด้วย

1.3 เขียนข้อรายการ จะระบุรายละเอียดในแต่ละขั้นตอน เช่น ขั้นเตรียมงาน ใช้อุปกรณ์อะไรบ้าง ขั้นปฏิบัติงานทำอะไรบ้าง ขั้นผลงานจะพิจารณาอะไรบ้าง

1.4 ศึกษาตัวแปรที่ส่งผลทำให้การปฏิบัติงานนั้นมีคุณภาพแตกต่างกัน ทั้งนี้เพื่อให้เกิดความยุติธรรมแก่ผู้เข้าสอบวัดทุกคน เช่น จากตัวอย่างแบบวัดผลงานภาคปฏิบัติ เรื่องการดองไข่เค็ม ตัวแปรที่ส่งผลต่อคุณภาพงาน ได้แก่ ชนิดของไข่เค็ม (ไข่เป็ด ไข่ไก่ ไข่นก กระทา) คุณภาพของไข่ (ไข่เก่า ไข่ใหม่) ชนิดของเกลือ (เกลือสินเธาว์ เกลือทะเล เกลือไอโอดีน) เป็นต้น

1.5 จัดรูปแบบเครื่องมือ คือ เลือกลักษณะของแบบวัดว่า แต่ละตอนจะมีลักษณะอย่างไร เช่น เรื่องการดองไข่เค็ม อาจจะมีลักษณะดังนี้

การเตรียมงาน	ใช้แบบตรวจสอบรายการ
การปฏิบัติงาน	ใช้แบบมาตราส่วนประมาณค่า

เวลา	ใช้แบบมาตราส่วนประมาณค่า
ผลงาน	ใช้แบบมาตราส่วนประมาณค่า

2. กำหนดคะแนนสำหรับแต่ละส่วน อาจทำเป็น 2 ขั้นตอน คือ

2.1 กำหนดคะแนนสำหรับแต่ละส่วน เช่น ในเรื่องการดองไข่เค็ม

การเตรียมงาน	20 คะแนน
การปฏิบัติงาน	40 คะแนน
เวลา	10 คะแนน
ผลงาน	30 คะแนน
รวม	100 คะแนน

2.2 กำหนดน้ำหนักสำหรับแต่ละข้อรายการ โดยให้น้ำหนักของทุกๆข้อในขั้นตอนหนึ่งๆ รวมกันเท่ากับสัดส่วนคะแนนในขั้น 2.1 ทั้งนี้โดยคำนึงถึงความยากของงานและความสำคัญของกิจกรรมที่ต้องปฏิบัติ เช่น ในเรื่องการดองไข่เค็ม ขั้นการเตรียมงาน แบ่งเป็น 7 รายการ กำหนดน้ำหนักคะแนนเป็นรายการละ 3 คะแนน ใน 6 รายการแรก และเป็น 2 คะแนน ในรายการที่ 7 รวมทั้งหมดเป็น 20 คะแนน

2.3 กำหนดเกณฑ์การตัดสิน ต้องกำหนดเกณฑ์การตรวจสอบพฤติกรรมในการปฏิบัติหรือคุณภาพของงานในลักษณะที่มองเห็นได้ วัดได้ โดยเฉพาะเกณฑ์การผ่าน ผลงานภาคปฏิบัติในเรื่องนั้น ส่วนเกณฑ์การผ่านในแต่ละขั้นตอน อาจจะมีด้วยตามความเหมาะสม

2.4 จัดรูปแบบเครื่องมือ คือเรียบเรียงข้อรายการต่างๆ ตามขั้นตอน กำหนดเกณฑ์กำหนดคะแนนและ/หรือน้ำหนักเข้าเป็นหมวดหมู่ จัดรูปแบบให้สะดวกในการใช้ พร้อมทั้งกำหนดคะแนนเกณฑ์ในการผ่านในเรื่องนั้นๆ

อำนาจ ฟ้าสว่าง (2550 : 38) ได้เสนอขั้นตอนในการสร้างเครื่องมือวัดทักษะภาคปฏิบัติไว้ดังนี้

1. ขั้นการวิเคราะห์จุดประสงค์และวิเคราะห์งาน แบ่งได้ดังนี้

1.1 วิเคราะห์วัตถุประสงค์ เพื่อให้ทราบว่าพฤติกรรมที่มุ่งให้เกิดกับผู้ถูกวัดด้านใดบ้าง และมีพฤติกรรมอยู่ในระดับใด

1.2 วิเคราะห์งานเป็นขั้นที่ต้องทำการวิเคราะห์ให้ละเอียดและถูกต้องตามขั้นตอน ของงาน โดยแบ่งเป็นตอน ๆ ดังนี้

1.2.1 ขั้นเตรียม จะต้องสามารถบ่งบอกให้ได้ว่าผู้ที่จะต้องถูกวัดในงานภาคปฏิบัตินั้น ๆ จะต้องเตรียมอะไรบ้างก่อนที่จะลงมือปฏิบัติซึ่งขึ้นอยู่กับระดับของผู้เรียนว่าจำเป็นต้องเตรียมอะไรและมีบทบาทหน้าที่นั้น ๆ หรือไม่

1.2.2 ขั้นปฏิบัติ จะต้องบ่งบอกให้ได้ว่าลำดับขั้นในการปฏิบัติงานนั้น ๆ เริ่มต้นจากลำดับอะไรเรียงลำดับเป็นขั้นตอน ตั้งแต่ขั้นแรกไปจนถึงขั้นสุดท้ายและบอกได้ว่าขั้นใดที่จะเป็นพฤติกรรมหลัก แล้วยังสามารถบอกได้ว่าขั้นใดเป็นพฤติกรรมรองและพฤติกรรมเสริมและบอกได้ว่าแต่ละขั้นมีเกณฑ์ใดเป็นสิ่งที่สังเกตการณ์ปฏิบัติงานนั้น ๆ ว่าได้ผลหรือไม่ได้ผล ตามที่กำหนด

1.2.3 ขั้นผลงาน ผลการปฏิบัติจะต้องวิเคราะห์ให้ทราบว่า มีสิ่งใดบ้างที่จะเป็น ตัวบ่งบอกให้ทราบว่า ผลงานนั้นมีคุณภาพและมีประสิทธิภาพ ซึ่งอาจสังเกตจำนวนที่ผลิตได้ หรือสังเกตจากเวลาที่ทำหรือจากความประณีตและคุณภาพของงาน

1.2.4 ขั้นกิจนิสัย ในการปฏิบัติงานเป็นการวัดผลเกี่ยวกับพฤติกรรมทางด้านจิตพิสัย ซึ่งในขั้นนี้บางที่จะสอดแทรกอยู่ในขั้นผลการปฏิบัติ

1.3 ตรวจสอบความแม่นยำตรงการวิเคราะห์งานสามารถทำได้ 2 อย่าง ดังนี้คือ

1.3.1 กลุ่มผู้ชำนาญงาน หรือกลุ่มครูผู้สอนในวิชานั้นประมาณ 1-6 คน ผู้เชี่ยวชาญในงานพิจารณานั้นการวิเคราะห์งานนั้นถูกต้องและจำเป็นหรือไม่

1.3.2 ผู้เชี่ยวชาญในการพิจารณาว่าการวิเคราะห์นั้นถูกต้องจำเป็นหรือไม่

1.4 การหาความสัมพันธ์ระหว่างพฤติกรรมที่มุ่งจะให้เกิดกับงานในแต่ละขั้นตอนนี้จะทำให้ทราบว่างานในแต่ละขั้นคือนั้น มุ่งให้เกิดพฤติกรรมในระดับใดโดยยึดระดับพฤติกรรมในการวิเคราะห์ 3 ระดับคือ ระดับเลียนแบบ ระดับประสานความคิดกับอวัยวะที่เกี่ยวข้องหรือภายใต้การควบคุมทำตามแบบและระดับขั้นชำนาญ

2. ขั้นการสร้างเครื่องมือ

2.1 เขียนขั้นรายการ รายละเอียดในขั้นเตรียม ขั้นปฏิบัติ ขั้นผลงาน และขั้นกิจนิสัย ให้กะทัดรัดและเข้าใจตรงกันสำหรับผู้ที่จะใช้เครื่องมือนี้

2.2 เขียนหลักการในแต่ละข้อรายการ ซึ่งเกณฑ์ตัดสินจะสูงหรือต่ำขึ้นอยู่กับระดับพฤติกรรมที่มุ่งเน้นให้เกิดในแต่ละขั้นตอน

2.3 กำหนดคะแนนในแต่ละข้อรายการ โดยพิจารณาจาก ข้อรายการว่าข้อใดเป็นพฤติกรรมหลัก น้ำหนักคะแนนก็ควรจะมากด้วย ข้อใดเป็นพฤติกรรมรอง น้ำหนักคะแนนก็ลดลงตามลำดับด้วย

3. ขั้นทดลองใช้เครื่องมือ เมื่อสร้างเครื่องมือเสร็จแล้วจำเป็นต้องมีการทดลองเครื่องมือ เพื่อที่จะได้นำไปหาคุณภาพของเครื่องมือซึ่งการทดลองอาจทำได้โดย

3.1 ให้กลุ่มผู้เชี่ยวชาญหลายคนทดลองใช้

3.2 ให้กลุ่มผู้สอนหรือผู้ที่เกี่ยวข้องกับงานนั้น ๆ ใช้

4. ขั้นวิเคราะห์หาคุณภาพ นำผลที่ได้จากการทดลองครั้งที่ 3 มาคำนวณหาค่าความเชื่อถือของเครื่องมือและหาเกณฑ์ปกติของเครื่องมือ

5. ขั้นนำเครื่องมือไปใช้หลังจากวิเคราะห์คุณภาพของเครื่องมือแล้ว

จากที่หลาย ๆ ท่านได้เสนอขั้นตอนในการสร้างเครื่องมือวัดทักษะภาคปฏิบัติดังกล่าวข้างต้นนั้น ผู้วิจัยได้สรุปเพื่อวิเคราะห์งานเป็นขั้นที่ต้องทำ เพื่อดำเนินการสร้างเครื่องมือวัดผลภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานเกษตร) สำหรับชั้นประถมศึกษาปีที่ 5 โดยผู้วิจัยได้ดำเนินการสร้างตามขั้นตอนของบุญชม ศรีสะอาด เป็นแนวทางในการสร้าง และใช้หลักการในการวิเคราะห์งานตามขั้นตอนของการวัดผลทักษะภาคปฏิบัติตามแนวคิดของ สมนึก ภัททิยธนี, สุรัชย์ สิงห์แผ่น และอำนาจ ฟ้าสว่าง ผู้วิจัยได้จัดรูปแบบการวัดเป็นขั้นตอน คือ ขั้นเตรียม ขั้นปฏิบัติงาน ขั้นผลงาน และกิจนิสัย

3. การหาคุณภาพและการตรวจสอบคุณภาพของเครื่องมือ

การหาคุณภาพของแบบวัดถือว่าเป็นสิ่งสำคัญยิ่ง เมื่อสร้างแบบวัดแล้วจึงจำเป็นต้องตรวจสอบคุณภาพของแบบวัดเพื่อให้ทราบว่าแบบวัดนั้นมีคุณภาพเพียงใด มีคุณสมบัติเหมาะสมสำหรับใช้ในการเก็บรวบรวมข้อมูลวิจัยหรือไม่ ถ้าแบบวัดหรือเครื่องมือที่สร้างขึ้นไม่มีคุณภาพจะไม่สามารถวัดได้ตรงกับสิ่งที่ต้องการวัดและเป็นข้อมูลที่เชื่อถือไม่ได้ จะทำให้การวัดนั้นไม่มีประโยชน์ สิ่งที่มีความสำคัญของแบบวัด คือ ความเที่ยงตรง อำนาจจำแนก ความเชื่อมั่น ดังนี้

3.1 ความเที่ยงตรง (Validity)

ความหมายของความเที่ยงตรง มีนักวิชาการให้ความหมายไว้ดังนี้

ศิริชัย กาญจนวาสี (2544 : 73) ได้ให้ความหมายของความเที่ยงตรงว่าเป็นความถูกต้องแม่นยำของแบบวัดในการวัดสิ่งที่ต้องการจะวัด (ขวัญยืน มูลศรี, 2548 : 50) ได้ให้ความหมายของความเที่ยงตรงว่าเป็นความสามารถของแบบวัดที่สามารถวัดได้ตรงตามลักษณะของสิ่งที่ต้องการจะวัด

ไพศาล วรคำ (2554 : 260) กล่าวว่า ความเที่ยงตรงคือความถูกต้องแม่นยำของเครื่องมือในการวัดสิ่งที่ต้องการวัด ความเที่ยงตรงจึงเปรียบเสมือนหัวใจของการทดสอบ

สมนึก ภัททิยธนี (2553) กล่าวถึงความเที่ยงตรงไว้ดังนี้ ความเที่ยงตรง (Validity) หมายถึง คุณภาพของแบบทดสอบ ที่สามารถวัดได้ตรงกับจุดมุ่งหมายที่ต้องการ หรือวัดในสิ่งที่ต้องการวัด ได้อย่างถูกต้องแม่นยำ ความเที่ยงตรงจึงเปรียบเสมือนหัวใจของการทดสอบ

