

สวัสดี สุคนธรังษี

การวัด
ในการ
จัดงาน
บุคคล

สวัสดิ์ สุกนธรัมย์ จบการศึกษาระดับมัธยมที่โรงเรียนวัดบวรนิเวศ เมื่อ พ.ศ. ๒๔๙๒ เคยผ่านการศึกษาระดับอุดมศึกษา ในคณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และ Institute of Public Administration, University of the Philippines ได้รับปริญญาดุษฎีบัณฑิต (Doctor of Philosophy) ทางสาขาวิชาการบริหารการศึกษา และรัฐประศาสนศาสตร์ จาก Indiana University ในปีการศึกษา พ.ศ. ๒๕๐๓

ประวัติการทำงาน เริ่มเข้ารับราชการในตำแหน่งอาจารย์ตรีวิทยาลัยวิชาการศึกษา เมื่อ พ.ศ. ๒๕๐๑ ตำแหน่งงานบางตำแหน่งที่เคยปฏิบัติในอดีต คือ หัวหน้าแผนกโรงเรียนฝึกหัดครู กรมการฝึกหัดครู หัวหน้ากองวิเทศการ สำนักงานปลัดกระทรวงพัฒนาการแห่งชาติ ผู้อำนวยการสถาบันภาษา กรมวิเทศสหการ และผู้อำนวยการสำนักวิจัย สถาบันบัณฑิตพัฒนบริหารศาสตร์

ปัจจุบัน ดำรงตำแหน่งอาจารย์ชั้นพิเศษ สถาบันบัณฑิตพัฒนบริหารศาสตร์ และเป็นผู้บรรยายพิเศษ วิทยาลัยป้องกันราชอาณาจักร

นอกจากหนังสือเล่มนี้แล้ว ผู้เขียนยังได้เรียบเรียงตำราอื่น ๆ อีกบางเล่ม คือ วิธีการวิจัยเบื้องต้น การบริหารด้วยการวิจัยปฏิบัติการ การบริหารเชิงทฤษฎีคณิตศาสตร์ และการสรรหาและคัดเลือกผู้บริหารของหน่วยงานธุรกิจ



มูลนิธิโครงการตำราสังคมศาสตร์และมนุษยศาสตร์
๔๓๓/๓๘ ถนนธรรมวโรดม แขวงบางกอกน้อย อพท. ๑๐๗๐๐
โทร. ๔๖๔-๕๖๖๘

มูลนิธิโครงการตำรา
สังคมศาสตร์และมนุษยศาสตร์

ราคา 40.00 บาท

ปก นันทะ เจริญพันธ์

การวัดในการจัดงานบุคคล

สวัสดิ์ สุขนรงค์ษ์, Ph.D.

สถาบันบัณฑิตพัฒนบริหารศาสตร์



โครงการตำรา
สังคมศาสตร์และมนุษยศาสตร์
สมาคมสังคมศาสตร์แห่งประเทศไทย
กรุงเทพมหานคร 2517

พิมพ์ครั้งที่หนึ่ง พ.ศ. 2517

ลิขสิทธิ์ภาษาไทยเป็นของโครงการตำราสังคมศาสตร์และมนุษยศาสตร์

สมาคมสังคมศาสตร์แห่งประเทศไทย

บริษัทสำนักพิมพ์ ไทยวัฒนาพานิช จำกัด

599 ถนนไมตรีจิต กรุงเทพมหานคร

เป็นผู้แทนจำหน่าย

รายนามคณะกรรมการบริหาร
โครงการตำราสังคมศาสตร์และมนุษยศาสตร์
สมาคมสังคมศาสตร์แห่งประเทศไทย

นายป๋วย อึ๊งภากรณ์	ประธาน
นายเสน่ห์ จามริก	รองประธาน
นางสาวกุสุมา สนิทวงศ์ ณ อยุธยา	กรรมการ
นางกุหลาบ มัลลิกะมาส	กรรมการ
นายโกศล ศรีสังข์	กรรมการ
นายจิรายุ อิศรางกูร ณ อยุธยา	กรรมการ
นายเจตนา นาควัชระ	กรรมการ
นายชาญวิทย์ เกษตรศิริ	กรรมการ
นายชิตยา สุวรรณชฎ	กรรมการ
นายวิทย์ วิศทเวทย์	กรรมการ
นายวีรยุทธ วิเชียรโชติ	กรรมการ
นายสมบัติ จันทรวงศ์	กรรมการ
นางสาวสวาท เสนาณรงค์	กรรมการ
นายสังเวียน อินทรวิชัย	กรรมการ
นายสายหยุด จำปาทอง	กรรมการ
นายเกริกเกียรติ พิพัฒน์เสรีธรรม	กรรมการและเหรัญญิก
นายสุลักษณ์ ศิวรักษ์	กรรมการและเลขานุการ

อนุกรรมการสาขาวิชาจิตวิทยา

นายวีรยุทธ วิเชียรโชติ	วิทยาลัยวิชาการศึกษา ประสานมิตร
นางนวลเพ็ญ วิเชียรโชติ	สถาบันระหว่างชาติสำหรับคณาจารย์
นางประนอม สโรชมาน	คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์
นายผวน สมปรีดา	โรงพยาบาลสมเด็จพระเจ้าพระยา
นางวนิดา จิตต์หมั่น	คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์
นายสมควร อภัยพันธ์	วิทยาลัยวิชาการศึกษา ประสานมิตร
นายสุข เดชชัย	ภาควิชาจิตวิทยา มหาวิทยาลัยเชียงใหม่
นางสาวสุภาพรรณ พรหมพิทักษ์	คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ขออุทิศหนังสือเล่มนี้

แต่

ครูทุกท่านผู้มีพระคุณ
ที่ได้อบรมสั่งสอนผู้เขียนมา

คำแถลงของโครงการตำรา ฯ

โครงการตำราฯ นี้ ก่อตั้งขึ้นด้วยความร่วมมือกันเองเป็นส่วนบุคคล โดยหมู่ผู้มีอาชีพสอนและผู้รักษานักศึกษาจากสถาบันต่าง ๆ จุติมุ่งหมายเบื้องต้นก็เพื่อส่งเสริมให้มีหนังสือตำราดี ๆ มากขึ้นในภาษาไทย โดยเฉพาะในทางวิชาสังคมศาสตร์และมนุษยศาสตร์ ทั้งนี้เพราะต่างก็เห็นพ้องต้องกันว่า หนังสือตำราภาษาไทยในระดับคุณภาพยังมีไม่เพียงพอ ถ้าส่งเสริมให้มีหนังสือเช่นนี้เพิ่มขึ้น ย่อมเป็นการปรับปรุงมาตรฐานการศึกษาในชั้นมหาวิทยาลัยไปด้วยดีโดยปริยาย ทั้งการส่งเสริมค่านี้ย่อมจะมีคุณค่าทางสร้างสรรค์ปัญญาความคิดริเริ่ม ในเรื่องราวเกี่ยวกับสังคม วัฒนธรรม เศรษฐกิจ และการเมืองอีกด้วย

พร้อมกันนี้ โครงการตำราสังคมศาสตร์และมนุษยศาสตร์ ก็มีเจตนาที่มุ่งจะทำหน้าที่เป็นแหล่งชุมนุมงานเขียนของนักวิชาการจากสถาบันต่าง ๆ ด้วย ทั้งนี้เพื่อให้ผลงานทางวิชาการได้เป็นที่รู้จักและเผยแพร่ออกไปโดยทั่วถึง ทั้งในค่านักสอนและผู้เรียน การดำเนินงานของโครงการตำราฯ มุ่งขยายความเข้าใจและความร่วมมือของบรรดานักวิชาการออกไปในวงกว้างยิ่งขึ้น ทั้งในค่านการกำหนดนโยบายสร้างตำรา การเขียน และการใช้ตำราเหล่านั้น เป็นที่หวังไว้ว่ากิจกรรมร่วมกันค่านนี้ อันเป็นภาระหน้าที่โดยตรงของนักวิชาการ จะเป็นเครื่องส่งเสริมและกระชับความสัมพันธ์อันพึงปรารถนาภายในวงวิชาชีพที่เกี่ยวข้องสืบไป วัตถุประสงค์และหลักการดังกล่าวนี้ เป็นหลักยึดถือในการก่อตั้ง การวางแผนและระเบียบดำเนินงาน ตลอดจนแนวทางแก้ไขปรับปรุงต่อไปในอนาคต