จากความหมายดังกล่าวข้างต้นสรุปได้ว่า ความเที่ยงตรง หมายถึง คุณภาพของแบบทดสอบ ที่สามารถวัดในสิ่งที่ต้องการวัดได้ตรงกับจุดมุ่งหมายที่ต้องการ

การตรวจสอบคุณภาพความเที่ยงตรง

การหาค่าความเที่ยงตรงของแบบทดสอบแยกเป็นแบบอิงกลุ่มและแบบอิงเกณฑ์ ดังนี้ (สมนึก ภัททิยธนี. 2553 : 217)

1. การหาค่าความตรงของแบบทดสอบอิงกลุ่ม วิธีหาค่าความเที่ยงตรง ที่นิยมใช้แบ่งเป็น 4 วิธี คือ

- 1.1 ความตรงตามเนื้อหา
- 1.2 ความเที่ยงตรงเชิงโครงสร้าง
- 1.3 ความเที่ยงตรงตามสภาพ
- 1.4 ความเที่ยงตรงตามการพยากรณ์

2. การหาค่าความตรงของแบบทดสอบอิงเกณฑ์ วิธีหาค่าความเที่ยงตรง ที่นิยมใช้แบ่งเป็น 2 วิธี คือ

2.1 ความเที่ยงตรงตามเนื้อหา แบ่งเป็น 2 กรณี คือ

2.1.1 อาศัยดุลยพินิจของผู้เชี่ยวชาญด้านเนื้อหาและด้านการวัดผล

2.1.2 อาศัยเทคนิคการตรวจสอบจากการทดลองหรือเทคนิคเชิงประจักษ์

แต่ที่นิยมใช้กันมากคือให้ผู้เชี่ยวชาญด้านเนื้อหาและด้านการวัดผลเป็นผู้พิจารณา ซึ่งจะพิจารณาสิ่งต่อไปนี้

1. จุดประสงค์เชิงพฤติกรรมที่กำหนดไว้ ครอบคลุมเนื้อหาหรือไม่

2. ข้อสอบที่จะวัดแต่ละข้อ วัดได้ตรงตามจุดประสงค์เชิงพฤติกรรมหรือไม่

ซึ่งทั้ง 2 กรณี โรวินเนลลี (Rovinelli) และแฮมเบิลตัน (Hambleton) ได้เสนอ

วิธีการพิจารณาเรียกว่า ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์เชิงพฤติกรรม

(IOC : Index of Item Objective Congruence)

2.2 ความเที่ยงตรงตาม โครงสร้าง เมื่อข้อสอบแบบอิงเกณฑ์ได้รับการตรวจสอบคุณภาพแล้วว่า มีความเที่ยงตรงตามเนื้อหาที่จะพิจารณาต่อไปอีกว่า แบบทดสอบฉบับนั้นทั้งฉบับ ซึ่งวัตถุประสงค์เดียวมีหลายข้อ มีความเที่ยงตรงต่อการสอบผ่านและไม่ผ่านของนักเรียนก่อนเรียนและหลังเรียนเรื่องนั้นหรือไม่ นั่นคือถ้านักเรียนที่เรียนรู้ได้ครบถ้วนตามจุดประสงค์ทำข้อสอบได้ถูกต้อง ก็แสดงว่าแบบทดสอบชุดนั้นมีความเที่ยงตรงตามโครงสร้าง

ในการหาค่าความเที่ยงตรงตามโครงสร้างมีหลายวิธี แต่ที่นิยมใช้คือวิธีของคาร์เวอร์ (Carver Method) โดยยึดแนวความคิดที่ว่า ผู้ที่เรียนแล้วน่าจะสอบผ่าน ผู้ที่ยังไม่ได้เรียนน่าจะสอบไม่ผ่าน วิธีการทำจะเริ่มจาก นำผลการสอบมาจัดลงในตาราง ดังนี้

	กลุ่มที่ยังไม่ได้เรียน	กลุ่มที่เรียนแล้ว
สอบผ่าน	b	a
สอบไม่ผ่าน	c	d

ความเที่ยงตรงของเครื่องมือเป็นสิ่งสำคัญที่สุดในการวิจัย เพราะถ้าเครื่องมือขาดคุณลักษณะด้านนี้แล้ว ย่อมทำให้การวัด ไม่ตรงกับจุดประสงค์ การหาค่าความเที่ยงตรงของแบบวัดจำแนกได้ แบ่งออกเป็น 3 ประเภท (ไพศาล วรคำ. 2554 : 260 - 266)

1. ความเที่ยงตรงเชิงเนื้อหา (Content Validity) หมายถึงคุณลักษณะของเครื่องมือที่สามารถวัดได้ตรงตามเนื้อหาที่จะวัด หรือเป็นดัชนีที่บ่งบอกว่าเนื้อหาของเครื่องมือหรือเนื้อหา ของข้อคำถามวัดได้ตรงตามเนื้อหาของเรื่องที่ต้องการวัด ดังนั้นประเด็นสำคัญของความเที่ยงตรงเชิงเนื้อหาจึงอยู่ที่การเลือกใช้กลุ่มตัวอย่างเนื้อเรื่องที่เป็นตัวแทน (Representative Sample) ของมวเนื้อเรื่องที่ต้องการวัด ว่าเป็นตัวแทนของเนื้อหาทั้งหมดและมีความเพียงพอ (Adequate) ต่อการวัดเนื้อเรื่องนั้นหรือไม่ การตรวจสอบความเที่ยงตรงเชิงเนื้อหาจึงอาศัยกระบวนการตรวจสอบ โดยกลุ่มผู้เชี่ยวชาญที่เป็นอิสระจากกัน ช่วยพิจารณาตัวอย่างเนื้อเรื่องในเครื่องมือวัดว่ามีขอบเขต ที่ครอบคลุมและเป็นตัวแทนมวลเนื้อเรื่องที่ต้องการวัดเพียงใด

2. ความเที่ยงตรงตามเกณฑ์สัมพันธ์ (Criterion-related Validity) เป็นความสอดคล้องสัมพันธ์กันระหว่างคะแนนจากเครื่องวัดที่ผู้วิจัยสร้างขึ้นกับเกณฑ์ภายนอก (Criterion) ที่สามารถใช้วัดคุณลักษณะที่ต้องการนั้นได้ เกณฑ์ภายนอกนี้อาจเป็นคะแนนจากแบบวัดอื่น หรือวิธีการอื่นๆที่วัดสภาพปัจจุบันหรือสภาพในอนาคตของกลุ่มตัวอย่างได้ตรงตามคุณลักษณะที่ต้องการวัด ความเที่ยงตรงเกณฑ์สัมพันธ์แบ่งออกได้เป็น 2 ประเภทคือ

2.1 ความเที่ยงตรงเชิงสภาพ หรือความเที่ยงตรงร่วมสมัย (Concurrent Validity) หมายถึง ความสอดคล้องสัมพันธ์กันระหว่างคะแนนที่ได้จากแบบวัดที่สร้างขึ้นกับคะแนนที่ได้จากแบบวัดอื่นๆที่กำหนดไว้แล้วในช่วงเวลาเดียวกัน หรือวิธีการอื่นๆที่วัดสภาพปัจจุบันของกลุ่มตัวอย่าง

2.2 ความเที่ยงตรงเชิงพยากรณ์ (Predictive Validity) หมายถึง ความสามารถของเครื่องมือที่จะบ่งบอกผลที่วัดในขณะนั้น ได้ถูกต้องตามสภาพที่แท้จริงในอนาคต โดยอาศัยความสัมพันธ์ระหว่างคะแนนของเครื่องมือกับคะแนนเกณฑ์สัมพันธ์ซึ่งจะปรากฏในอนาคต เช่นแบบทดสอบวัดความถนัดทางการเรียนที่สร้างขึ้นเพื่อทำนายผลการเรียนในอนาคต ก็อาจใช้คะแนนเฉลี่ยสะสมปีสุดท้ายเป็นเกณฑ์สัมพันธ์ ซึ่งการคำนวณหาความเที่ยงตรงเชิงพยากรณ์ต้องเสียเวลาของผู้วิจัยอาจเลือกแบบทดสอบที่มีผู้สร้างไว้แล้วและมีความเที่ยงตรงเชิงพยากรณ์สูงมาเป็นเครื่องมือเปรียบเทียบ โดยนำเครื่องมือดังกล่าวมาวัดกับกลุ่มตัวอย่าง พร้อมกับแบบวัดที่ผู้วิจัยสร้างขึ้น แล้วหาความสัมพันธ์ระหว่างคะแนนระหว่างแบบวัดทั้งสองฉบับ

3. ความเที่ยงตรงเชิงทฤษฎีหรือความเที่ยงตรงเชิงโครงสร้าง (Construct) หมายถึง ความสามารถของเครื่องมือที่สามารถวัดได้ตรงตามขอบเขต หรือครบตามคุณลักษณะย่อย ๆ ของ สิ่งที่ต้องการวัดที่ระบุไว้ในทฤษฎีเกี่ยวกับคุณลักษณะนั้น ๆ ซึ่งโดยทั่วไปตัวแปรที่เป็นคุณลักษณะ (Trait) มักจะมีโครงสร้างขององค์ประกอบในเชิงทฤษฎี บางทีจึงถูกเรียกว่า ความเที่ยงตรงเชิงโครงสร้าง การหาความเที่ยงตรงเชิงทฤษฎีนิยมใช้กับเครื่องมือวัดตัวแปรคุณลักษณะหรือตัวแปรแฝงที่มีการนิยามเชิงทฤษฎี เช่น เซวรณ์ปัญญา เจตคติ ความเชื่อ ค่านิยม เซวรณ์อารมณ์ เป็นต้น โดยคุณลักษณะเหล่านี้สังเกตโดยตรงไม่ได้ จะสังเกตได้เฉพาะผลที่เกิดขึ้นเท่านั้น การตรวจสอบความเที่ยงตรงเชิงทฤษฎีหรือความเที่ยงตรงเชิงโครงสร้างสามารถดำเนินการได้หลากหลายวิธี เช่น

3.1 วิธีตัดสิน โดยผู้เชี่ยวชาญ เป็นการให้กลุ่มผู้เชี่ยวชาญตรวจสอบถึงความเหมาะสมของทฤษฎีที่นำมาใช้นิยาม ผังข้อคำถามและคุณภาพของข้อคำถาม ซึ่งเป็นหลักฐานเบื้องต้นที่นำมาใช้สนับสนุนความเที่ยงตรงเชิงทฤษฎี ผู้วิจัยจึงต้องนำเสนอทฤษฎีและนิยาม

ของคุณลักษณะ ที่ต้องการวัด ผังข้อคำถามของแบบวัดและแบบวัดนั้นต่อผู้เชี่ยวชาญ เพื่อทำการตรวจสอบทฤษฎี นิยาม โครงสร้าง องค์ประกอบของคุณลักษณะที่มุ่งวัดนั้นว่ามีความเหมาะสมสอดคล้องกับทฤษฎีอื่น อันเป็นที่ยอมรับกันทั่วไปหรือไม่ เพียงไร ผังข้อคำถามมีความครอบคลุมและเป็นตัวแทนของคุณลักษณะอันพึงประสงค์ที่ต้องการวัด ได้ดีเพียงไร คุณภาพการเขียนข้อคำถามแต่ละข้อเป็นไปตามผังข้อคำถามหรือไม่ และสามารถวัดได้ตรงกับคุณลักษณะที่ต้องการวัด ได้ดีเพียงไร ถ้าผู้เชี่ยวชาญมีความเห็นว่าข้อคำถามนั้นสามารถวัดคุณลักษณะที่ต้องการวัด ได้สูงกว่า 80 % ของความคิดเห็นของผู้เชี่ยวชาญทั้งหมด แสดงว่าข้อคำถามนั้นใช้ได้

3.2 วิธีเปรียบเทียบคะแนนระหว่างกลุ่มรู้จัก (Comparing the of Known Groups) หากคุณลักษณะที่ต้องการวัดนั้นมีความแตกต่างกันระหว่างกลุ่มบุคคลอย่างเห็นได้ชัด หรือผู้วิจัยทราบแน่ชัดว่า คุณลักษณะที่ต้องการวัดนั้นมีในกลุ่มบุคคลกลุ่มหนึ่ง และ ไม่มีในกลุ่มบุคคลอีกกลุ่มหนึ่ง เช่น กลุ่มที่ประสบความสำเร็จกับกลุ่มที่ไม่ประสบความสำเร็จ หรือกลุ่มที่มีประสบการณ์กับกลุ่มที่ไม่มีประสบการณ์ เป็นต้น การเปรียบเทียบคะแนนที่วัดได้ระหว่างกลุ่มที่ทราบแน่ชัดแล้วว่า มีคุณลักษณะที่ต้องการวัดแตกต่างกัน (Known Groups) ก็จะเป็นหลักฐานส่วนหนึ่งที่ใช้สนับสนุนความเที่ยงตรงเชิงทฤษฎีได้ โดยถ้าเครื่องมือสามารถวัดคุณลักษณะที่สนใจนั้น ได้จริง ผลการวัดจะต้องมีความแตกต่างระหว่างกลุ่ม การเปรียบเทียบคะแนนระหว่างกลุ่มนี้อาจใช้วิธีการทางสถิติเพื่อทดสอบความแตกต่างระหว่างกลุ่ม เช่น การทดสอบที (t-test) การวิเคราะห์ความแปรปรวน หรือการทดสอบไคสแควร์ เป็นต้น