โครงการตำราฯ นี้มีฐานะเป็นโครงการหนึ่งของสมาคมสังคมศาสตร์แห่งประเทศไทย ตั้งขึ้นเพื่อดำเนินงานจัดพิมพ์หนังสือตำราในนามของสมาคมฯ ทั้งนี้โดยมีคณะกรรมการบริหารทำหน้าที่เป็นเอกเทศ ในชั้นนี้ โครงการฯ ได้จัดรูปงานแบ่งออกเป็นสาขาวิชา รวม ๘ สาขา คือ ๑. สาขาวิชาภูมิศาสตร์ ๒. สาขาวิชาประวัติศาสตร์ ๓. สาขาวิชาเศรษฐศาสตร์ ๔. สาขาวิชารัฐศาสตร์ ๕. สาขาวิชาสังคมวิทยาและมานุษยวิทยา ๖. สาขาวิชาปรัชญา ๗. สาขาวิชาจิตวิทยา ๘. สาขาวิชาวรรณคดี แต่ละสาขาวิชามีอนุกรรมการทำหน้าที่พิจารณาวางแผนนโยบายการสร้างตำราในสาขาของตน กำหนดชื่อและเรื่องหนังสือ ตลอดจนจัดหาและกำหนดตัวบุคคลผู้เขียน เพื่อนำเสนอต่อคณะกรรมการบริหาร อนุกรรมการสาขาวิชานี้ ประกอบด้วยนักวิชาการในสาขาที่เกี่ยวข้องจากมหาวิทยาลัยและสถาบันต่าง ๆ นอกเหนือไปจากดำเนินงานจัดทำตำราแล้ว อนุกรรมการสาขาวิชาเหล่านี้ ยังทำหน้าที่ส่งเสริมให้นักวิชาการได้เข้ามามีส่วนร่วมในโครงการให้มากขึ้นเท่าที่จะพึงกระทำได้

โครงการตำราฯ มีนโยบายส่งเสริม การแปล เรียบเรียง และวิจัย จุดประสงค์ สำคัญก็เพื่อเร่งรัดให้ได้มีหนังสือตำราจากงานทุกประเภท อย่างไรก็ตามก็ ในขั้นแรกย่อมจะ ต้องเน้นหนักไปในด้านแปลเป็นส่วนใหญ่ เพื่อเผยแพร่หลักวิชาความรู้ ทั้งในขั้นพื้นฐาน และขั้นสูงใหม่ ๆ ให้ถึงมือนักศึกษาและผู้สนใจทั่วๆ ไป เป็นที่เชื่อแน่ว่าการส่งเสริมงาน แปลจะเป็นทางหนึ่ง ซึ่งจะช่วยปูพื้นฐานสำหรับการริเริ่ม คิดค้น ขีดเขียน ทางด้าน วิชาสังคมศาสตร์และมนุษยศาสตร์ ในวงกว้างต่อไป

ทางค้ำทุนทรัพย์สำหรับจ่ายในขั้นดำเนินงานนั้น โครงการฯ นี้ได้รับความช่วยเหลือในระยะต้นจากมูลนิธิร็อกกี้เฟลเลอร์ ในระยะต่อไป โครงการตำราฯ มุ่งอาศัยกำลัง ทุนจากผลประโยชน์อันพึงได้ จากการจำหน่ายหนังสือที่โครงการฯ นี้ตีพิมพ์ มาใช้เป็นทุน หมุนเวียนต่อไป แต่ในขณะเดียวกัน โครงการตำราสังคมศาสตร์และมนุษยศาสตร์ก็มีใช้ เป็นกิจการแสวงผลกำไร หากมีความมุ่งประสงค์ให้นักศึกษาได้มีโอกาสซื้อหาหนังสือตำรา ได้ในราคาถูกย่อมเยาพอสมควร เพราะฉะนั้น รายได้จักหาพอเพียงกับรายจ่ายไม่ จึงต้อง หวังพึ่งแหล่งการสนับสนุนทางค้ำทุนทรัพย์ต่อไปอีก สิ่งทีคณะกรรมการโครงการตำราฯ หวังก็คือในขั้นต่อไป แหล่งดังกล่าวนี้จักมาจากภายในประเทศของเราเอง หากนักวิชา การได้รับความสนับสนุนให้ได้มีผลงานทางวิชาการปรากฏออกมาเช่นนี้ ย่อมจะเป็นแบบ อย่างอันดีงามสำหรับอนาคตของการศึกษาของประเทศชาติสืบไป



ประธานกรรมการ

โครงการตำราสังคมศาสตร์และมนุษยศาสตร์

คำนำ

การเรียบเรียงตำราเล่มนี้ ได้ใช้เวลาประมาณ 2 ปี ได้เริ่มลงมือเขียนตั้งแต่เดือน มิถุนายน พุทธศักราช 2512 เป็นต้นมา และเพิ่งเสร็จเป็นรูปเล่มในเดือนเมษายน พุทธศักราช 2514

การที่ได้พยายามเรียบเรียงตำรานี้ขึ้นก็เพราะพิจารณาแล้วเห็นว่า เป็นวิชาการที่ควรจะมีเผยแพร่ในประเทศไทย ในปัจจุบัน เป็นที่ประจักษ์แก่ผู้มีหน้าที่รับผิดชอบงานด้านต่าง ๆ ในการบริหารปกครองประเทศแล้วว่า คำขวัญที่ดิเรกมีปากของนักบริหารทั้งหลายว่า ต้องจัดคนให้เหมาะสมกับงาน เป็นถ้อยคำที่ฟังแล้วไพเราะ แต่ปฏิบัติไม่ได้สำเร็จด้วยเพียงแค่นี้กันอย่างง่าย ๆ ทั้งนี้ก็เพราะ การที่จะปฏิบัติเรื่องนี้ให้สัมฤทธิ์ผล สมเจตนาธรรมณ์ ต้องอาศัยหลักการและทฤษฎีอันมีรายละเอียดสลับซับซ้อนหลายอย่าง ดังอธิบายไว้ในหนังสือเล่มนี้แล้ว

ในการเรียบเรียงตำราเล่มนี้ ได้พยายามวิเคราะห์หลักการและทฤษฎีต่าง ๆ ทั้งหมดในเรื่องการวัดและประเมินบุคคล เท่าที่จะมีประโยชน์เกี่ยวข้องกับการจัดงานบุคคล การที่พยายามทำให้ครบถ้วนในเรื่องนี้ ก็เพื่อประโยชน์แก่ผู้ที่จะค้นคว้าวิจัยในโอกาสต่อไป ในแง่ของนักบริหาร อาจไม่จำเป็นที่จะต้องทราบทุกอย่างที่กล่าวไว้ในหนังสือเล่มนี้โดยละเอียด แต่อย่างน้อย ก็ควรจะสามารสร้างแนวคิดของตนเองให้เป็นเหตุเป็นผลสอดคล้องกับหลักการ และทฤษฎีที่ได้มีการพิสูจน์ไว้แล้ว ทั้งนี้จะทำให้ได้ประโยชน์ทั้งในแง่ความยุติธรรมที่ผู้อยู่ใต้บังคับบัญชาพึงได้รับ และในด้านความสำเร็จ ลุล่วง ได้ผลดีของการปฏิบัติงานร่วมกันระหว่างบุคคลทุกฝ่ายในหน่วยงานต่าง ๆ ทั่วประเทศ

เมื่อเริ่มเขียนหนังสือเล่มนี้ ตกลงว่าจะมีผู้ร่วมเขียน 2 ท่าน คือผู้ช่วยศาสตราจารย์
สมศรี จุลละนันท์, M.A., M.S. และผู้ช่วยศาสตราจารย์เพ็ญศรี วายวานนท์, M.A.
แต่เหตุการณ์และเวลาที่ผ่านไป ทำให้ผู้ที่ร่วมเขียนรายแรกต้องย้ายสถานที่ทำงานไป
และผู้ที่ร่วมเขียนรายที่สองก็ได้ร่วมงานแต่เพียงในระยะแรก เนื้อหาในบทที่ 2 เป็น
ผลของการที่ท่านทั้งสองนี้ได้สละเวลามาร่วมเขียน แต่อย่างไรก็ดี หากมีข้อบกพร่อง
ในบทที่ 2 นี้ ผู้เขียนก็เป็นผู้รับผิดชอบแต่ผู้เดียว เช่นเดียวกับที่อาจจะมีในบทอื่นๆ

สวัสดิ์ สุคนธรังษี

สถาบันบัณฑิตพัฒนบริหารศาสตร์

16 ตุลาคม 2516

สารบัญเรื่อง

	หน้า
บทที่ 1 ลักษณะเบื้องต้นของการวัดผล	1
การวัดผลกับการประเมินบุคคล	1
การวัดผลกับแนวคิดทางคณิตศาสตร์	3
มิติและระดับการวัดผล	5
ประวัติการวัดผลบุคคลโดยย่อ	9
ความแตกต่างระหว่างบุคคล และลักษณะทางจิตวิทยาที่วัดได้	10
วิธีการต่างๆ ในการวัดผลบุคคล	15
เกณฑ์การวัด	24
บทที่ 2 การวิเคราะห์งานและการประเมินผลงาน	30
ความหมายของการวิเคราะห์งาน	30
ความมุ่งหมายของการวิเคราะห์งาน	31
ขั้นตอนเบื้องต้นในการวิเคราะห์งาน	32
วิธีการรวบรวมข้อมูล	39
ลักษณะของข้อมูลที่เป็นประโยชน์ต่อเครื่องมือวัดผล	45
การวิเคราะห์งานกับคำอธิบายงาน และข้อกำหนดของคนทำงาน	48
การประเมินผลการปฏิบัติงาน	54
บทที่ 3 ตัวแปรของการวัดผล	60
คุณสมบัติของเครื่องมือวัด	60
ความสัมพันธ์ระหว่างค่าความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์	64

	หน้า
มิติของการวัด	67
ตารางสัมพัทธ์แบบเอฟและแบบอาร์	70
ทฤษฎี Spearman และ Thurstone ในการแยกมิติ	75
บทที่ 4 ค่าความเชื่อมั่นของเครื่องมือวัด	80
ค่าผิดพลาดมาตรฐานของการวัด	81
สัมประสิทธิ์ของความเชื่อมั่น	86
ความแปรปรวนที่มีผลกระทบต่อค่าความเชื่อมั่น	94
การแปลความหมายของสัมประสิทธิ์ความเชื่อมั่น	96
ระดับความเชื่อมั่นที่ควรเชื่อถือได้	103
บทที่ 5 ค่าความเที่ยงตรงของเครื่องมือวัด	106
ความเที่ยงตรงกับความแปรปรวน	107
ประเภทความเที่ยงตรง	111
ขั้นตอนดำเนินการในการสร้างเครื่องมือวัดงานบุคคลที่เที่ยงตรง	118
การแปลความหมายของค่าความเที่ยงตรง	120
ปัญหาบางประการเกี่ยวกับความเที่ยงตรง	125
บทที่ 6 มาตรฐานวัดความสำเร็จของงาน	127
แนวคิดในการสร้างมาตรฐานวัดความสำเร็จของงาน	128
มาตรฐานวัดความสำเร็จของงานที่ได้โดยวิธีปรนัย	132
มาตรฐานวัดความสำเร็จของงานที่ได้จากการวินิจฉัยส่วนบุคคล	136
การประเมินมาตรฐานวัดความสำเร็จของงาน	144
บทที่ 7 การวิเคราะห์องค์ประกอบ	147
แนววิเคราะห์ทางคณิตศาสตร์	148
รูปแบบในการวิเคราะห์องค์ประกอบ	162
วิธีการแปลความหมายขององค์ประกอบ	167
บทที่ 8 ทฤษฎีการวัดเชาว์	174
ความเป็นมาของทฤษฎีการวัดเชาว์	175
ทฤษฎีโครงสร้างของเชาว์	179
วิธีการวัดเชาว์	186
เชาว์กับการปฏิบัติงาน	191

	หน้า
บทที่ 9 ความถนัดพิเศษและทักษะทางกาย	197
ความถนัดในงานสารบรรณ	198
ความถนัดด้านเครื่องจักรกล	202
แบบทดสอบผสมที่ใช้วัดความถนัดพิเศษกับความเที่ยงตรง	204
ความคล่องของกล้ามเนื้อ	206
ความสามารถของประสาทตา	208
บทที่ 10 ทฤษฎีการวัดบุคลิกภาพ	211
วิธีการในการวัดบุคลิกภาพ	214
ทฤษฎีองค์ประกอบและทฤษฎีลำดับลดหลั่นของบุคลิกภาพ	223
บทที่ 11 เทคนิคการวัดทัศนคติ	230
ความหมายของทัศนคติ	231
หลักการของเครื่องวัดช่วงเท่ากัน	233
หลักการของเครื่องวัดค่าร่วม	237
แนวคิดในการวัดทัศนคติ	240
บทที่ 12 เทคนิคการวัดความสัมฤทธิ์ผล	248
การวางแผนในการสร้างแบบทดสอบ	249
การวิเคราะห์รายการคำถาม	252
การให้ค่าคะแนน	258
ลักษณะของแบบทดสอบมาตรฐานที่วัดความสัมฤทธิ์ผล	260
บทที่ 13 ปัญหาในการบริหารงานวัดผลบุคคล	263
หน่วยวัดผล	264
การดำเนินการวัดผล	269
บทที่ 14 การใช้ผลการวัดในงานบุคคล	279
ความสัมพันธ์ระหว่างข้อวินิจฉัยสั่งการกับการใช้เครื่องมือวัด	279
เครื่องมือวัดในขั้นตอนต่างๆ ของงานบุคคล	284
การวัดเจ้าหน้าที่สำนักงาน	290
การวัดเจ้าหน้าที่ระดับผู้บริหารงาน	293
บรรณานุกรม	302

สารบัญตาราง

	หน้า
ตารางที่ 1	
1 ค่าคะแนนตามเกณฑ์เปอร์เซ็นต์ไต่ล้งของแบบทดสอบ ชุด DAT สำหรับเพศชาย	26
2 ศักยภาพในการคาดคะเนของเครื่องมือวัด	63
3 ค่าสหสัมพันธ์สมมติระหว่างเครื่องวัดกับองค์ประกอบ	68
4 ตารางสัมพันธ์แบบเอฟ ของชุดเครื่องมือวัดความเป็นผู้นำ (สมมติ)	71
5 ตารางสัมพันธ์แบบอาร์ ของชุดเครื่องมือวัดความเป็นผู้นำ (สมมติ)	72
6 ตารางสัมพันธ์แบบอาร์ ระหว่างเครื่องมือวัด 9 ชุด ของกองทัพอากาศสหรัฐ	73
7 ตารางสัมพันธ์แบบเอฟ ของชุดเครื่องมือวัดความมีปฏิภุมานะ	75
8 ค่าผิดพลาดมาตรฐานของการวัดเทียบกับค่าความเชื่อมั่น	86
9 วิธีคำนวณสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบมาตรฐาน บางแบบ	94
10 ความแปรปรวนกับวิธีหาค่าความเชื่อมั่นวิธีต่าง ๆ	96
11 ค่าผิดพลาดมาตรฐานของ Revised Stanford-Binet	99
12 ค่าผิดพลาดมาตรฐานของ Wechsler Intelligence Scale for Children	100
13 เปรียบเทียบค่าความเชื่อมั่นของวิธีคำนวณบางแบบ	102
14 ค่าความแม่นยำของสัมประสิทธิ์ความเชื่อมั่น	104
15 ค่าสูงสุดของ r_{xy} เมื่อเทียบกับค่า r_{xx} และ r_{yy}	110

	หน้า
16 ค่าความเที่ยงตรงเทียบกับมาตรฐานวัดความสำเร็จของงาน	121
17 โอกาสผิดพลาดของเครื่องมือวัดที่มีค่าความเที่ยงตรงระดับต่างๆ ในการคัดเลือกคนไว้ 50 เปอร์เซนต์	122
18 ความเที่ยงตรงเทียบกับค่ามาตรฐานวัดความสำเร็จของงาน	123—124
19 ตารางแบบเอฟสองมิติ	157
20 ตารางแบบเอฟสามมิติ	160
21 ค่าองค์ประกอบเมื่อหมุน vector 50°	162
22 ตารางแบบเอฟแสดงด้วยตัวอักษร	168
23 ค่าองค์ประกอบที่คำนวณโดยวิธีช่วงทแยง	171
24 ค่าองค์ประกอบที่หมุนแกนแล้ว	172
25 องค์ประกอบที่เกี่ยวกับการรู้ของเซวาน์	183
26 ความสัมพันธ์ระหว่างเซวาน์กับอาชีฟ	193—194
27 เกณฑ์เฉลี่ยของความสามารถในการคะแนนของแบบทดสอบเซวาน์	195
28 ความตรงประเด็นในการคาดคะเนของแบบทดสอบด้านต่างๆ	205
29 ดัชนีความยากของคำถามที่ปรับค่าความเผชิญแล้ว	156—257
30 ค่าสมมติของความสามารถของบุคคลสามคนเป็นเปอร์เซนต์ไต้	285
31 คะแนนคัดเลือกกับผลการอบรม	288
32 ผลการวัดเซวาน์กับความสามารถในหน้าที่นักบริหาร	300