3.3 วิธีการเปรียบเทียบคะแนนจากการทดลอง (Comparing the Scores from an Experiment) โดยทั่วไปทฤษฎีต่างๆ จะสามารถพยากรณ์หรือคาดการณ์ผลที่จะตามมาจากปรากฏการณ์ใดๆ ได้ หรือหากมีการเปลี่ยนแปลงเงื่อนไขของการจัดกระทำตามการทดลองก่อนและหลังการได้จัดกระทำกับตัวแปรทดลองแล้ว เช่น ตามทฤษฎีคาดหมายได้ว่า คะแนนความวิตกกังวลของบุคคลจะแปรเปลี่ยนไปตามสถานการณ์ที่เผชิญ ถ้าสร้างสถานการณ์ให้เกิดความวิตกกังวลในระดับต่างๆ ให้กับกลุ่มตัวอย่าง แล้วใช้แบบวัดความวิตกกังวลก็น่าจะได้คะแนนความวิตกกังวลที่ระดับต่างๆ กันตามสภาพการณ์ที่สร้างขึ้น แบบวัดที่สามารถให้คะแนนการวัดได้สอดคล้องกับสถานการณ์ที่ทดลองตามความคาดหมายของทฤษฎี ก็จะมี ความเที่ยงตรงเชิงทฤษฎี

3.4 วิธีวิเคราะห์เมตริกซ์ลักษณะหลากหลาย (Multi-trait Multi-method Matrix : MTMM) เป็นการตรวจสอบความเที่ยงตรงเชิงทฤษฎีที่อาศัยการวิเคราะห์ความสัมพันธ์ระหว่าง

การวัดหลาย ๆ ลักษณะ (Multi-trait) โดยใช้วิธีการวัดหลายๆ วิธีหรือแบบวัดหลายๆ ชุด (Multi-method) โดยมุ่งตรวจสอบความเหมาะสมของเครื่องมือหลายๆ ชุดในการวัดลักษณะใดลักษณะหนึ่งที่สนใจศึกษา

จากที่กล่าวมาผู้วิจัยได้ใช้การตรวจสอบคุณภาพความเที่ยงตรงของแบบวัดทักษะภาคปฏิบัติ ดังนี้

1. ค่าความเที่ยงตรงตามเนื้อหา โดยอาศัยดุลยพินิจของผู้เชี่ยวชาญด้านเนื้อหา และ ด้านการวัดผล โดยใช้สูตร IOC หาค่าเฉลี่ยดัชนีความสอดคล้องของผู้เชี่ยวชาญทั้งหมดตามวิธีของ โรวินเนลลี และแฮมเบิลตัน (Rovinelli and Hambleton) อ้าง โดย (สมนึก ภัททิยธนี. 2553 : 220) ดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ IOC แทน ดัชนีความสอดคล้องระหว่างข้อสอบกับจุดประสงค์
 $\sum R$ แทน ผลรวมคะแนนความคิดเห็นของผู้เชี่ยวชาญทั้งหมด
 N แทน จำนวนผู้เชี่ยวชาญทั้งหมด

2. ค่าความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) เพื่อตรวจสอบความเที่ยงตรงเชิงโครงสร้าง (Construct Validity) ของแบบวัดทักษะการปฏิบัติ ใช้สถิติ t-test แบบ Independent Samples (สุรวาท ทองบุ. 2550 : 129-130)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

เมื่อ t แทน ค่าที่ใช้พิจารณาในการแจกแจงแบบที่

X_1	แทน	ค่าเฉลี่ยของกลุ่มที่มีคุณลักษณะตามที่ต้องการศึกษา
X_2	แทน	ค่าเฉลี่ยของกลุ่มที่ไม่มีคุณลักษณะตามที่ต้องการศึกษา
S_1^2	แทน	ความแปรปรวนของกลุ่มที่มีคุณลักษณะตามที่ต้องการศึกษา
S_2^2	แทน	ความแปรปรวนของกลุ่มที่ไม่มีคุณลักษณะตามที่ต้องการศึกษา
n_1	แทน	จำนวนคนในกลุ่มที่มีคุณลักษณะตามที่ต้องการศึกษา
n_2	แทน	จำนวนคนในกลุ่มที่ไม่มีคุณลักษณะตามที่ต้องการศึกษา

ถ้าค่า t ที่คำนวณได้มีนัยสำคัญทางสถิติที่ระดับ .05 แสดงว่าเครื่องมือที่มีความเที่ยงตรงเชิงโครงสร้างสูง สามารถนำไปใช้วัดคุณลักษณะที่ต้องการศึกษาจากกลุ่มตัวอย่างจริงได้

3.2 อำนาจจำแนก (Discrimination Index)

มีนักการศึกษาให้ความหมายอำนาจจำแนกไว้ดังนี้

ล้วน สายยศ, และอังคณา สายยศ (2543 : 299) ได้ให้ความหมายของค่าอำนาจจำแนกว่าเป็นความสามารถในการแยกลักษณะของคน 2 กลุ่ม นั่นคือแยกคนที่มีคุณลักษณะนั้นสูงกับคน ที่มีคุณลักษณะนั้นต่ำ

บุญเชิด ภัฏโณนันทพงษ์ (มปป. : 135) ให้ความหมายของอำนาจจำแนกว่าเป็นค่าสหสัมพันธ์ระหว่างคะแนนรวมของการตอบแบบทดสอบฉบับหนึ่งกับคะแนนของการตอบข้อสอบแต่ละข้อในแบบทดสอบนั้น

สมนึก ภัททิยชนี (2553 : 71) ได้ให้ความหมายของค่าอำนาจจำแนกว่า อำนาจจำแนกหมายถึง ความสามารถของข้อสอบในการจำแนกผู้สอบที่มีคุณลักษณะหรือความสามารถแตกต่างกันออกจากกันได้

ไพศาล วรคำ (2554 : 293) กล่าวว่า อำนาจจำแนกหมายถึงคุณลักษณะของข้อสอบหรือข้อคำถามที่สามารถแยกปริมาณของคุณลักษณะที่ต้องการวัดที่มีอยู่ในแต่ละบุคคลได้

จากความหมายของอำนาจจำแนกดังกล่าวข้างต้น สรุปได้ว่าอำนาจจำแนก เป็นคุณสมบัติของแบบวัดที่สามารถจำแนกความสามารถที่แตกต่างของคนสองกลุ่มออกจากกันได้

การตรวจสอบคุณภาพอำนาจจำแนก

เครื่องมือที่นิยมหาค่าอำนาจจำแนกได้แก่ แบบทดสอบ และแบบสอบถาม ซึ่งมีหลายวิธีจำแนกตามลักษณะของเครื่องมือดังนี้ (ไพศาล วรคำ. 2554 : 294 -299)

1. การหาค่าอำนาจจำแนกแบบอิงกลุ่ม ต้องการจำแนกบุคคลออกตามระดับความสำคัญ

1.1 เทคนิคร้อยละ 50 อาศัยแบ่งผู้สอบออกเป็น 2 กลุ่ม เท่ากัน

1.2 เทคนิค 27% ในการใช้เทคนิคร้อยละ 50 ถ้าผู้สอบจำนวนมากค่อนข้างยุ่งยากในการวิเคราะห์ข้อมูลจากผู้สอบทั้งหมด การนำคะแนนกลุ่มสูงกลุ่มต่ำมาวิเคราะห์เพียงร้อยละ 27 ก็จะได้ค่าประมาณที่ใกล้เคียง

1.3 การหาค่าสหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม

1.4 การหาสหสัมพันธ์แบบ Point Biserial เป็นการประยุกต์การหาสัมประสิทธิ์สหสัมพันธ์ของตัวแปรแบ่งสองแท้ เนื่องจากคะแนนรายข้อนั้นมีค่าสองค่าคือ 0 และ 1

2. การหาค่าอำนาจจำแนกแบบอิงเกณฑ์ ต้องการจำแนกกลุ่มบุคคลที่รอบรู้หรือไม่ผ่านเกณฑ์ออกจากกลุ่มบุคคลที่ไม่รอบรู้หรือไม่ผ่านเกณฑ์

2.1 ดัชนีอำนาจจำแนกของเบรนนาน (Brennan's Index : B-Index) เป็นการหาค่าอำนาจจำแนกแบบอิงเกณฑ์ตามแนวคิดการสอบครั้งเดียว แล้วพิจารณาความสามารถของข้อสอบในการแยกคนกลุ่มผ่านเกณฑ์กับไม่ผ่านเกณฑ์ออกจากกัน โดยหาค่าอำนาจจำแนกได้จากความแตกต่างระหว่างสัดส่วนของผู้ผ่านเกณฑ์ตอบถูกกับสัดส่วนของผู้ไม่ผ่านเกณฑ์ตอบถูก

2.2 ดัชนีความไวของข้อสอบ (Sensitive Index : S) เป็นการหาอำนาจจำแนกแบบอิงเกณฑ์ตามแนวคิดการสอบสองครั้ง โดยสอบก่อนเรียนและหลังเรียน ข้อสอบที่มีอำนาจจำแนก จะสามารถแยกความรอบรู้ของผู้สอบก่อนเรียนและหลังเรียนได้ดี โดยก่อนเรียนผู้เรียนไม่ควร จะตอบถูก เพราะยังไม่มีความรู้ ส่วนหลังเรียนผู้เรียนควรตอบถูก เพราะเกิดการเรียนรู้ หรือมีความรอบรู้แล้ว การที่ข้อสอบสามารถตรวจสอบความรอบรู้ที่เพิ่มขึ้นได้ดี นิยมเรียกว่า ความไวของข้อสอบ (Sensitivity)

3. การหาอำนาจจำแนกของแบบสอบอัตนัย ค่าคะแนนในแต่ละข้อจะมีได้หลายค่า การหาค่าอำนาจจำแนกของแบบสอบอัตนัยสามารถหาได้จากสูตรของวิทนีและซาเบอร์ส (Whitney and Sabers)

4. การหาอำนาจจำแนกของแบบสอบถาม

วิธีการหาอำนาจจำแนกที่กล่าวมาเป็นการหาอำนาจจำแนกของแบบทดสอบ ซึ่งมีการตรวจให้คะแนนแบบ 0-1 สำหรับแบบสอบถามซึ่งมักจะมีคำตอบเป็นมาตรฐานค่า (Rating Scale) จึงมีวิธีการหาอำนาจจำแนกแตกต่างไปจากแบบทดสอบ แต่การหาค่าอำนาจจำแนกของแบบทดสอบบางวิธีก็สามารถนำมาใช้กับแบบสอบถามได้เหมือนกัน เช่น การหาสหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม (Item – total Correlation) อีกวิธีหนึ่งที่นิยมใช้ในการหาค่าอำนาจจำแนกของแบบสอบถามก็คือ การทดสอบนัยสำคัญทางสถิติระหว่างกลุ่มสูงกับกลุ่มต่ำด้วยสถิติทดสอบที (t-test) โดยคัดผู้ได้คะแนนสูงสุดและต่ำสุดมาร้อยละ

25-30 ของจำนวนผู้ที่ตอบแบบสอบถามทั้งหมด แล้วนำคะแนนของกลุ่มสูงและกลุ่มต่ำแต่ละข้อมาเปรียบเทียบค่าเฉลี่ย ถ้าค่าเฉลี่ยของทั้งสองกลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติก็แสดงว่าข้อคำถามนั้นมีอำนาจจำแนก แต่ถ้าไม่มีนัยสำคัญทางสถิติ หรือค่าสถิติที่ได้มีค่าเป็นลบ ก็แสดงว่าข้อคำถามนั้นมีอำนาจจำแนกค่าหรือไม่สามารถจำแนกคุณลักษณะที่ต้องการวัดได้ การทำการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มด้วยวิธีนี้จะไม่สามารถบอกค่าดัชนีอำนาจจำแนกได้ บอกได้เพียงว่าข้อคำถามนั้นมีอำนาจจำแนกหรือไม่เท่านั้น การหาค่าจำแนกในวิธีนี้อาจทำได้โดยอาศัยความสัมพันธ์ระหว่างสถิติทดสอบที่กับค่าสัมประสิทธิ์สหสัมพันธ์

จากที่กล่าวมาข้างต้น การหาค่าอำนาจจำแนกสามารถหาได้หลายวิธี ซึ่งผู้วิจัยได้ตรวจสอบคุณภาพของแบบวัดทักษะภาคปฏิบัติ ที่มีลักษณะค่าของคะแนนในแต่ละข้อมีหลายค่า ผู้วิจัยจึงเลือกวิธีการหาค่าอำนาจจำแนกตามลักษณะการตรวจให้คะแนนแบบข้อสอบอัตนัย โดยแบ่งผู้เข้าสอบออกเป็นสองคือกลุ่มสูงและกลุ่มต่ำด้วยเทคนิค 25% และหาค่าอำนาจจำแนก โดยใช้สูตร ของวิทนีและซาเบอร์ส (Whitney and Sabers) อ้างโดย (ไพศาล วรรคมา. 2554 : 302)

$$D = \frac{S_H - S_L}{N(X_{max} - X_{min})}$$

เมื่อ	D	แทน	ค่าอำนาจจำแนกของแบบวัด
	S_H	แทน	ผลรวมคะแนนในกลุ่มสูง
	S_L	แทน	ผลรวมคะแนนในกลุ่มต่ำ
	N	แทน	จำนวนคนในกลุ่มสูงหรือกลุ่มต่ำ
	X_{max}	แทน	คะแนนสูงสุดในข้อนั้น
	X_{min}	แทน	คะแนนต่ำสุดในข้อนั้น

3.3 ความเชื่อมั่น (Reliability)

มีนักการศึกษาหลายท่านให้ความหมายของความเชื่อมั่นไว้ดังนี้

วิญญา วิศาลภรณ์ (2540 : 120) ให้ความหมายของความเชื่อมั่นว่า หมายถึง ความมั่นคงความสม่ำเสมอหรือความคงเส้นคงวาของผลจากการวัด

ศิริชัย กาญจนวาสิ (2544 : 120) ให้ความหมายของความเชื่อมั่นว่า หมายถึง เป็นความคงที่หรือความคงที่คงวาของผลที่ได้จากการวัดซ้ำ ถ้าต้องการวัดสิ่งเดียวกันหลาย ๆ ครั้ง ถ้าได้ค่าค่อนข้างคงเส้นคงวาสูงขึ้นถือว่าแบบวัดมีความเชื่อมั่นมากขึ้น

สมนึก ภักดิ์ทิษณี (2553 : 69) ให้ความหมายของความเชื่อมั่นว่า หมายถึง ลักษณะของแบบทดสอบทั้งสองฉบับที่สามารถวัดได้คงเส้นคงวาไม่เปลี่ยนแปลง ไม่ว่าจะทำการสอบใหม่กี่ครั้งก็ตาม

ไพศาล วรคำ (2554 : 272) ให้ความหมายของความเชื่อมั่นว่า หมายถึง ความคงที่ของผลที่ได้จากการวัดด้วยเครื่องมือชุดใดชุดหนึ่งในการวัดหลาย ๆ ครั้ง

จากความหมายของความเชื่อมั่นดังกล่าวข้างต้น สรุปได้ว่า ความเชื่อมั่นหมายถึง ความคงเส้นคงวาของการวัด ไม่ว่าจะวัดซ้ำกี่ครั้งก็มีค่าเหมือนเดิม

การตรวจสอบคุณภาพความเชื่อมั่น

การตรวจสอบค่าความเชื่อมั่น ของแบบวัดเริ่มพัฒนามาจากนิยาม คือ เป็น ความสัมพันธ์ระหว่างค่าการวัดหลายๆ ครั้ง ภายใต้แนวคิด 3 แนวคิด ดังนี้ (ไพศาล วรคำ, 2544 : 273 – 278)

1. การวัดความคงที่ (Measure of Stability) เป็นหารหาความเชื่อมั่นจากการสอบซ้ำ (Test – retest) โดยการหาความสัมพันธ์ระหว่างคะแนนสอบครั้งแรกและครั้งที่สอง ที่ทำการทดสอบกับกลุ่มตัวอย่างเดิม ด้วยเครื่องมือฉบับเดิม โดยทิ้งช่วงระยะห่างในการสอบทั้งสองครั้งพอประมาณเพื่อให้ผู้สอบจำข้อสอบไม่ได้ และไม่ให้เรียนรู้เพิ่มเติม จากนั้นนำคะแนนมาหาค่าสัมประสิทธิ์ของความคงที่ (Coefficient of Stability) ของคะแนนการสอบซ้ำ

2. การวัดความสมมูลกัน (Measure of Equivalence) เพื่อแก้ปัญหาของการสอบซ้ำจึงใช้เครื่องมือสองฉบับที่คล้ายกัน หรือคู่ขนานกัน (Parallel Test) มาใช้แทน โดยข้อความของแบบทดสอบทั้ง 2 ฉบับ มีลักษณะเป็นข้อความที่สมมูลกัน (Equivalent) คือ มีจำนวนข้อเท่ากัน วัดในเรื่องเดียวกันและมีค่าสถิติของข้อสอบแต่ละคู่ (ข้อที่สมมูลกันระหว่างแบบสอบสองฉบับ) เท่ากัน นำมาทดลองใช้กับกลุ่มตัวอย่างหนึ่งในเวลาเดียวกัน จากนั้นนำคะแนนจากแบบสอบทั้งสองฉบับมาหาความสัมพันธ์กัน

3. การวัดความสอดคล้องภายใน (Measure of Internal Consistency) เป็นการประมาณค่าความเชื่อมั่นจากการทดลองใช้เครื่องมือเพียงครั้งเดียว ด้วยแบบวัดฉบับเดียว และวัดกับกลุ่มตัวอย่างเดียว แนวคิดการวัดความสอดคล้องภายในนี้พัฒนามาจากการวัดความสมมูลกันโดยมีความเชื่อว่า เมื่อเครื่องมือวัดในสิ่งเดียวกัน และสร้างข้อคำถามให้มี

ลักษณะสมมูลกัน การแบ่งเครื่องมือหรือแบบวัดออกเป็นสองส่วนก็จะเหมือนกับแบบวัดคู่ขนาน ทำให้เกิดการหาความเชื่อมั่นด้วยวิธีแบ่งครึ่งข้อสอบ (Split-half) ขึ้น ต่อจากนั้นก็ยังสามารถแนวคิดออกไปอีกว่า หากข้อคำถามมีลักษณะสมมูลกัน การทำข้อสอบแต่ละข้อก็เหมือนกับการวัดหนึ่งครั้ง หากมีข้อคำถามจำนวน k ข้อ ก็เหมือนการวัดซ้ำจำนวน k ครั้ง ความสัมพันธ์ระหว่างผลการวัดทุกข้อคำถามในแบบวัด จึงน่าจะบ่งบอกความเชื่อมั่นของแบบวัดได้ ดังนั้นการพิจารณาความสอดคล้องภายในของแบบวัด หรือความเป็นเอกพันธ์ (Homogeneity) จึงเป็นการหาความเชื่อมั่นอีกแบบหนึ่ง

จากที่กล่าวมาแล้วสรุปได้ว่า การตรวจสอบคุณภาพความเชื่อมั่นสามารถหาได้ 3 วิธี ประกอบด้วย ความเชื่อมั่นแบบการวัดความคงที่ ความเชื่อมั่นแบบการวัดความสมมูลกัน และความเชื่อมั่นโดยการวัดความสอดคล้องภายใน ซึ่งผู้วิจัยได้ตรวจสอบคุณภาพเชื่อมั่นโดยการวัดความสอดคล้องภายใน หากค่าความเชื่อมั่นของผู้สังเกต 2 คน โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's Product-moment Correlation Coefficient) (สมนึก ภัททิยธนี. 2551 : 254) ดังนี้

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

เมื่อ	r_{xy}	แทน	สัมประสิทธิ์สหสัมพันธ์ระหว่างผู้สังเกต 2 คน
	$\sum X$	แทน	ผลรวมของคะแนน จากการสังเกตของผู้สังเกตคนที่ 1
	$\sum Y$	แทน	ผลรวมของคะแนน จากการสังเกตของผู้สังเกตคนที่ 2
	$\sum XY$	แทน	ผลรวมของผลคูณระหว่าง คะแนน จากการสังเกตของผู้สังเกตคนที่ 1 กับคะแนนจากผู้สังเกตคนที่ 2
	$\sum X^2$	แทน	ผลรวมของกำลังสองของ X
	$\sum Y^2$	แทน	ผลรวมของกำลังสองของ Y
	X	แทน	คะแนนจากผู้สังเกตคนที่ 1
	Y	แทน	คะแนนจากผู้สังเกตคนที่ 2
	N	แทน	จำนวนผู้เข้าสอบ

และค่าความเชื่อมั่นของแบบวัดทักษะภาคปฏิบัติ โดยหาค่าสัมประสิทธิ์แอลฟา (Alpha - coefficient) ตามวิธีของ ครอนบาค (Cronbach) จากสูตร (บุญชม ศรีสะอาด. 2553 : 117) ดังนี้

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum S_i^2}{S_t^2} \right)$$

เมื่อ α แทน ค่าสัมประสิทธิ์ความเชื่อมั่นของเครื่องมือวัด

k แทน จำนวนข้อของเครื่องมือวัด

$\sum S_i^2$ แทน ผลรวมของความแปรปรวนของแต่ละข้อ

S_t^2 แทน ความแปรปรวนของคะแนนรวม

4. เกณฑ์ปกติ (Norm)

4.1 ความหมายของเกณฑ์ปกติ

ล้วน สายยศ และอังคณา สายยศ (2539 : 312) ได้กล่าวว่า เกณฑ์ปกติ หมายถึง ข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนจากประชากรที่นิยามไว้เป็นอย่างดีแล้ว และเป็นคะแนนตัวที่จะบอกระดับความสามารถของผู้สอบว่าอยู่ในระดับใดของกลุ่มประชากร

ฉัตรศิริ ปิยะพิมลสิทธิ์ (2548 : 182) ได้ให้ความหมายของเกณฑ์ปกติว่า เป็น ข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนจากประชากรที่นิยามไว้เป็นอย่างดีแล้ว และเป็นคะแนนที่บอกระดับความสามารถของผู้สอบว่าอยู่ในระดับใดของกลุ่มประชากร แต่ในทางปฏิบัติประชากรที่นิยามไว้เป็นกลุ่มตัวอย่างที่ดี เป็นกลุ่มตัวอย่างที่ดีของประชากร แต่ต้องมีจำนวนที่มากพอที่จะเป็นตัวแทนของประชากร ไม่อย่างนั้นเกณฑ์ปกติก็ไม่น่าเชื่อ

สมนึก กัททิษฐี (2553 : 269) ได้ให้ความหมายของเกณฑ์ปกติว่า เป็นข้อเท็จจริงทางสถิติที่บรรยายการแจกแจงของคะแนนจากประชากรที่นิยามไว้เป็นอย่างดี และเป็นคะแนนตัวที่จะบอกระดับความสามารถของผู้เข้าสอบว่าอยู่ระดับใดของกลุ่มประชากร แต่ต้องมีจำนวนมากพอที่จะเป็นตัวแทนของประชากร ไม่อย่างนั้นเกณฑ์ปกติก็จะเชื่อถือไม่ได้

จากที่กล่าวมาดังกล่าว ผู้วิจัยได้สรุปความหมายของเกณฑ์ปกติ ว่าเป็นคะแนนที่บอกระดับทักษะภาคปฏิบัติ กลุ่มการงานอาชีพและเทคโนโลยี (งานเกษตร) ของนักเรียนระดับชั้นประถมศึกษาปีที่ 5 ในสังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษา ร้อยเอ็ด เขต 2 ว่ามี

ทักษะภาคปฏิบัติ กลุ่มการทำงานอาชีพและเทคโนโลยี (งานเกษตร) อยู่ในระดับใดเมื่อเทียบกับกลุ่มประชากร

4.2 หลักเกณฑ์ในการสร้างเกณฑ์ปกติ มี 3 ประการ ดังนี้

4.2.1 ความเป็นตัวแทนที่ดี กลุ่มตัวอย่างที่นำมาใช้สร้างเกณฑ์ปกติเกิดจากการสุ่มที่เป็นตัวแทนที่ดีของประชากร ในการสุ่มตัวอย่างประชากรที่นิยมทำได้หลายวิธี เช่น การสุ่มอย่างง่าย สุ่มแบบเป็นระบบ สุ่มแบบแบ่งชั้น หรือสุ่มแบบแบ่งกลุ่ม ทั้งนี้ต้องเลือกสุ่มตามความเหมาะสมโดยการพิจารณาประชากรเป็นสำคัญ

4.2.2 มีความเที่ยงตรง เมื่อนำคะแนนสอบไปเทียบกับเกณฑ์ปกติที่ทำไว้แล้ว สามารถแปลความหมายได้ตรงกับความเป็นจริง ความสอดคล้องของคะแนนการสอบกับเกณฑ์ปกติตามความเป็นจริงจึงเป็นสิ่งสำคัญมากในการแปลความหมายของคะแนนการสอบแต่ละครั้ง

4.2.3 มีความทันสมัย เกณฑ์ปกติขึ้นอยู่กับความสามารถของประชากรกลุ่มนั้น การพัฒนาที่อยู่ตลอดเวลา เทคโนโลยี สภาพแวดล้อม อาหารการกิน สิ่งเหล่านี้ จะช่วยให้คนแก่ หรือ อ่อน ได้ ดังนั้นเกณฑ์ปกติที่เคยศึกษาไว้นานแล้วหลายปี อาจไม่ตรงกับความเป็นจริง จึงต้องสร้างขึ้นมาใหม่ให้ทันสมัย โดยทั่วไปแล้วเกณฑ์ปกติควรเปลี่ยนทุก ๆ 5 ปี