สารบัญแผนภาพ

	หน้า
แผนภาพที่ 1	
1 เปรอร์เซ็นไทล์เทียบกับอัตราร้อยละของผู้ถูกวัด	27
2 แบบสอบถามหน้าที่และความรับผิดชอบ	37-38
3 รายการตรวจสอบกิจกรรมการใช้เหตุผลในการแก้ปัญหา เพื่อปฏิบัติงาน	48
4 ตัวอย่างแบบฟอร์มรายงานการวิเคราะห์งาน	50-51
5 รายงานอธิบายและกำหนดงาน	51-52
6 มาตรฐาน ชั้นปฐมและชั้นรอง	131
7 ตัวอย่างวิธีการให้ค่าแบบกราฟ	137
8 ลักษณะองค์ประกอบแสดงด้วยกราฟ	154
9 vectors หลายเส้นที่มีเพียงสองมิติ	156
10 vectors ที่แสดงมิติที่ถูกต้อง	158
11 vectors สามเส้นที่มีสามมิติ	158
12 vectors หลายเส้นที่มีสามมิติ	159
13 แผนภูมิความแปรปรวนร่วมของ Cattell	165
14 องค์ประกอบที่หมุนแกนแล้ว	171
15 โครงสร้างของเซาวันที่มีสามมิติ	180
16 ลำดับลดหลั่นขององค์ประกอบของเซาวัน	185
17 โครงแบบของลำดับลดหลั่นองค์ประกอบสามชั้น	186
18 ตัวอย่าง Minnesota Clerical Test	200
19 ตัวอย่างแบบทดสอบ MPFB	203
20 ลำดับลดหลั่นของบุคลิกภาพ	228
21 ผลการวัดบุคลิกแสดงเป็นตาราง	276
22 ผลการวัดบุคลิกแสดงเป็นกราฟ	277

สารบัญแผนภูมิ

แผนภูมิที่ 1	ค่าความแปรปรวนขององค์ประกอบร่วมกันที่สมมติ	69
2	องค์ประกอบตามทฤษฎี Spearman	77

บทที่ 1

ลักษณะเบื้องต้นของการวัดผล

การวัดผลบุคคลเป็นสาขาวิชาการที่อาศัยผลการศึกษาวิจัยของวิชาการในสาขาต่างๆ หลายสาขา วิชาคณิตศาสตร์และจิตวิทยา นับเป็นวิชาการสาขาที่มีบทบาทสำคัญต่อ วิชาการวัดผลบุคคล ฉะนั้น ก่อนที่จะพิจารณารายละเอียดของวิชานี้ ซึ่งจะได้กล่าวไว้ในบทต่อๆ ไป จึงควรที่จะพิจารณาหลักการและผลการวิจัยบางส่วนที่เป็นพื้นฐานของ วิชานี้ ในบทนี้จะกล่าวถึงความหมายของการวัดผล ระดับของการวัด ประวัติ ผลการ วิจัยทางจิตวิทยาในเรื่องความแตกต่างของบุคคล และลักษณะทางจิตวิทยาที่วัดได้ ตลอดจนลักษณะต่างๆ ของเครื่องมือวัดที่ถือเป็นมาตรฐานได้ เพื่อเป็นการสร้างความ เข้าใจพื้นฐานของการวัดผลเสียก่อน

การวัดผลกับการประเมินบุคคล

ในตำราภาษาอังกฤษที่ใช้อ้างอิง ได้อธิบายความหมายของการวัดผลว่าเป็นการ เปรียบเทียบปริมาณ เพื่อแสดงค่าตัวเลข และต่างกับการประเมินซึ่งหมายความถึง ขบวนการในการตัดสินความสำคัญของปรากฏการณ์อย่างหนึ่ง เมื่อเทียบกับปรากฏ การณ์อีกอย่างหนึ่งโดยมีมาตรฐานที่กำหนดไว้¹ การอธิบายความหมายของการวัดผล เช่นนี้ มีลักษณะคล้ายคลึงกับที่มักนิยามกันในหนังสือวิชาการทางการวัดผลต่างๆ ซึ่ง มักจะกล่าวว่า การวัดผลเป็นการจัดค่าตัวเลขให้แก่วัตถุ หรือเหตุการณ์ โดยมีกฎเกณฑ์ ส่วนการประเมินผล เป็นการตัดสินค่านิยมของการกระทำ² การที่กล่าวว่า การวัดผล

¹ Good, C.V., ed., *Dictionary of Education*, pp. 209, 337–338.

² Guilford, J.P., *Psychometric Methods*, p. 5.
Shaw, M.E. and Wright, J.M., *Scales for the Measurement of Attitudes*, p. 15.
Travers, R.M.W., *Educational Measurement*, pp. 6–8.

เป็นการจัดค่าตัวเลขโดยมีกฎเกณฑ์นั้น หมายความว่า ในการวัดผลทุกครั้ง ผู้วัดจะต้องจัดประเภทของสิ่งที่ถูกวัดได้ ในกรณีการวัดผลบุคคล ไม่ว่าจะเป็นการวัดว่าบุคคลหนึ่ง ๆ ได้มีคุณสมบัติอะไรบ้างแล้ว หรือเป็นการวัดว่า บุคคลหนึ่ง ๆ จะมีคุณสมบัติอะไรบ้างในโอกาสต่อไป ผู้วัดก็ต้องจัดบุคคลที่ถูกวัดออกเป็นประเภท ๆ ได้ จึงจะเรียกว่ามีการวัดผล ประเภทของบุคคลที่จะวัดเมื่อถูกวัดแล้วนั้น ขึ้นอยู่กับความมุ่งหมายของการวัด เช่น ถ้ามีความมุ่งหมายที่จะรับเข้าทำงาน หรือไม่รับเข้าทำงาน ก็จะต้องจัดบุคคลที่ถูกวัดออกเป็นสองประเภท ได้แก่ ประเภทที่ควรให้ทำงาน และประเภทที่ไม่ควรให้ทำงาน แต่ถ้ามีความมุ่งหมายที่จะวัดเพื่อให้ทำงานในหน้าที่ต่าง ๆ ประเภทที่จัดแบ่ง ก็จะต้องมีตั้งแต่สองประเภทขึ้นไป เช่น ประเภทที่จะต้องฝึกอบรมก่อน ประเภทที่จะต้องเข้าทำงานในตำแหน่งต่าง ๆ ส่วนที่กล่าวมา การประเมินเป็นการตัดสินค่านิยมนั้น หมายความว่า ผู้ประเมินต้องการทราบว่า การที่สิ่งที่ถูกประเมินเป็นเช่นนั้นดีหรือไม่ แต่ไม่ได้วัดว่า สิ่งที่ถูกประเมินนั้น เป็นอย่างไร ในกรณีของการประเมินบุคคล ผู้ประเมินต้องการทราบว่า การที่บุคคลหนึ่ง ๆ มีคุณสมบัติอย่างหนึ่ง ๆ นั้น เป็นที่พึงปรารถนาหรือไม่ ผู้ประเมินไม่ต้องการทราบว่า บุคคลนั้น ๆ มีการเปลี่ยนแปลงอย่างไร

ในการวัดผลบุคคล ค่าตัวเลขของสิ่งที่ถูกวัดจะแสดงปริมาณของคุณสมบัติทั้งหลายของบุคคล กล่าวอีกนัยหนึ่ง การวัดบุคคลไม่ใช่เป็นการวัดตัวบุคคล แต่เป็นการวัดคุณสมบัติต่าง ๆ ของบุคคล เช่น ความฉลาด ความกระตือรือร้น เป็นต้น การวัดผลบุคคลจึงต่างกับการวัดผลทางกายภาพ (physical measurement)³ เพราะว่าการวัดผลทางกายภาพนั้น ลักษณะคุณสมบัติของเครื่องมือวัดจะตรงกับลักษณะคุณสมบัติของสิ่งที่ถูกวัดผลของการวัดจะแน่นอน เช่น เครื่องมือวัดความยาว ความหนัก หรือความหนักก็ย่อมวัดคุณสมบัติในเรื่องนั้น ๆ ออกมาได้โดยตรง แต่ในการวัดบุคคลนั้น คุณสมบัติของเครื่องมือวัด ไม่ได้ตรงกับคุณสมบัติของสิ่งที่วัด สิ่งที่จะวัดในการวัดบุคคล เป็นเรื่องของพฤติกรรม หรือ ลักษณะนิสัยทางจิตวิทยา สิ่งที่จะวัดเช่นนั้นจึงเป็นเรื่องของนามธรรม

การวัดที่ถูกต้อง จะต้องเป็นการวัดที่มีมาตรฐาน ดังได้กล่าวมาแล้วว่า ค่าตัวเลขที่ผู้วัดกำหนดให้แก่วัตถุ หรือเหตุการณ์ หรือคุณสมบัติต่าง ๆ ของบุคคล จะต้องดำเนินไปโดยมีกฎเกณฑ์ กฎเกณฑ์นี้หมายถึง การดำเนินการอย่างมีระเบียบมาตรฐานในขอบวน

³ Berdie, R.F.; Layton, W.L.; Swanson, E.O.; and Hagenah, T., *Testing in Guidance and Counselling*, p. 69.