4.3 ชนิดของเกณฑ์ปกติ

เกณฑ์ปกติแบ่งได้ตามลักษณะของประชากรและตามลักษณะของการใช้สถิติ เปรียบเทียบดังนี้ (สมนึก ภัททิยธนี. 2553 : 271-272)

เกณฑ์ปกติแบ่งชนิดตามลักษณะของประชากร แบ่งได้ดังนี้

1. เกณฑ์ปกติระดับชาติ (National Norms) การสร้างเกณฑ์ปกติระดับชาตินั้นใช้ประชากรที่นิยมไว้มากมายทั่วประเทศ เช่นหาเกณฑ์ปกติวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 4 ก็ต้องสร้างเกณฑ์ปกติจากนักเรียนชั้นประถมศึกษาปีที่ 4 ทั่วประเทศ จำนวนนักเรียนที่จะต้องจึงมีจำนวนมาก

2. เกณฑ์ปกติระดับท้องถิ่น (Local Norms) เป็นการสร้างเกณฑ์ปกติระดับเล็กลงมา เช่นระดับจังหวัด หรือระดับอำเภอ เป็นประโยชน์ในการเปรียบเทียบคะแนนของผู้สอบกับคนทั้งจังหวัดหรือทั้งอำเภอ

3. เกณฑ์ปกติของโรงเรียน (School Norms) โรงเรียนบางแห่งมีขนาดใหญ่ นักเรียนแต่ละชั้นเรียนมีจำนวนมากเมื่อสร้างแบบทดสอบแต่ละวิชาของแต่ละระดับชั้น ได้ดีมีคุณภาพแล้วจะสร้างเกณฑ์ปกติของโรงเรียนก็ได้ กรณีสร้างเกณฑ์ปกติของโรงเรียนเดียวหรือในกลุ่มโรงเรียนเดียวกัน เรียกว่า เกณฑ์ปกติของโรงเรียน ใช้ประเมินเปรียบเทียบกับนักเรียนแต่ละคนกับนักเรียนส่วนรวมของโรงเรียน และใช้ประเมินการพัฒนาของโรงเรียนได้ด้วย โดยพิจารณาจากผลการสอบแต่ละปีว่าเด่นหรือด้อยกว่าปีที่สร้างเกณฑ์ปกติเอาไว้

เกณฑ์ปกติแบ่งชนิดตามลักษณะของการใช้สถิติเปรียบเทียบ ได้ดังนี้

1. เกณฑ์ปกติเปอร์เซ็นต์ไทล์ (Percentile Norms) เกณฑ์ปกติแบบนี้สร้างจากคะแนนดิบที่มาจากประชากร หรือกลุ่มตัวอย่างที่เป็นตัวแทนที่ดี แล้วดำเนินการตามวิธีการสร้างเกณฑ์ปกติทั่วไปเมื่อหาค่าเปอร์เซ็นต์ไทล์เสร็จก็หยุดแค่นั้น เกณฑ์ปกติแบบนี้เป็นคะแนนจัดอันดับเท่านั้น จะนำไปวกลบกันไม่ได้ แต่สามารถเปรียบเทียบและแปลความหมายได้ เช่น เด็กคนหนึ่งสอบได้ 25 คะแนน ได้เทียบกับเกณฑ์ปกติตรงกับตำแหน่งเปอร์เซ็นต์ไทล์ที่ 80 แสดงว่าถ้ามีคนเข้าสอบ 100 คน เด็กคนนี้มีความสามารถเหนือกว่าคนอื่นอยู่ 80 คน (เขาอ่อนกว่าคนอื่นเพียง 20 คน)

2. เกณฑ์ปกติคะแนนที (T-Score Norms) เป็นคะแนนมาตรฐานที่สามารถนำมาวกลบและหาคะแนนเฉลี่ยได้ มีความเหมาะสมในการแปลความหมาย คือ มีค่าตั้งแต่ 0 ถึง 100 มีค่าเฉลี่ย 50 ส่วนเบี่ยงเบนมาตรฐาน 10 เรียกคะแนนชนิดนี้ว่า คะแนน T ปกติ (Normalized T-Score)

3. เกณฑ์ปกติสเตไนน์ (Stanine Norms) คะแนนแบบนี้เป็นคะแนนมาตรฐานชนิดหนึ่ง แต่มีค่าเพียง 9 ตัว (Standard Nine Point) คะแนนเฉลี่ยอยู่ที่ 5 มีส่วนเบี่ยงเบนมาตรฐานประมาณ 2 คะแนน

4. เกณฑ์ปกติตามอายุ (Age Norms) แบบทดสอบมาตรฐานบางอย่างหาเกณฑ์ปกติตามอายุ เพื่อดูพัฒนาการในเรื่องเดียวกันว่า อายุต่างกันจะมีพัฒนาการอย่างไร หรืออายุเท่ากันจะมีพัฒนาการต่างกันหรือไม่

5. เกณฑ์ปกติตามระดับชั้น (Grade Norms) เป็นการหาเกณฑ์ปกติตามระดับชั้นในโรงเรียน แบบทดสอบที่จะทำเกณฑ์ปกติชนิดนี้ได้ต้องเป็นเนื้อหาเดียวกัน วิชาที่นิยมสร้างเกณฑ์ปกติชนิดนี้มักจะเป็นวิชาพื้นฐาน

จากที่กล่าวมาแล้วสรุปได้ว่า เกณฑ์ปกติที่แบ่งตามลักษณะการใช้สถิติมีหลายชนิด คือ เกณฑ์ปกติเปอร์เซ็นต์ไทล์ เกณฑ์ปกติคะแนนที เกณฑ์ปกติสเตไนน์ เกณฑ์ปกติตาม

อายุและเกณฑ์ปกติตามระดับชั้น ซึ่งในการวิจัยครั้งนี้ ผู้วิจัยได้สร้างเกณฑ์ปกติของคะแนน โดยใช้ตำแหน่งเปอร์เซ็นต์ไทล์เทียบกับคะแนนมาตรฐานที่ปกติ และคำนวณหาค่าคะแนน T ปกติ โดยอาศัยสมการพยากรณ์

4.4 วิธีสร้างเกณฑ์ปกติ

1. วิธีเขียนกราฟจากคู่อันดับ โดยการเขียนกราฟจากคู่อันดับ ระหว่างคะแนนสอบกับคะแนนมาตรฐาน T ปกติ

2. โดยอาศัยสมการพยากรณ์ โดยการพิจารณาคะแนนสอบและคะแนนมาตรฐาน T ปกติ แต่ละคู่ มีลักษณะเป็นคู่อันดับที่มีความสัมพันธ์กันสูง เมื่อทดสอบความสัมพันธ์ระหว่างคะแนนสอบและคะแนน T ปกติ ที่เป็นสมการเส้นตรงได้ดังนี้ (สมนึก กัททิยชนี. 2553 : 272)

$$T_c = a + bx \dots\dots\dots (1)$$

เมื่อ
$$b = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2} \dots\dots\dots (2)$$

และ
$$a = \bar{Y} - b\bar{X} \dots\dots\dots (3)$$

เมื่อ	N	แทน	จำนวน X (คะแนนสอบ)
	T _c	แทน	คะแนน T ปกติ ที่คำนวณจากสมการเส้นตรงอยู่ในรูปฟังก์ชันของคะแนนสอบ
	a	แทน	ตำแหน่งที่เส้นตรงตัดแกน Y
	b	แทน	ความชันของเส้นตรง (ค่าสัมประสิทธิ์การทำนาย หรือการพยากรณ์)
	X	แทน	คะแนนสอบ
	\bar{X}	แทน	ค่าเฉลี่ยของคะแนนสอบ
	Y	แทน	คะแนนมาตรฐาน T ปกติ
	\bar{Y}	แทน	ค่าเฉลี่ยของคะแนนมาตรฐาน T ปกติ

จากที่กล่าวมาข้างต้น ผู้วิจัยใช้หลักการสร้างเกณฑ์ปกติโดยการคำนวณตำแหน่งเปอร์เซ็นต์ไทล์ จากคะแนนดิบ เทียบกับคะแนนมาตรฐานที่ปกติ และคำนวณหาค่า คะแนน T ปกติ โดยอาศัยสมการพยากรณ์ โดยใช้สูตรการหาตำแหน่งเปอร์เซ็นต์ไทล์ดังนี้ (สมนึก ภัททิยธนี. 2553 : 266)

$$PR = \left(cf + \frac{1}{2}f \right) \frac{100}{N}$$

เมื่อ	PR	แทน	ตำแหน่งเปอร์เซ็นต์ไทล์
	f	แทน	ความถี่ของคะแนน
	cf	แทน	ความถี่สะสม
	N	แทน	จำนวนกลุ่มตัวอย่าง

4.5 ขั้นตอนการสร้างเกณฑ์ปกติ (Norms) โดยอาศัยสมการพยากรณ์

4.5.1 การแปลงคะแนนดิบให้เป็นคะแนน T ปกติ มีขั้นตอนดังนี้

1) สร้างตารางแจกแจงความถี่ โดยเรียงคะแนนจากมากไปน้อย แล้วนำคะแนนของนักเรียนแต่ละคนมาลงรอยขีด (Tally)

2) หาค่า f และ cf

3) หาค่า $cf + \frac{1}{2}f$ (จะหาค่า $cf + \frac{1}{2}f$ ของชั้น ต้องใช้ค่า cf ที่อยู่ก่อนหน้านั้น แต่ใช้ค่า f ของชั้นนั้น)

4) เอาค่า $cf + \frac{1}{2}f$ ไปคูณด้วย $\frac{100}{N}$ ได้เป็น $\left(cf + \frac{1}{2}f \right) \frac{100}{N}$ ค่าที่ได้

เรียกว่า ตำแหน่งเปอร์เซ็นต์ไทล์ (Percentile Rank = PR) แสดงถึงค่าของพื้นที่ใต้โค้งการแจกแจงซึ่งมีค่าทั้งหมดเป็น 1 หรือ 100 %

5) นำค่า $\left(cf + \frac{1}{2}f \right) \frac{100}{N}$ หรือ PR ที่ได้ ไปเทียบเป็นค่า T จากตาราง

เทียบตำแหน่งเปอร์เซ็นต์ไทล์ ตามตารางที่ 1

ตารางที่ 1 ตารางเทียบตำแหน่งเปอร์เซ็นต์ไทล์เป็นคะแนนมาตรฐาน T ปกติ

T	0	1	2	3	4	5	6	7	8	9
1	.003	.004	.007	.011	.016	.023	.034	.04	.069	.097
2	.13	.19	.26	.35	.47	.62	.82	1.07	1.39	1.79
3	2.28	2.87	3.59	4.46	5.48	6.68	8.08	9.68	11.51	13.57
4	15.87	18.41	21.19	24.20	27.43	30.85	34.46	38.21	42.07	46.02
5	50.0	53.98	57.93	61.79	65.54	69.15	72.57	75.80	78.81	81.59
6	84.13	86.43	88.49	90.32	91.92	93.32	94.52	95.54	96.41	97.13
7	97.72	98.21	98.61	98.93	99.18	99.38	99.53	99.65	99.74	99.81
8	99.865	99.903	99.931	99.952	99.966	99.977	99.984	99.989	99.993	99.995

ที่มา : สมนึก ภัททิยชนี (2553 : 266)

วิธีเทียบเปอร์เซ็นต์ไทล์เป็นคะแนนมาตรฐาน T ปกติ

1. ตารางเทียบตำแหน่งเปอร์เซ็นต์ไทล์ (PR) เป็นคะแนน T ปกติ ข้างต้นนี้ เป็นตารางที่ปรับรูปแบบมาจากตารางการเปลี่ยนค่าเปอร์เซ็นต์ไทล์ เป็นคะแนน T ปกติ ซึ่งมีอยู่ในหนังสือตำราทั่วไป ทั้งนี้เพื่อความสะดวกในการใช้

2. ค่าของคะแนน T ตามแนวตั้ง (แถวซ้ายมือ เลข 1 - 8) แสดงหลักสิบ และตามแนวนอน (แถวบน เลข 0 - 9) แสดงหลักหน่วย

3. ให้นำค่าตำแหน่งเปอร์เซ็นต์ไทล์ที่คำนวณได้ : $\left(cf + \frac{1}{2} f \right) \frac{100}{N}$ มาเทียบกับค่าเปอร์เซ็นต์ไทล์ที่อยู่ในตารางนี้ซึ่งมีค่าทศนิยม 2-3 ตำแหน่ง โดยพิจารณาค่าที่ตรงกัน (หากไม่มีค่าที่ตรงกันให้ใช้ค่าที่ใกล้เคียงที่สุด)

4. ให้อ่านคะแนน T หลักสิบ จากแนวตั้ง (แถวซ้ายมือ) และรวมกับหลักหน่วยจากแนวนอน (แถวบน) เช่น ถ้าตำแหน่งเปอร์เซ็นต์ไทล์ มีค่า 91.92 จะได้คะแนน T = 64 หรือถ้าตำแหน่งเปอร์เซ็นต์ไทล์ มีค่า 13.57 จะได้คะแนน T = 39 เป็นต้น

5. หากตำแหน่งเปอร์เซ็นต์ไทล์ที่คำนวณได้ไม่ตรงกับค่าใด ๆ ในตารางนี้ ให้เลือกเอาค่าในตารางนี้ที่ใกล้เคียงมากที่สุด ไม่ว่าจะใกล้เคียงกับค่าที่น้อยกว่าหรือมากกว่าก็ตาม เช่นตำแหน่งเปอร์เซ็นต์ไทล์ มีค่า 2.0 จะได้คะแนน T = 29 (เพราะ 2.0 ใกล้เคียง 1.79 มากกว่า 2.28)

4.5.2 สร้างสมการพยากรณ์ ดังนี้

$$T_c = a + bx \quad \dots\dots\dots (1)$$

เมื่อ
$$b = \frac{N\sum XY - \sum X\sum Y}{N\sum X^2 - (\sum X)^2} \quad \dots\dots\dots (2)$$

และ
$$a = \bar{Y} - b\bar{X} \quad \dots\dots\dots (3)$$

- เมื่อ N แทน จำนวน X (คะแนนสอบ)
- T_c แทน คะแนน T ปกติ ที่คำนวณจากสมการเส้นตรงอยู่ในรูป ฟังก์ชันของคะแนนสอบ
- a แทน ตำแหน่งที่เส้นตรงตัดแกน Y
- b แทน ความชันของเส้นตรง (ค่าสัมประสิทธิ์การทำนาย หรือ การพยากรณ์)
- X แทน คะแนนสอบ
- \bar{X} แทน ค่าเฉลี่ยของคะแนนสอบ
- Y แทน คะแนนมาตรฐาน T ปกติ
- \bar{Y} แทน ค่าเฉลี่ยของคะแนนมาตรฐาน T ปกติ

4.5.3 คำนวณคะแนน T ปกติ จากคะแนนสอบโดยอาศัยสมการ ดังนี้

$$T_c = a + bx \quad \dots\dots\dots (1)$$

เมื่อ
$$b = \frac{N\sum XY - \sum X\sum Y}{N\sum X^2 - (\sum X)^2} \quad \dots\dots\dots (2)$$

และ
$$a = \bar{Y} - b\bar{X} \quad \dots\dots\dots (3)$$

4.5.4 ขยายคะแนน T ปกติ เพื่อให้ครอบคลุมคะแนน โดยอาศัยสมการ

$$T_c = a + bx$$

4.6 การแปลความหมายของคะแนน T ปกติ

ระดับทักษะภาคปฏิบัติ กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี ของนักเรียน ระดับชั้นประถมศึกษาปีที่ 5 สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาร้อยเอ็ด เขต 2 แปลความหมายของคะแนนที่ได้จากแบบวัดทักษะภาคปฏิบัติ กลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี โดยการกำหนดตำแหน่งเปอร์เซ็นต์ไทล์ จากคะแนนดิบ มาเทียบกับคะแนนมาตรฐานที่ปกติ (Normalized T-score) และกำหนดค่า คะแนน T ปกติ โดยอาศัยสมการพยากรณ์ แล้วพิจารณาตามเกณฑ์ดังนี้ (ชวาล แพรัตกุล. 2528 : 53)

ต่ำกว่า T35	แปลว่า	มีทักษะภาคปฏิบัติต่ำ
ตั้งแต่ T35 ถึง T44	แปลว่า	มีทักษะภาคปฏิบัติค่อนข้างต่ำ
ตั้งแต่ T45 ถึง T54	แปลว่า	มีทักษะภาคปฏิบัติปานกลาง
ตั้งแต่ T55 ถึง T64	แปลว่า	มีทักษะภาคปฏิบัติค่อนข้างสูง
ตั้งแต่ T65 เป็นต้นไป	แปลว่า	มีทักษะภาคปฏิบัติสูงมาก

5. งานวิจัยที่เกี่ยวข้อง

5.1 งานวิจัยในประเทศไทย

ในการสร้างแบบวัดภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานเกษตร) เรื่องการปลูกพืชผักสวนครัว สำหรับชั้นประถมศึกษาปีที่ 5 นี้ มีงานวิจัยที่เกี่ยวข้องได้แก่

พงษ์พิมล คำลอย (2545 : 104-105) ได้พัฒนาแบบวัดทักษะภาคปฏิบัติวิชาวิทยาศาสตร์ (ว101) ชั้นมัธยมศึกษาปีที่ 1 กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดสำนักงานการประถมศึกษาอำเภอน้ำยืน จังหวัดอุบลราชธานี จำนวน 90 คน ซึ่งได้มาจากการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบสังเกตการณ์ ปฏิบัติงานวิชาวิทยาศาสตร์ และแบบประเมินผลการปฏิบัติงานวิชาวิทยาศาสตร์ สถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน คำนีความสอดคล้องของข้อคำถามกับพฤติกรรม ค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย และ t-test การตรวจสอบของแบบสังเกตการณ์ปฏิบัติงานวิชาวิทยาศาสตร์ ทั้ง 9 ฉบับ มีค่าเฉลี่ย 2.33,3.25,3.75,2.32,3.97,1.73, 2.45,2.97,และ 2.58 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่า 0.75,0.86,1.22,0.88,1.28,0.57, 0.70,1.04และ0.67 ตามลำดับ ค่าความแม่นยำ เิงเนื้อหา โดยให้ผู้เชี่ยวชาญเป็นผู้พิจารณา

ความสอดคล้องระหว่างจุดประสงค์กับเนื้อหา มีค่า IOC ตั้งแต่ 0.67 ถึง 1.00 ค่าอำนาจจำแนกรายข้อโดยการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างคะแนนกลุ่มสูงกับกลุ่มต่ำด้วย t-test มีค่าตั้งแต่ 2.07 ถึง 4.58 มีนัยสำคัญทางสถิติที่ระดับ .05 และค่าเชื่อถือได้ของแบบสังเกตจากผู้สังเกต 2 คนมีค่าตั้งแต่ 0.83 ถึง 0.94 แต่ฉบับมีนัยสำคัญทางสถิติที่ระดับ .01 แบบประเมินผลการปฏิบัติงานวิชาวิทยาศาสตร์ ได้ค่าอำนาจจำแนกรายข้อโดยการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างคะแนนกลุ่มสูงกับกลุ่มต่ำด้วย t-test จากการทดสอบครั้งที่ 1 มีค่าตั้งแต่ 1.48 ถึง 14.00 ทุกข้อรายการมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้นข้อที่ 5 ซึ่งต้องปรับปรุงแก้ไขเพื่อใช้ทดสอบครั้งต่อไป การทดสอบครั้งที่ 2 มีค่าตั้งแต่ 1.47 ถึง 14.00 ทุกข้อรายการมีนัยสำคัญทางสถิติที่ระดับ .05 ยกเว้นข้อที่ 5 ซึ่งต้องปรับปรุงแก้ไขและการตรวจสอบครั้งที่ 3 ของแบบสังเกตการปฏิบัติงานวิชาวิทยาศาสตร์ทั้ง มีค่าเฉลี่ย 7.13, 5.67, 4.83, 5.91, 6.05, 5.88, 5.33 และ 5.63 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่า 2.05, 1.55, 1.41, 1.43, 1.89, 2.07, 1.52, 1.63 และ 1.64 ตามลำดับ ค่าความแม่นยำตรงเชิงเนื้อหา โดยให้ผู้เชี่ยวชาญเป็นผู้พิจารณาความสอดคล้องระหว่างจุดประสงค์กับเนื้อหา มีค่า IOC ตั้งแต่ 0.67 ถึง 1.00 ค่าอำนาจจำแนกรายข้อโดยการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างคะแนนกลุ่มสูงกับกลุ่มต่ำด้วย t-test มีค่าตั้งแต่ 2.26 ถึง 14.00 มีนัยสำคัญทางสถิติที่ระดับ .05 และค่าเชื่อถือได้ของแบบสังเกตจากผู้สังเกต 2 คนมีค่าตั้งแต่ 0.94 ถึง 0.99 และมีนัยสำคัญทางสถิติที่ระดับ .01

ชนศักดิ์ เคนบุปผา (2546 : 86-88) ได้พัฒนาแบบวัดภาคปฏิบัติกลุ่มการทำงานพื้นฐานอาชีพ(งานเกษตร) ชั้นประถมศึกษาปีที่ 6 กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ของโรงเรียนบ้านนาชุม โรงเรียนบ้านกองโพนและโรงเรียนบ้านไทรงาม สังกัดสำนักงานเขตพื้นที่การศึกษาอุบลราชธานี เขต 2 จำนวน 60 คนซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน เครื่องมือที่พัฒนาประกอบด้วย แบบวัดภาคปฏิบัติกลุ่มการทำงานพื้นฐานอาชีพ (งานเกษตร) ประกอบด้วยแบบทดสอบจำนวน 4 ฉบับ ได้แก่ ฉบับที่ 1 เรื่องการขยายพันธุ์พืชโดยการตอนกิ่ง จำนวน 14 ข้อ ฉบับที่ 2 เรื่องการขยายพันธุ์พืชโดยการติดตา จำนวน 12 ข้อ ฉบับที่ 3 เรื่องการขยายพันธุ์พืชโดยการต่อกิ่ง จำนวน 13 ข้อ และฉบับที่ 4 เรื่องการขยายพันธุ์พืชโดยการทาบกิ่ง จำนวน 12 ข้อ สถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความคลาดเคลื่อนมาตรฐานในการวัด การวิเคราะห์ความแปรปรวน ค่าสัมประสิทธิ์ของเพียร์สัน ค่าความยาก ค่าอำนาจจำแนก และค่าดัชนีความ

สอดคล้องระหว่างจุดประสงค์เชิงพฤติกรรมกับงานที่กำหนดให้ปฏิบัติและระหว่างเกณฑ์การให้คะแนนกับรายละเอียดการให้คะแนน ดังนี้

ค่าความสอดคล้องระหว่างจุดประสงค์เชิงพฤติกรรมกับงานที่กำหนดให้ปฏิบัติ ตั้งแต่ ฉบับที่ 1 ตั้งแต่ 0.87 ถึง 1.00 ฉบับที่ 2 ตั้งแต่ 0.75 ถึง 1.00 ฉบับที่ 3 ตั้งแต่ 0.62 ถึง 1.00 ฉบับที่ 4 ตั้งแต่ 0.62 ถึง 1.00

ค่าความสอดคล้องระหว่างเกณฑ์ การให้คะแนนกับรายละเอียดการให้คะแนน ฉบับที่ 1 ตั้งแต่ 0.50 ถึง 1.00 ฉบับที่ 2 ตั้งแต่ 0.75 ถึง 1.00 ฉบับที่ 3 ตั้งแต่ 0.62 ถึง 1.00 ฉบับที่ 4 ตั้งแต่ 0.62 ถึง 1.00

ค่าอำนาจจำแนก ฉบับที่ 1 ตั้งแต่ 0.47 ถึง 0.93 ฉบับที่ 2 ตั้งแต่ 0.40 ถึง 0.84 ฉบับที่ 3 ตั้งแต่ 0.33 ถึง 0.80 ฉบับที่ 4 ตั้งแต่ 0.47 ถึง 0.87

ค่าความยาก ฉบับที่ 1 ตั้งแต่ 0.6- ถึง 0.77 ฉบับที่ 2 ตั้งแต่ 0.53 ถึง 0.80 ฉบับที่ 3 ตั้งแต่ 0.60 ถึง 0.83 ฉบับที่ 4 ตั้งแต่ 0.47 ถึง 0.77

ค่าความเชื่อถือได้ของแบบวัด ฉบับที่ 1 มีค่าเท่ากับ .99 ฉบับที่ 2 มีค่าเท่ากับ .97 ฉบับที่ 3 มีค่าเท่ากับ .96 ฉบับที่ 4 มีค่าเท่ากับ .94

ค่าความคลาดเคลื่อนมาตรฐานในการวัด ฉบับที่ 1 มีค่าเท่ากับ 1.57 ฉบับที่ 2 มีค่าเท่ากับ 0.82 ฉบับที่ 3 มีค่าเท่ากับ 0.91 ฉบับที่ 4 มีค่าเท่ากับ 1.36

ความเป็นปรนัยในการตรวจให้คะแนนรายข้อมีค่าสัมประสิทธิ์สหสัมพันธ์ ฉบับที่ 1 ตั้งแต่ 0.90 ถึง 1.00 ฉบับที่ 2 ตั้งแต่ 0.388 ถึง 0.915 ฉบับที่ 3 ตั้งแต่ 0.28 ถึง 1.00 ฉบับที่ 4 ตั้งแต่ 0.53 ถึง 0.96

ความเป็นปรนัยในการตรวจให้คะแนนทั้งฉบับมีค่าสัมประสิทธิ์สหสัมพันธ์ ฉบับที่ 1 ตั้งแต่ 0.96 ถึง 0.97 ฉบับที่ 2 ตั้งแต่ 0.91 ถึง 0.95 ฉบับที่ 3 ตั้งแต่ 0.87 ถึง 0.96 ฉบับที่ 4 ตั้งแต่ 0.77 ถึง 0.94