การวัด เช่น การกำหนดวิธีที่ผู้ถูกวัดจะต้องปฏิบัติ การจัดเวลาในการวัด วิธีการให้ค่าตัวเลข วิธีการแปลความหมายของตัวเลขที่วัดได้ เป็นต้น การวัดโดยมีกฎเกณฑ์เช่นนี้จึงเรียกว่า เป็นการวัดที่มีมาตรฐาน (standard measurement) อย่างไรก็ตามกฎเกณฑ์ที่จะใช้เป็นหลักในการวัดนี้ย่อมกำหนดไม่ได้แน่นอนตายตัว⁴ ผู้วัดคนหนึ่ง ๆ จำเป็นต้องตั้งสมมติฐานของตนเองให้สอดคล้องกับสิ่งที่จะวัด และเมื่อได้ทดสอบสมมติฐานนี้ตามหลักวิธีการวิจัยแล้ว ก็จะทำให้ถือได้ว่า สามารถใช้สมมติฐานนี้เป็นกฎเกณฑ์ในการวัดเรื่องนั้น ๆ สมมติฐานที่จะใช้เป็นกฎเกณฑ์ในการวัดได้นี้ จะต้องแสดงได้ชัดเจนว่า สามารถวัดปริมาณของคุณสมบัติของวัตถุ หรือเหตุการณ์นั้น ๆ ได้

การประเมินสัมพันธภาพกับการวัดอย่างใกล้ชิด การประเมินที่อาศัยค่าของการวัดที่มีมาตรฐานประกอบ ถือว่าเป็นการประเมินที่มีระเบียบ (formal evaluation) การประเมินเช่นนี้ต้องมีเกณฑ์ (criteria) แน่แน่นอน ผลที่ได้จากการวัดที่มีมาตรฐานเมื่อนำมาประกอบการพิจารณา จะทำให้ทราบได้ว่าวัตถุ หรือเหตุการณ์ หรือบุคคลที่ประเมินดีตามที่ต้องการหรือไม่ วิธีการต่าง ๆ ที่จะกล่าวในบทต่อไปทั้งหมด เป็นวิธีที่จะหาค่าตัวเลขโดยการวัดที่มีมาตรฐานทั้งสิ้น อย่างไรก็ตาม เป็นที่น่าเสียดายว่า ผู้ที่มีหน้าที่ประเมินบางครั้งอาจจะใช้วิธีการประเมินที่ขาดหลักเกณฑ์ (informal evaluation) การประเมินโดยวิธีนี้ หมายความว่า ผู้ประเมินใช้แต่ดุลพินิจของตนเองเท่านั้นในการตัดสินค่านิยมของสิ่งที่จะถูกประเมิน เช่น ผู้บังคับบัญชาอาจใช้ดุลพินิจพิจารณาความเหมาะสมของผู้อยู่ใต้บังคับบัญชา แล้วจึงดำเนินการเลื่อนขั้นตำแหน่ง หรือโยกย้ายสับเปลี่ยนตำแหน่ง การประเมินเช่นนี้ไม่ได้อาศัยค่าตัวเลขที่ได้จากการวัดที่มีมาตรฐานประกอบการพิจารณาการที่เป็นเช่นนี้ก็อาจเป็นเพราะว่า การหาค่าตัวเลขโดยการวัดที่มีมาตรฐานนั้นเป็นเรื่องที่ปฏิบัติได้ยาก เพราะมีกฎเกณฑ์หลายอย่าง และยังอาจจะต้องสิ้นเปลืองเวลาและเงินทองเป็นจำนวนมาก แต่อย่างไรก็ดีในแง่วิชาการ การที่จะให้การประเมินบุคคลมีหลักเกณฑ์โดยอาศัยการวัดที่มีมาตรฐานนั้น แม้จะยากลำบากสิ้นเปลืองเวลาและเงินทอง ก็เป็นสิ่งที่ควรสนับสนุนเป็นอย่างยิ่ง

การวัดผลกับแนวคิดทางคณิตศาสตร์

การที่การวัดผลต้องแสดงค่าตัวเลขของสิ่งที่ถูกวัดได้นั้น ทำให้ต้องอาศัยหลักทางคณิตศาสตร์มาช่วยในการดำเนินการ ความจริงแล้ว คณิตศาสตร์กับการวัดผลเป็นสิ่งที่ใกล้เคียงกันมาก การศึกษาวิชาการวัดผลจะไม่เห็นจริงถ้าหากว่าไม่ศึกษาคณิตศาสตร์

⁴ Nunnally, J.C., *Psychometric Theory*, p. 4.

ควบคุมไปด้วย คณิตศาสตร์เป็นสิ่งที่มนุษย์ประดิษฐ์ขึ้น เพื่ออธิบายปรากฏการณ์ต่าง ๆ ของโลก ในทางคณิตศาสตร์มีหลักการอยู่ประการหนึ่งที่เรียกว่า ความเสมอเหมือนกันของการปรากฏ (isomorphism)⁵ หลักนี้อธิบายว่า การที่มนุษย์ใช้ตัวเลขต่าง ๆ แทนค่าปรากฏการณ์ของโลกได้นั้น ก็เพราะว่า โครงสร้างของธรรมชาติมีคุณสมบัติที่เสมอเหมือนกันกับระบบตรรกวิทยาของคณิตศาสตร์ คณิตศาสตร์ไม่ได้มีอยู่ในโลกนี้มาก่อน มนุษย์ กล่าวอีกนัยหนึ่งก็คือ ปรากฏการณ์ของโลกไม่ได้เกิดขึ้นตามหลักคณิตศาสตร์ เพราะฉะนั้น จึงกล่าวได้ว่า มนุษย์ไม่ได้ค้นพบคณิตศาสตร์ แต่มนุษย์ประดิษฐ์คณิตศาสตร์ ปรากฏการณ์บางอย่างของโลกมีความเสมอเหมือนกันกับระบบคณิตศาสตร์อย่างใกล้ชิด แต่บางอย่างก็มีความเสมอเหมือนกันในระดับต่ำ เช่น ในการนำหลักการแจกแจงภายในเส้นโค้งปกติมาเทียบกับปรากฏการณ์ของโลก มนุษย์ก็สามารถทดสอบคุณสมบัติของความเหมาะสม (goodness of fit) โดยวิธีหา chi square ถ้าค่า chi square ที่ได้มีปริมาณน้อย ก็แปลว่ารูปแบบของเส้นโค้งปกติที่สร้างขึ้นนั้นน่าจะใช้ได้กับข้อมูลที่วิจัย แต่ถ้าค่าที่ได้มีปริมาณมาก ก็แปลว่ารูปแบบที่สร้างขึ้นนั้นมีความน่าจะเป็นน้อย การที่ได้ทราบว่ารูปแบบของเส้นโค้งปกติที่สร้างขึ้นนี้ใช้ได้กับข้อมูลที่วิจัยหรือไม่ ทำให้มนุษย์สามารถลงความเห็นเกี่ยวกับปรากฏการณ์ต่าง ๆ ของโลก และคาดคะเนสิ่งต่าง ๆ ได้

ตามหลักความเสมอเหมือนกันของการปรากฏนี้ ทำให้นักวิชาการสามารถนำตัวเลขมาแสดงค่าของการวัดวัตถุหรือเหตุการณ์ต่าง ๆ ได้ โดยมีข้อแม้ว่า คุณสมบัติของตัวเลขที่นำมาใช้นั้นจะต้องมีความเสมอเหมือนกันกับคุณสมบัติของวัตถุหรือเหตุการณ์ที่จะวัด โดยหลักคณิตศาสตร์ที่เป็นที่อ้างอิงกันทั่วไป ตัวเลขมีคุณสมบัติ 3 ประการ คือ แสดงความมีตัวตน (identity) แสดงค่าลำดับ (rank order) และแสดงผลของการเพิ่ม (additivity) ในการแสดงลักษณะทั้งสามประการนี้ อาจจำแนกออกเป็นรูปแบบได้ถึง 9 แบบ คือ

1. $a = b$ หรือ $a \neq b$
2. ถ้า $a = b$ ฉะนั้น $b = a$
3. ถ้า $a = b$ และ $b = c$ ฉะนั้น $a = c$
4. ถ้า $a < b$ ฉะนั้น $b < a$
5. ถ้า $a > b$ และ $b > c$ ฉะนั้น $a > c$
6. ถ้า $a = p$ และ $b > 0$ ฉะนั้น $a + b > p$

⁵ Guilford, *op. cit.*, pp. 6-7.

7. $a + b = b + a$

8. ถ้า $a = p$ และ $b = q$ ฉะนั้น $a + b = p + q$

9. $(a + b) + c = a + (b + c)$

รูปแบบ 3 แบบแรก เป็นการแสดงความมีตัวตนทางคณิตศาสตร์ รูปแบบที่ 1 แสดงว่า สิ่งหนึ่งอาจจะเหมือนกับสิ่งหนึ่ง หรืออาจจะไม่เหมือนกับอีกสิ่งหนึ่ง รูปแบบที่ 2 แสดงความเสมอเหมือนกันของสิ่งที่ต่างกัน และรูปแบบที่ 3 แสดงผลของความเสมอเหมือนกันระหว่างสิ่งต่าง ๆ หลาย ๆ สิ่ง รูปแบบอีก 2 แบบถัดไปนั้นเป็นการแสดงลำดับทางคณิตศาสตร์ รูปแบบที่ 4 แสดงลักษณะการขาดความเสมอเหมือนกัน และรูปแบบที่ 5 แสดงผลของการขาดความเสมอเหมือนกันระหว่างสิ่งหลาย ๆ สิ่ง ส่วนรูปแบบ 4 แบบที่เหลือนั้น แสดงผลของการเพิ่ม รูปแบบที่ 6 แสดงโอกาสที่จะเพิ่มระหว่างสิ่งสองสิ่ง ถึงแม้ว่าสิ่งใดสิ่งหนึ่งจะมีค่าเท่ากับศูนย์ก็ตาม เมื่อนำสิ่งนั้นไปเพิ่มกับสิ่งอื่น ๆ ก็มิได้ทำให้ค่าของสิ่งอื่น ๆ นั้นด้อยลงไป รูปแบบที่ 7 แสดงให้เห็นว่าคุณสมบัติในการเพิ่มของสิ่งต่าง ๆ นั้น ไม่ได้ขึ้นอยู่กับลำดับที่ของการจัดตั้งเหล่านั้น รูปแบบที่ 8 แสดงว่า ถ้าหากว่าสิ่งหนึ่งสิ่งใดมีความเสมอเหมือนกับสิ่งอื่น ๆ แล้ว ก็ย่อมสามารถนำไปแทนค่าในการเพิ่มได้ รูปแบบสุดท้ายนั้น ชี้ให้เห็นว่า แม้จะมีสิ่งต่าง ๆ มากมายหลาย ๆ อย่าง เมื่อต้องการทราบผลลัพธ์ของการเพิ่ม มนุษย์ก็อาจจะนำมาจัดเรียงลำดับได้ตามความพอใจ

มิติและระดับการวัดผล

การวัดลักษณะต่าง ๆ ของร่างกาย เช่น น้ำหนักหรือความสูง เป็นการวัดที่จะหาค่าของสิ่งต่าง ๆ แต่ละสิ่ง โดยไม่ต้องอาศัยผลการวัดของสิ่งอื่น ๆ การวัดเช่นนี้เป็น การวัดมิติเดียว (unidimensional) ผลที่วัดได้ในเรื่องนั้น ๆ มีความมากมายแสดงเป็นค่าตัวเลขได้โดยตัวของตัวเอง แต่การวัดผลบุคคล ซึ่งจะต้องแสดงค่าตัวเลขของลักษณะและคุณสมบัติของบุคคล เป็นการวัดสิ่งซึ่งไม่ได้แสดงค่าได้ด้วยตัวของมันเอง ลักษณะและคุณสมบัติของบุคคล เช่น เซวน์ ความถนัดพิเศษทางธรรมชาติ ความสัมพันธ์ผลในการปฏิบัติงาน และลักษณะต่าง ๆ ของบุคลิกภาพ เป็นสิ่งที่ประกอบด้วยหลายมิติ โดยเฉพาะลักษณะต่าง ๆ ของบุคลิกภาพนั้น เป็นสิ่งที่เกี่ยวโยงกันอย่างใกล้ชิด ได้มีการพิสูจน์ไว้แน่ชัดแล้วว่า หากจะมีการวัดลักษณะต่างของบุคลิกภาพ เช่น อารมณ์ (temperament) จะต้องมีการวัดองค์ประกอบอย่างน้อยถึง 10 อย่าง ได้แก่

⁶ Shaw and Wright, *op. cit.*, pp. 6–10.

กิจกรรมทั่วไป (general activity) การควบคุมตน (restraint) ความรู้สึกว่าตนสูงกว่าผู้อื่น (ascendance) การชอบสมาคม (sociability) ความมั่นคงทางอารมณ์ (emotional stability) การมองโลกในแง่ของความเป็นจริง (objectivity) ความเป็นมิตร (friendliness) ความช่างคิด (thoughtfulness) ความสัมพันธ์กับบุคคลอื่น (personal relations) และความเป็นผู้ชาย (masculinity)⁷ แม้ลักษณะอื่นๆ ของบุคคล เช่น เซวน์ หรือความถนัดพิเศษทางธรรมชาติก็เช่นเดียวกัน ย่อมไม่สามารถวัดได้เพียงมิติเดียว ในเรื่องของเซวน์นั้น ได้มีผู้พิสูจน์ว่า จะต้องวัดอย่างน้อย 3 มิติ⁸ และในเรื่องความถนัดพิเศษทางธรรมชาตินี้ ก็ล้วนแต่มีหลายมิติตามแต่ลักษณะพิเศษหนึ่งๆ การวัดเช่นนั้น เป็นการวัดหลายมิติ (multidimensional)⁹ การที่ลักษณะและคุณสมบัติของบุคคลมีหลายมิติเช่นนั้น ย่อมเป็นการเห็นได้ชัดว่า ทำให้การวัดบุคคลเป็นเรื่องที่ยุ่งยากซับซ้อน ค่าตัวเลขที่วัดได้เกี่ยวกับลักษณะและคุณสมบัติของบุคคลในเรื่องหนึ่งๆ ไม่ว่าจะเป็นการวัดความสามารถในการปฏิบัติงาน หรือความสนใจที่จะปฏิบัติงาน จะถือเอาเป็นสิ้นสุดเหมือนอย่างการวัดน้ำหนักหรือความสูงของบุคคลไม่ได้

ค่าตัวเลขที่วัดผลได้ในเรื่องต่างๆ และโอกาสต่างๆ มิได้มีคุณภาพที่เทียบกันทุกครั้ง ค่าตัวเลขจะมีคุณภาพสูงต่อเมื่อการวัดนั้นๆ มีกฎเกณฑ์ในการให้ค่ารัศม หรือกล่าวอีกอย่างหนึ่งในแง่คณิตศาสตร์ว่า ใช้รูปแบบทางคณิตศาสตร์ทั้ง 9 แบบอย่างมากที่สุดเท่าที่จะทำได้ การวัดผลเช่นนั้นเรียกได้ว่าเป็นการวัดที่มีคุณภาพสูง ฉะนั้นจึงแบ่งการวัดออกเป็น 4 ระดับ คือ การวัดที่แสดงประเภทหรือชนิด (nominal scale) มาตรการลำดับ (ordinal scale) มาตรการช่วงระยะ (interval scale) และมาตรการอัตราส่วน (ratio scale)¹⁰

การวัดที่แสดงประเภทหรือชนิด การวัดในระดับนี้ แสดงค่าตัวเลขที่ใช้แทนความเป็นตัวตนของสิ่งที่ถูกวัด เช่น วัดได้ผลออกเป็นกลุ่มที่ 1 และกลุ่มที่ 2 หรือประเภทที่ 1 และประเภทที่ 2 การวัดประเภทนี้ใช้รูปแบบทางคณิตศาสตร์เพียง 3 แบบแรก ซึ่งแสดงได้แต่เพียงว่าสิ่งต่างๆ ที่วัดนั้น เท่ากันหรือต่างกันอย่างใด การวัดในระดับนี้เป็น การวัดที่มีคุณภาพต่ำที่สุด แต่เนื่องจากการวัดในระดับนี้ทำได้สะดวกที่สุด จึงเป็นที่

⁷ Guion, R.M., *Personal Testing*, p. 317.

⁸ *Ibid.*, pp. 66—77.

⁹ Guilford, *op. cit.*, pp. 470—471.

¹⁰ Schuessler, K.F., *Social Research Methods*, pp. 32—34.
Guilford, *op. cit.*, pp. 11—17.

แพร่หลายกันในระหว่างผู้บริหารงานบุคคลหลายกลุ่ม และบางครั้งก็ทำให้เกิดความเข้าใจผิด เชื่อว่าผลของการวัดนั้น ๆ เชื่อถือได้แล้ว

มาตราลำดับ เมื่อผู้วัดจัดสิ่งต่าง ๆ ให้เป็นลำดับตามมาตรฐานการวัดอย่างหนึ่งอย่างใด เช่น จัดให้เป็นลำดับที่ 1 ที่ 2 ที่ 3 ไปเรื่อย ๆ จนถึงที่สุดท้าย การจัดลำดับอย่างนี้เรียกว่า จัดโดยอาศัยมาตราลำดับ ในการจัดลำดับเช่นนี้ ผู้วัดย่อมไม่ได้แยกแยะความแตกต่างในเชิงปริมาณระหว่างลำดับหนึ่งถึงอีกลำดับหนึ่ง กล่าวคือ ไม่ทราบว่าที่ 1 มากกว่าที่ 2 เท่ากับที่ 2 มากกว่าที่ 3 หรือเปล่า แต่ผู้วัดสนใจแต่เพียงว่าที่ 1 มากกว่าที่ 2 และที่ 2 มากกว่าที่ 3