วีระ พันธุ์สุวรรณ (2546 : 92-93) ได้สร้างแบบวัดภาคปฏิบัติวิชาดนตรี และตรวจสอบคุณภาพของแบบวัดภาคปฏิบัติวิชาดนตรี สำหรับนักเรียนชั้นประถมศึกษาปีที่ 6 กลุ่มตัวอย่างที่ใช้ในการศึกษารั้งนี้คือนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนตระการพิชผล สำนักงานการประถมศึกษา จำนวน 70 คน ซึ่งได้มาโดยการสุ่มอย่างง่าย สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าอำนาจจำแนกรายข้อ คำนีความสอดคล้องของข้อคำถามกับเนื้อหาและค่าความเชื่อถือได้ ผลการวิจัยพบว่า แบบวัดภาคปฏิบัติวิชาดนตรี ทั้ง 7 ชั้น 32 ข้อ มีค่าดัชนีความสอดคล้องข้อคำถามกับเนื้อหา ตั้งแต่ .60 ถึง 1.00

มีค่าอำนาจจำแนกรายข้อตั้งแต่ .22 ถึง .65 และข้อคำถามกับเนื้อหา ตั้งแต่ .6 ถึง 1.00 มีค่าอำนาจจำแนกรายข้อตั้งแต่ .22 ถึง .65 และค่าความเชื่อถือได้ .90

ภัทรพร มะลิซ้อน (2547 : 131-149) ได้สร้างเครื่องมือวัดภาคปฏิบัติการทดลอง เรื่อง สารในชีวิตประจำวัน กลุ่มสาระการเรียนรู้วิทยาศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 จำนวน 2 ชนิด คือ แบบสังเกตวัดภาคปฏิบัติการทดลอง และแบบทดสอบการแก้ปัญหาภาคปฏิบัติ กลุ่มตัวอย่าง เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2546 ของโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาร้อยเอ็ดเขต 2 ได้มาโดยการสุ่มแบบหลายขั้นตอน จำนวน 208 คน ผลการวิจัยปรากฏว่าแบบสังเกตวัดภาคปฏิบัติการทดลอง อำนาจจำแนกรายข้อ มีค่า ตั้งแต่ 0.36 ถึง 0.93 ความเชื่อมั่นของแบบสังเกต มีค่าตั้งแต่ 0.87 ถึง 0.92 ความเชื่อมั่นของผู้สังเกต 3 คน มีค่าตั้งแต่ 0.98 ถึง 0.99 ความเที่ยงตรงเชิงสภาพ มีค่าตั้งแต่ 0.71 ถึง 0.85 ค่าเฉลี่ยของคะแนนจากผู้สังเกต 3 คน มีค่าตั้งแต่ 11.33 ถึง 12.09 ส่วนเบี่ยงเบนมาตรฐาน มีค่าตั้งแต่ 4.17 ถึง 4.94 ความคลาดเคลื่อนมาตรฐานในการวัด มีค่าตั้งแต่ 0.40 ถึง 0.48 แบบทดสอบการแก้ปัญหาภาคปฏิบัติ มีค่าความเชื่อมั่นตั้งแต่ 0.90 ถึง 0.92 ค่าเฉลี่ยของแบบทดสอบ มีค่าตั้งแต่ 6.04 ถึง 6.78 ส่วนเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 3.35 ถึง 3.56 ค่าความคลาดเคลื่อนมาตรฐานในการวัด มีค่าตั้งแต่ 0.46 ถึง 0.49

อาทิตย์ คั่นเจริญ (2548 : 105-112) ได้สร้างแบบวัดทักษะขั้นพื้นฐานในการเล่น วอลเลย์บอล ของนักเรียนระดับช่วงชั้นที่ 3 (ม.1-ม 3) ตามหลักสูตรการศึกษาขั้นพื้นฐาน พุทธศักราช 2544 กลุ่มตัวอย่างแบ่งเป็น 2 กลุ่ม คือกลุ่มที่ใช้ในการร่วมพิจารณาแบบวัดทักษะจำนวน 10 คน จากผู้เชี่ยวชาญและกลุ่มตัวอย่างที่ใช้ในการทดลองแบบวัดทักษะขั้นพื้นฐานกีฬา วอลเลย์บอล ซึ่งเป็นนักเรียนระดับช่วงชั้นที่ 3 (ม.1-ม 3) ภาคเรียนที่ 2 ปีการศึกษา 2548 ของโรงเรียนวัดแสมคำ สำนักงานเขตบางขุนเทียน กรุงเทพมหานคร ได้มาโดยวิธีการสุ่มแบบแบ่งชั้น ระดับชั้นละ 10 คน รวม 30 คน สถิติที่ใช้ในการศึกษาค้นคว้า ได้แก่ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ค่าความคลาดเคลื่อนมาตรฐาน ค่าดัชนีความสอดคล้องระหว่างจุดประสงค์เชิงพฤติกรรมกับข้อปฏิบัติ ค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติกับเกณฑ์การให้คะแนน ค่าดัชนีความสอดคล้องระหว่างผู้ประเมิน ผลการศึกษาค้นคว้า พบว่า แบบวัดทักษะขั้นพื้นฐานในการเล่น วอลเลย์บอล ประกอบด้วยแบบวัดทักษะ 10 ฉบับ ค่าดัชนีความสอดคล้องระหว่างจุดประสงค์เชิงพฤติกรรมกับข้อปฏิบัติตั้งแต่ 0.77 – 1.00 มีค่า ค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติกับเกณฑ์การให้คะแนน ตั้งแต่ 0.77 – 1.00

ผู้ประเมิน (RAI) เท่ากับ 0.88 – 0.91 แสดงว่าแบบวัดทักษะขั้นพื้นฐานในการเล่น วอลเลย์บอลมีคุณภาพตามเกณฑ์

กษมาศ คงตุก (2549 : 103-110) ได้สร้างแบบวัดผลการเรียนรู้ในกลุ่มสาระ การเรียนรู้การงานอาชีพและเทคโนโลยี ของนักเรียนชั้นประถมศึกษาปีที่ 6 กลุ่มเป้าหมายที่ ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 โรงเรียนวัดทัศนารุณสุนทรภิราม เขตราช เทวี สังกัดกรุงเทพมหานคร ภาคเรียนที่ 1 ปีการศึกษา 2548 โดยเลือกแบบเจาะจง จำนวน 2 ห้องเรียน มีนักเรียนจำนวน 86 คน เป็นชาย 41 คน เป็นหญิง 45 คน เครื่องมือที่ใช้ใน การศึกษาได้แก่ แบบวัดผลการเรียนรู้ แบบวัดเจตคติต่อการเรียนวิชาคอมพิวเตอร์ แบบ ประเมินทักษะการปฏิบัติ แบบประเมินตนเองในการทำงาน แบบประเมินโดยเพื่อนในการ ทำงาน แบบบันทึกการพูดคุยกับนักเรียน แบบบันทึกการอภิปรายในชั้นเรียน แบบบันทึกการ ตรวจสอบผลงาน ผลการศึกษาพบว่า แบบวัดผลการเรียนรู้มีค่าความตรงเชิงพินิจตั้งแต่ 0.80 – 1.00 , มีค่าความยาก (p) .33 - .74 ค่าอำนาจจำแนก (r) .29 - .81 จำนวน 37 ข้อ ค่าความ เชื่อมั่น .88 แบบวัดเจตคติต่อวิชาคอมพิวเตอร์ แบบประเมินทักษะการปฏิบัติ แบบประเมิน ตนเองในการทำงาน แบบประเมินโดยเพื่อนในการทำงาน ค่าความเชื่อมั่น .78 .81 .83 และ .85 ตามลำดับ บันทึกการพูดคุยด้านความสะดวกในการใช้ ความคล่องตัวในการใช้และเวลา ในการใช้ พบว่าครูร้อยละ 70 มีความเห็นว่าสะดวก แบบบันทึกการอภิปรายในชั้นเรียน นักเรียนมีความเห็นว่าสะดวกและคล่องตัวดี

สุรัชย์ สิงห์แผ่น (2550: 56-59) ได้สร้างแบบวัดภาคปฏิบัติกลุ่มสาระการงาน อาชีพและเทคโนโลยี วิชาช่างเชื่อมโลหะไฟฟ้า ชั้นมัธยมศึกษาปีที่ 5 กลุ่มตัวอย่างที่ใช้ใน การศึกษาค้นคว้า เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 โรงเรียนพรเจริญ เขตพื้นที่การศึกษา หนองคาย เขต 3 ซึ่งได้มาโดยการเลือกแบบเจาะจง (Purposive Sampling) จำนวน 72 คน ภาคเรียนที่ 2 ปีการศึกษา 2549 เครื่องมือที่ใช้ในการศึกษาค้นคว้า มี 2 ชนิด ได้แก่ แบบ สังเกตปฏิบัติงาน จำนวน 5 ฉบับ ประกอบด้วย ฉบับที่ 1 การเริ่มต้นอาร์ค (Strike the arc) มี 12 ข้อ ฉบับที่ 2 ท่าเชื่อม (Welding Position) ในการเชื่อมโลหะด้วยไฟฟ้า มี 14 ข้อ ฉบับที่ 3 ความรู้เกี่ยวกับการออกแบบมี 7 ข้อ ฉบับที่ 4 การผลิตชิ้นงาน มี 14 ข้อ และฉบับที่ 5 การนำเสนอผลงานมี 16 ข้อ และแบบทดสอบวัดความเข้าใจ ชนิดเลือกตอบ 5 ตัวเลือก 1 ฉบับ 40 ข้อ ทำการทดลองใช้เครื่องมือ 3 ครั้ง ครั้งที่ 1 ใช้กลุ่มตัวอย่าง 30 คน ครั้งที่ 2 ใช้กลุ่ม ตัวอย่าง 20 คน เพื่อหาความเชื่อมั่นของแบบสังเกตการณ์ปฏิบัติงาน ค่าความเชื่อมั่นของผู้ สังเกตให้คะแนน 2 คนและค่าความคลาดเคลื่อนมาตรฐานใน การวัด และครั้งที่ 3 ใช้กลุ่ม

ตัวอย่าง 22 คน เพื่อเป็นการหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าอำนาจจำแนกรายข้อ ค่าความเชื่อมั่นของแบบทดสอบวัดความเข้าใจ ผลการศึกษาค้นคว้า ปรากฏว่าแบบสังเกตการณ์ปฏิบัติงาน จำนวน 5 ฉบับที่พัฒนาขึ้นมีค่าเฉลี่ยตั้งแต่ 7.02 ถึง 7.59 จากคะแนนเต็ม 10 คะแนน และค่าเฉลี่ยของแบบสังเกตการณ์ปฏิบัติงาน ฉบับที่ 2 และ 4 มีค่า 17.02 และ 17.32 ตามลำดับ จากคะแนนเต็ม 20 คะแนน ส่วนเบี่ยงเบนมาตรฐานมีค่าตั้งแต่ 0.86 ถึง 1.38 ค่าความเชื่อมั่นของแบบสังเกตการณ์ปฏิบัติงาน ตั้งแต่ 0.63 ถึง 0.91 ค่าความเชื่อมั่นของผู้สังเกตให้คะแนน 2 คน มีค่าตั้งแต่ 0.48 ถึง 0.85 ค่าความคลาดเคลื่อนมาตรฐานของการวัด มีค่าตั้งแต่ 0.51 ถึง 0.69 แบบทดสอบวัดความเข้าใจ จำนวน 1 ฉบับ ชนิดเลือกตอบ 5 ตัวเลือก จำนวน 40 ข้อ มีค่าอำนาจจำแนกรายข้อ ตั้งแต่ 0.20 ถึง 0.94 ค่าความเชื่อมั่นทั้งฉบับ มีค่า 0.94 ค่าความคลาดเคลื่อนมาตรฐานของการวัดมีค่า 0.09 เมื่อนำไปทดลอง ใช้กับการวัดภาคปฏิบัติครั้งที่ 3 ได้ค่าเฉลี่ย 25.91 จากคะแนนเต็ม 40 คะแนน ส่วนเบี่ยงเบนมาตรฐาน มีค่า 8.52