มาตราช่วงระยะ เมื่อผู้วัดระบุแต่เพียงความแตกต่างระหว่างหน่วยแต่ละ 2 หน่วย เรียกว่าผู้วัดนั้นได้ใช้มาตราช่วงระยะเป็นแนวการบันทึกข้อสังเกต ผู้วัดอาจจะบันทึกแต่เพียงความแตกต่างระหว่างคนที่ 1 กับคนที่ 2 และที่ 3 ความแตกต่างระหว่างคนที่ 2 และคนที่ 3 กับคนที่ 4 และความแตกต่างระหว่าง...ไปเรื่อย ๆ ฉะนั้นค่าตัวเลขที่วัดได้ก็เป็น 2, 1, 2, 1 ปริมาณที่วัดไว้นี้ อาจจะเพิ่มขึ้นหรือลดลงเท่าไรก็ได้ เช่น อาจจะเป็น 4, 2, 4, 2 หรือ 20, 10, 20, 10 หรือ ฯลฯ การวัดเช่นนี้ทำให้ทราบปริมาณความแตกต่างของหน่วยแต่ละ 2 หน่วยได้ชัดเจนยิ่งขึ้น แต่อย่างไรก็ดี การวัดแต่เพียงความแตกต่างระหว่างหน่วยแต่ละ 2 หน่วยนี้ ย่อมอาจจะทำให้การพิจารณาข้อเท็จจริงบิดเบือนไปก็ได้ เช่น ความแตกต่างเป็นปริมาณ 10 อาจจะมีค่าไม่เหมือนกันระหว่าง 1,000 กับ 1,010 และระหว่าง 10 กับ 20

มาตราอัตราส่วน เมื่อผู้วัดทราบปริมาณแน่นอนของหน่วยแต่ละหน่วย ก็อาจจะบันทึกผลของการวัดออกมาเป็นมาตราอัตราส่วนได้ ในการบันทึกโดยอาศัยมาตราอัตราส่วนนี้ ผู้วัดจะนำปริมาณของแต่ละหน่วยมาจัดเรียงตามลำดับแล้วตีค่าของปริมาณของหน่วยแต่ละสองหน่วยที่ใกล้เคียงกัน เมื่อได้ทราบค่าความแตกต่างที่สัมพันธ์กันทั้งหมดนี้แล้ว ก็แสดงตัวเลขให้เห็นว่าแต่ละหน่วยมีค่าเป็นปริมาณเท่าตัวของหน่วยที่ลำดับรองลงไป ยกตัวอย่างเช่น พนักงานพิมพ์ดีด 6 คน พิมพ์ข้อความอย่างเดียวกันในเวลาเท่ากัน ด้วยเครื่องพิมพ์ชนิดเดียวกันที่มีสภาพใกล้เคียงกัน ผลปรากฏว่าพนักงานทั้ง 6 คน พิมพ์คำผิดเป็นจำนวนดังนี้ คนที่ 1 พิมพ์ผิด 10 คำ คนที่ 2 และที่ 3 คนละ 8 คำ คนที่ 4 ผิด 7 คำ คนที่ 5 ผิด 5 คำ คนที่ 6 ผิด 4 คำ เมื่อผู้วัดนำปริมาณเหล่านี้มาพิจารณาก็จะเห็นว่า คนที่ 1 พิมพ์ผิดมากกว่าคนที่ 2 และที่ 3 เป็นจำนวน 2 คำ มากกว่าคนที่ 4 เป็นจำนวน 3 คำ มากกว่าคนที่ 5 จำนวน 5 คำ และมากกว่าคนที่ 6 จำนวน

6 คำ ส่วนคนที่ 2 และที่ 3 พึมพำผิดมากกว่าคนที่ 4 เป็นจำนวน 1 คำ และมากกว่า...
เป็นดังนี้ไปเรื่อย ๆ ผู้วัดก็จะสามารถบันทึกลงไปได้ว่าผู้ใดผิดมากกว่าผู้ใดเป็นกี่เท่าตัว
และในขณะเดียวกัน ก็อาจนำปริมาณการผิดนี้มารวมกันแล้วหาอัตราเฉลี่ยการวัด โดย
อาศัยมาตราอัตราส่วนนี้ย่อมได้เปรียบกว่ามาตราลำดับ เพราะไม่เพียงแต่ทราบว่ามีผู้ใด
หรือสิ่งใดควรอยู่ลำดับใด แต่ยังได้ทราบเพิ่มว่ามี ความแตกต่างระหว่างบุคคลหรือสิ่ง
ของนั้น ๆ อย่างไรอีกด้วย

ในการวัดผลบุคคล นักทฤษฎีในการวัดได้คิดค่าดัชนี เพื่อใช้เฉพาะกับการวัด
บุคคลหลายอย่าง ซึ่งควรนำมาพิจารณาในที่นี้ ได้แก่ค่าระดับเชาว์ (Intelligence
quotient) และสัมประสิทธิ์สหสัมพันธ์ ค่าระดับเชาว์นั้น เป็นอัตราส่วนระหว่าง
ระดับของเชาว์ที่วัดได้โดยเทียบเป็นจำนวนปีของอายุ กับอายุจริงของผู้ถูกวัด ค่าระดับ
เชาว์นั้นเมื่อพิจารณาเห็น ๆ จะเห็นได้ว่าเป็นการวัดที่แสดงอัตราส่วน เพราะสามารถ
แสดงได้ว่าบุคคลมีเชาว์ในระดับอายุเท่าใด แต่แท้ที่จริงแล้ว ค่าระดับเชาว์นั้นไม่ได้
แสดงช่วงระยะที่เท่า ๆ กันของเชาว์ระหว่างอายุต่าง ๆ ฉะนั้นค่าที่วัดได้จึงเป็นเพียง
การแสดงค่าตัวเลขตามมาตราลำดับเท่านั้น ส่วนสัมประสิทธิ์สหสัมพันธ์นั้น แม้จะ
แสดงค่าศูนย์ได้ แต่ก็ไม่สามารถแสดงความห่างเท่า ๆ กันระหว่างช่วงระยะต่าง ๆ ได้
 ฉะนั้น ก็ยังถือเป็นการวัดในระดับมาตราอัตราส่วนไม่ได้ เป็นเพียงการวัดในระดับ
มาตราลำดับเช่นเดียวกัน ต่อเมื่อนำค่านี้มายกกำลังสอง ทำให้กลายเป็นค่าความแปร
ปรวนขององค์ประกอบร่วมกัน ดังจะได้อธิบายรายละเอียดไว้ในบทที่ 3 ค่านี้ก็จะกลาย
เป็นค่าที่แสดงผลการวัดในระดับมาตราอัตราส่วน¹¹

ตามทฤษฎีทางจิตมิติ (psychometrics) ซึ่งเป็นหลักของการวัดผลบุคคลระบุว่า การ
วัดผลและประเมินบุคคลนั้น จะต้องมีความมาตรฐานสองประการจึงจะถือได้ว่าเป็นการวัด
ผลและประเมินบุคคลที่ดี ประการแรก จะต้องมีการระบุและจำกัดความมิติที่จะวัดให้
แน่นอนลงไป ในกรณีเรื่องลักษณะและคุณสมบัติของบุคคลที่มีหลายมิติเช่นนี้ หาก
จะวัดผลให้ได้ค่าเป็นที่เชื่อถือได้ ก็หมายความว่า จะต้องพิสูจน์ให้เห็นได้ว่า ค่าตัวเลข
นั้น ๆ แสดงผลของการวัดมิติใดบ้าง และประการที่สอง การวัดผลจะต้องแสดงค่า
ตัวเลขในระดับที่มีคุณภาพสูง มาตราของการวัดที่ได้กล่าวมาทั้งหมดนั้น แสดงให้
เห็นแล้วว่า การวัดผลและการวิจัยทางสังคมศาสตร์ยังประสบปัญหาในการที่จะใช้มาตรา
อัตราส่วน นับว่าการวัดผลทางสังคมศาสตร์ยังเสียเปรียบการวัดผลทางวิทยาศาสตร์

¹¹ *Ibid.*, pp. 17-18.