นิภารัตน์ เฟิงอารีย์ (2550 : 89-97) ได้สร้างเครื่องมือวัดภาคปฏิบัติในกิจกรรมการทดลองวิทยาศาสตร์ชั้นประถมศึกษาปีที่ 6 เพื่อหาคุณภาพเครื่องมือวัดภาคปฏิบัติที่ผู้วิจัยสร้างขึ้น ซึ่งได้มาโดยการสุ่มแบบหลายขั้นตอน จำนวน 221 คน สถิติที่ใช้ในการวิเคราะห์ข้อมูลได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าอำนาจจำแนก ค่าความเชื่อมั่น และค่าความคลาดเคลื่อนมาตรฐานของการวัดผลการวิจัยมีดังนี้ แบบสังเกตวัดภาคปฏิบัติการทดลอง จำนวน 4 ฉบับ อำนาจจำแนกรายข้อแบบสังเกตภาคปฏิบัติมีค่าตั้งแต่ 0.50 ถึง 0.68 ค่าความเชื่อมั่นของแบบสังเกตภาคปฏิบัติ มีค่าตั้งแต่ 0.73 ถึง 0.83 ความเชื่อมั่นของผู้สังเกต 2 คน มีค่าตั้งแต่ 0.79 ถึง 0.95 และค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าตั้งแต่ 0.64 ถึง 0.80 แบบทดสอบวัดความรู้ความเข้าใจกลุ่มสาระการเรียนรู้วิทยาศาสตร์ จำนวน 4 ฉบับ ฉบับละ 10 ข้อ คะแนนเต็มฉบับละ 10 ซึ่งเป็นแบบเลือกตอบ ชนิด 4 ตัวเลือก ค่าเฉลี่ยมีค่าเท่ากับ 21.42 ส่วนเบี่ยงเบนมาตรฐาน มีค่าเท่ากับ 6.92 ค่าอำนาจจำแนกรายข้อมีค่าตั้งแต่ 0.22 ถึง 0.76 ค่าความเชื่อมั่นของแบบทดสอบวัดความรู้ความเข้าใจวิทยาศาสตร์ มีค่าเท่ากับ 0.82 และค่าความคลาดเคลื่อนมาตรฐานในการวัดมีค่าเท่ากับ 2.93

อำนวยการ ฟ้างสว่าง (2550 : 122-127) การเรียนรู้ได้สร้างและหาคุณภาพแบบวัดภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานบ้าน) ชั้นประถมศึกษาปีที่ 4 กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นประถมศึกษาปีที่ 4 สังกัดสำนักงานเขตพื้นที่การศึกษาอุบลราชธานี เขต 3 จำนวน 95 คน ซึ่งได้มาโดยวิธีการสุ่มแบบหลายขั้นตอน

สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าดัชนีความสอดคล้องระหว่างจุดประสงค์เชิงพฤติกรรมย่อยที่จะวัดพฤติกรรมหลักของแบบวัดภาคปฏิบัติ ค่าดัชนีความสอดคล้องระหว่างเกณฑ์ที่กำหนดกับรายละเอียดการให้คะแนนของแบบวัดภาคปฏิบัติ ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน ค่าสัมประสิทธิ์ความเชื่อถือ ค่าความยาก ค่าอำนาจจำแนกและค่าความคลาดเคลื่อนในการวัด ผลการวิจัยพบว่าเครื่องมือที่พัฒนาประกอบด้วย แบบวัดภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานบ้าน) จำนวน 4 ฉบับ ได้แก่เรื่องที่ 1 การทำความสะอาดบ้าน จำนวน 14 ข้อ ฉบับที่ 2 เรื่องการซักผ้า จำนวน 14 ข้อ ฉบับที่ 3 เรื่องการซ่อมแซมเสื้อผ้า จำนวน 13 ข้อ และฉบับที่ 4 เรื่องการประกอบอาหาร จำนวน 14 ข้อ แบบวัดทั้ง 4 ฉบับที่สร้างขึ้นมีค่าความยากตั้งแต่ 0.6 – 0.80 ค่าอำนาจจำแนกตั้งแต่ 0.20 - 0.67 ค่าสัมประสิทธิ์สหสัมพันธ์ของผู้สังเกตให้คะแนน 3 คน มีค่าตั้งแต่ 0.93 – 0.97, 0.95 - 0.97 และ 0.94–0.98 ตามลำดับ ค่าสัมประสิทธิ์ความเชื่อถือได้มีค่าตั้งแต่ 0.98 – 0.99 ค่าเฉลี่ยของคะแนนตั้งแต่ 15.75, 16.53, 15.22 และ 15.71 ค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐานตั้งแต่ 0.04 – 0.07 และเกณฑ์การให้คะแนนแต่ละขั้นคือ ขั้นเตรียม 15% ขั้นปฏิบัติ 35% ขั้นผลงาน 30% และขั้นกิจนิสัย 20% สรุปผลแบบวัดภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานบ้าน) มีค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยงตรง ค่าความเชื่อถือได้อยู่ในเกณฑ์ที่มีคุณภาพและสอดคล้องกับสภาพจริง จึงสามารถนำไปใช้เป็นเครื่องมือในการวัดภาคปฏิบัติงานบ้าน ชั้นประถมศึกษาปีที่ 4 ได้

5.2 งานวิจัยต่างประเทศ

บัวแมน (Booman. 1991 : 194) ได้สร้างแบบทดสอบวัดทักษะการปฏิบัติงานวิชาฟิสิกส์ในระดับมัธยมศึกษาตอนปลายโดยศึกษาจากกลุ่มตัวอย่าง 219 คนซึ่งเป็นนักเรียนที่เรียนวิชาฟิสิกส์จากโรงเรียนในชนบท ในเมือง ทั้งโรงเรียนรัฐบาลและเอกชนในนิวยอร์ก ซึ่งทดสอบในเดือนพฤษภาคม ค.ศ. 1990 โดยใช้แบบทดสอบจำนวน 7 ฉบับ ซึ่งได้ตรวจสอบความเที่ยงตรงตามเนื้อหา โดยให้ผู้เชี่ยวชาญเป็นครูผู้สอนวิชาฟิสิกส์ในระดับมัธยมศึกษาตอนปลายเป็นผู้ตรวจสอบ ปรากฏว่าแบบทดสอบมีความเที่ยงตรงตามเนื้อหาและหาความเที่ยงตรงตามโครงสร้าง โดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน มีค่าเท่ากับ 0.49 ส่วนความเชื่อมั่นของแบบทดสอบหาโดยใช้สูตรวิธีหาสัมประสิทธิ์แอลฟาของ Cronbach (Cronbach Alpha Coefficient) มีค่าเท่ากับ 0.81

ไนท์ (Knight, 1992 : 4302) ได้สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนระดับ 5 ในมิชิแกน กลุ่มตัวอย่างที่ใช้ในครั้งนี้เป็นนักเรียนในระดับ 5 จำนวน 37 คน ข้อสอบที่สร้างขึ้นมีจำนวน 26 ข้อ โดยให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงตามเนื้อหา ปรากฏว่าทุกข้อมีคะแนนเฉลี่ยเกิน 0.75 หากค่าอำนาจจำแนกรายข้อโดยใช้สูตรสหสัมพันธ์แบบพอยท์ - ไบซีเรียล (Point - Biserian Correlatoin) มีค่าอยู่ระหว่าง 0.19 ถึง 0.74 สำหรับค่าความเชื่อมั่นของผู้สังเกต 2 คน หาโดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน มีค่าเท่ากับ 0.78 ส่วนความเชื่อมั่นของแบบทดสอบหาโดยใช้สูตร KR-20 มีค่าเท่ากับ 0.70

อาดัมส์ (Adams, 2000 : 2321) ได้ทำการศึกษาเพื่อสร้างและทดสอบผลกระทบของรูปแบบการประเมินผลใหม่ที่มีต่อวิชาชีพ โดยการสังเคราะห์องค์ประกอบที่ดีที่สุดของรูปแบบที่มีอยู่ 5 รูปแบบ ได้แก่ (Kirkpatrick 1986 1960 1994 ; Cevro 1984 1986 Grotelueschen 1986; Jackson ; และ Robinson 1989) ในการพัฒนาแนวทางและเครื่องมือชุดหนึ่งเพื่อสนับสนุนรูปแบบ และการทดสอบด้วยโปรแกรมการศึกษาทางวิชาชีพต่อเนื่อง สำหรับ นายธนาคาร รูปแบบการประเมินผลกระทบที่อาศัยพฤติกรรมนั้น อาศัยระดับการประเมิน 4 ระดับ ของ Kirkpatrick ซึ่งใช้ในระดับตรงกันข้าม ได้แก่ผลที่เกิดขึ้น พฤติกรรมการเรียนรู้ และปฏิกิริยา ซึ่งมีคำถามที่ถามมาก่อนในระหว่างและหลังโปรแกรมการศึกษา วิชาชีพต่อเนื่องเพื่อหาสารสนเทศการประเมินจาก ผู้ร่วมวิจัยและผู้จัดการของตนให้มีแนวทางการกำหนดสูตรที่ละชั้นๆ การพัฒนาและการทดสอบรูปแบบการประเมินผลกระทบที่อาศัยพฤติกรรมผลการศึกษพบว่า รูปแบบการประเมินดังกล่าวนี้ได้รับคำอธิบายว่าเป็นวิธีการประเมินที่มีผลกระทบที่เป็นเชิงเส้นตรง เพื่อให้ตรงกับการทำนายที่นำเสนอเป็นการเปลี่ยนแปลงองค์การจากยุคอุตสาหกรรมที่เจริญก้าวหน้าเป็นเชิงเส้นตรง ไปจนถึงยุคสารสนเทศไปยุคความรู้ที่ไม่เพิ่มขึ้นเป็นเชิงเส้นตรง

ซีก (Cheak, 2001 : 3135) ได้ทำการศึกษาและพัฒนาแบบทดสอบภาคสนามเพื่อวัดความคิดเชิงวิจารณ์ของนักเรียนในขณะที่กำลังเรียนรู้หนังสือ ในสาขาสิ่งแวดล้อมศึกษา โดยใช้แบบทดสอบการคิดเชิงวิจารณ์ในสิ่งแวดล้อมศึกษาทดสอบภาคสนามกับกลุ่มตัวอย่างจากนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 และ 5 และนักศึกษาระดับวิทยาลัย การจัดลำดับนักเรียนเหล่านี้ จัดตามการรับรู้ของครูอาจารย์ด้านการให้เหตุผลเกี่ยวกับระดับความสามารถในการคิดเชิงวิจารณ์ ของนักเรียนนักศึกษาเป็นรายบุคคล แบบทดสอบการคิดเชิงวิจารณ์ในสิ่งแวดล้อมศึกษาฉบับนี้มี 3 ตอน วัดความสามารถการคิดเชิงวิจารณ์ 3 ด้าน คือ การลงข้อสรุป

การลงข้อวินิจฉัยและการระบุความสำเร็จหรือคติ ผลการวิเคราะห์ข้อมูลด้วยสถิติหลายชนิด พบว่าตัวแปรอิสระทั้ง 2 ตัวแปร คือระดับความสามารถการคิดเชิงวิจารณ์มีคะแนนสูงกว่า นักศึกษาที่ไม่มีความคิดเชิงวิจารณ์และนักศึกษาที่อายุมากกว่าและให้เหตุผลได้มากกว่าทำ คะแนนได้สูงกว่านักศึกษาที่อายุน้อยกว่าและให้เหตุผลได้น้อยกว่า

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการสร้างเครื่องมือวัดทักษะภาคปฏิบัติ พบว่า มีงานวิจัยที่เกี่ยวข้องกับการสร้างเครื่องมือวัดภาคปฏิบัติค่อนข้างน้อย โดยเฉพาะการสร้าง แบบวัดทักษะภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี ที่เกี่ยวกับวิชางาน เกษตร จากงานวิจัยดังกล่าวข้างต้นพอสรุปได้ว่า การสร้างเครื่องมือวัดทักษะภาคปฏิบัติ นั้น ต้องการคุณภาพของเครื่องมือได้แก่ ค่าความเที่ยงตรง ค่าอำนาจจำแนก ค่าความเชื่อมั่นของ แบบวัด ค่าความเชื่อมั่นของผู้ประเมินผลการปฏิบัติงาน ดังนั้นการสร้างแบบวัดทักษะ ภาคปฏิบัติกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี (งานเกษตร) สำหรับนักเรียนชั้น ประถมศึกษาปีที่ 5 ครั้งนี้ ผู้วิจัยตรวจสอบคุณภาพของเครื่องมือ คือความเที่ยงตรงเชิงเนื้อหา โดยให้ผู้เชี่ยวชาญเป็นผู้ตรวจสอบ โดยใช้สูตร IOC หาค่าเฉลี่ยดัชนีความสอดคล้องของผู้เชี่ยวชาญทั้งหมดตามวิธีของ โรวินELLI และแฮมเบิลตัน (Rovinelli and Hambleton) ใช้ เทคนิคกลุ่มรู้จัก (Known-group Technique) ค่าอำนาจจำแนกของแบบวัดทักษะภาคปฏิบัติ ที่มีลักษณะค่าของคะแนนในแต่ละข้อมีหลายค่า ตามลักษณะการตรวจให้คะแนนแบบข้อสอบ อัตนัย โดยแบ่งผู้เข้าสอบออกเป็นสองคือกลุ่มสูงและกลุ่มต่ำด้วยเทคนิค 25% และหาค่า อำนาจจำแนก โดยใช้สูตร ของวิทนีและซาเบอร์ส (Whitney & Sabers) ในการทดสอบความ แตกต่างระหว่างค่าเฉลี่ยของทั้งสองกลุ่ม หาค่าความเชื่อมั่นของผู้สังเกตผลการปฏิบัติงาน โดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน และหาค่าความเชื่อมั่นของแบบวัดทักษะภาคปฏิบัติ โดย หาค่าสัมประสิทธิ์แอลฟา (Alpha - coefficient) ตามวิธีของ Cronbach