กายภาพและชีวภาพอยู่ อย่างไรก็ดี กล่าวได้สั้น ๆ ว่าการวัดผลที่แสดงแต่เพียงค่าตัวเลขเป็นประเภทหรือชนิด ย่อมเป็นการวัดผลที่พึงหลีกเลี่ยงทางวิชาการ¹²

ประวัติการวัดผลบุคคลโดยย่อ

การวัดผลบุคคล เป็นวิชาการที่เกิดหลังจากคณิตศาสตร์และจิตมิติได้มีความเจริญก้าวหน้าขึ้น วิชาคณิตศาสตร์ที่ว่าด้วยความน่าจะเป็นได้เริ่มศึกษากันในศตวรรษที่ 17 Bernoulli เป็นบุคคลแรกที่ได้พิมพ์เอกสารทางวิชาการเกี่ยวกับเรื่องนี้ ในศตวรรษที่ 18 De Moivre ได้คิดทฤษฎีการแจกแจงปกติ Gauss เป็นบุคคลสำคัญในศตวรรษที่ 19 ที่ได้คิดเส้นโค้งปกติขึ้น ทำให้เป็นหลักในการวิจัยทางสังคมศาสตร์ รวมทั้งการวัดผลบุคคล บุคคลสำคัญคนแรกที่นำเอาหลักคณิตศาสตร์มาประยุกต์กับจิตวิทยา ทำให้สามารถวัดค่าทางจิตวิทยาออกเป็นตัวเลขได้คือ Sir Francis Galton ผลงานของนักวิชาการผู้นี้ในปลายศตวรรษที่ 19 เป็นเรื่องเกี่ยวกับการวัดลักษณะและคุณสมบัติต่างๆ ของบุคคล Karl Pearson ได้มีส่วนทำให้งานของ Galton สมบูรณ์ขึ้น โดยได้คิดวิธีหาสัมประสิทธิ์สหสัมพันธ์และคะแนนมาตรฐาน ทำให้การวัดผลทางจิตวิทยามีค่าตัวเลขที่แน่นอนยิ่งขึ้น บุคคลสำคัญในศตวรรษที่ 20 หลายคนได้ร่วมกันคิดค้นวิธีการวัดให้เป็นที่เชื่อถือได้ยิ่งขึ้น เช่น James M. Cattell และ E.L. Thorndike ทั้งสองคนนี้เป็นลูกศิษย์ของ Wilhelm Wundt ซึ่งเป็นนักจิตวิทยาชาวเยอรมัน บุคคลในประเทศต่างๆ เช่น ฝรั่งเศส และสหรัฐ ได้มีส่วนส่งเสริมวิชาการด้านนี้ให้กว้างขวางยิ่งขึ้น เช่น Alfred Binet และ Lewis Terman แต่การศึกษาค้นคว้าทั้งหมดนี้ยังเป็นไปเพื่อประโยชน์ในแง่จิตวิทยาโดยตรงเป็นประการสำคัญ ต่อมาเมื่อเกิดสงครามโลกขึ้น จึงได้มีการพยายามนำเอาความรู้ในวิชานี้มาประยุกต์กับการบริหารงานบุคคล ในระยะสงครามโลกครั้งที่หนึ่ง ผลงานของ Binet และ Cattell ได้นำมาทดลองใช้ในการคัดเลือกแยกประเภทบุคคลที่เข้ารับราชการทหาร เพื่อจัดงานให้ทหารแต่ละคนปฏิบัติในตำแหน่งที่เหมาะสม เครื่องมือวัดที่เป็นแม่บทในการดำเนินงานในครั้งนั้นก็คือ Binet Intelligence Test ซึ่งสร้างขึ้นในปี ค.ศ. 1905 นับเป็นแบบทดสอบแบบแรกที่ใช้วัดเชาวน์ Terman ได้นำแบบทดสอบนี้มาแปลและดัดแปลงเป็นภาษาอังกฤษ และให้ชื่อใหม่ว่า Stanford-Binet ในปี ค.ศ. 1916 แบบทดสอบใหม่ที่ดัดแปลงนี้มีชื่อเสียงมาก และเป็นที่เชื่อถือกันมาเป็นระยะเวลานานจนถึงกับมีการปรับปรุงใหม่อีก 2 ครั้ง คือ ในปี ค.ศ. 1937 และ 1960

¹² Ghiselli, E.E., *Theory of Psychological Measurement*, pp. 37—66.

สงครามโลกครั้งที่สอง มีบทบาทสำคัญที่ทำให้การวัดผลบุคคลเจริญก้าวหน้ามากขึ้น ระยะเวลา ค.ศ. 1915—1930 นับว่าเป็น “ยุคทอง” ของการวัดผลบุคคล ในยุคนี้ได้มีเครื่องมือวัดผลบุคคลเกิดขึ้นมากมาย การวิจัยในวงการทหารของสหรัฐนับได้ว่าเป็นส่วนที่สำคัญที่สุดในการขยายความรู้ทางวิชาการวัดผลบุคคล เครื่องมือวัดบุคคลที่เด่นในยุคนี้เป็นของกองทัพบกสหรัฐ เช่น Army Alpha และ Army Beta ซึ่งสร้างขึ้นในปี ค.ศ. 1918 นอกจากนี้ สถาบันอุดมศึกษาอเมริกันก็ได้มีส่วนในการค้นคว้าวิชาการเรื่องนี้ด้วย เช่น มหาวิทยาลัยมินนิโซตา และมหาวิทยาลัยสแตนฟอร์ด เป็นต้น ในปัจจุบัน ผลการค้นคว้าในเรื่องการวัดผลบุคคล ทำให้มีเครื่องมือวัดผลบุคคลในด้านต่าง ๆ ทุกด้าน ไม่ว่าจะเป็นในเรื่องเชาวน์ ความถนัดพิเศษทางธรรมชาติ ความสัมพันธ์ผล หรือลักษณะต่าง ๆ ของบุคลิกภาพ แบบทดสอบที่สร้างขึ้นและมีในท้องตลาดนับได้เป็นพันชุด¹³ วิสาหกิจ และสถาบันอุดมศึกษาหลายแห่งนับได้ว่าเป็นแหล่งที่จะค้นคว้า หรือจัดหาเครื่องมือวัดบุคคลแบบต่าง ๆ เช่น Educational Testing Service, California Test Bureau, Science Research Associates, มหาวิทยาลัยโคลัมเบีย และมหาวิทยาลัยไอโอวา เป็นต้น

ความแตกต่างระหว่างบุคคล และลักษณะทางจิตวิทยาที่วัดได้

สิ่งต่าง ๆ และปรากฏการณ์ต่าง ๆ ของโลกย่อมมีความต่างกัน ความแตกต่างกันระหว่างสิ่งต่าง ๆ นี้ อาจจะเป็นในแง่ คุณภาพ หรือชนิด หรือปริมาณ หรือความบ่อย หรือระดับ ในแง่ใดแง่หนึ่ง โดยทั่ว ๆ ไปก็มักจะกล่าวถึงในแง่ปริมาณและคุณภาพ เมื่อกล่าวถึงในแง่คุณภาพก็ต้องหมายถึงว่าต่างกันเชิงประเภท และเมื่อกล่าวถึงปริมาณก็จะหมายความว่าต่างกันในเรื่องจำนวนซึ่งวัดได้ การวัดผลเป็นการที่จะต้องใช้ค่าตัวเลข เพื่อชี้ให้เห็นความแตกต่างระหว่างสิ่งต่าง ๆ ในแง่ปริมาณ

ในหน่วยงานต่าง ๆ บุคคลที่ปฏิบัติงาน ต่างก็ไม่เหมือนกัน ถ้ากล่าวอย่างหยาบ ๆ โดยใช้สามัญสำนึก ก็อาจกล่าวได้ว่า บุคคลต่าง ๆ ที่ปฏิบัติงานในหน่วยงานหนึ่ง ๆ ต่างกันในเรื่องผลการปฏิบัติงานและพฤติกรรมที่แสดงออก แม้จะจัดให้บุคคลต่าง ๆ ปฏิบัติงานโดยใช้เครื่องจักรชนิดเดียวกัน ขนาดเดียวกัน ผลงานของบุคคลเหล่านี้ก็ยิ่งอาจต่างกันอีก และพฤติกรรมอื่น ๆ ของบุคคลเหล่านั้น เช่น ความเอาใจใส่ ความพยายาม ก็ยิ่งอาจต่างกันออกไป

¹³ Berdie, et. al., *op. cit.*, pp. 69—70.

ในทางจิตวิทยาถือว่า การที่บุคคลที่ปฏิบัติงานในหน่วยงานต่างกันนั้น เป็นเพราะแต่ละคนมีลักษณะและคุณสมบัติทางจิตวิทยาไม่เหมือนกัน คุณสมบัติทางจิตวิทยาของแต่ละคนเป็นผลทำให้บุคคลหนึ่ง ๆ แสดงความฉลาด หรือความสามารถเฉพาะตัว มีความรู้วิชาการตลอดจนมีบุคลิกภาพที่ไม่เหมือนบุคคลอื่น ถ้ากล่าวในแง่การวิจัยก็อธิบายได้ว่า ตัวแปรต่าง ๆ มากมายหลายตัว มีส่วนทำให้บุคคลแสดงพฤติกรรมต่าง ๆ กัน และเป็นหน้าที่ของผู้วัดผลบุคคลที่จะต้องค้นคว้าหาตัวแปรแต่ละตัวเหล่านี้ ตัวแปรเหล่านี้ก็คือ คุณสมบัติของพฤติกรรม (traits) ต่าง ๆ นั้นเอง จากผลการวิจัยจึงทำให้มีการตั้งชื่อคุณสมบัติของพฤติกรรมเหล่านั้นขึ้น สมควรกล่าวไว้ในที่นี้ด้วยว่าในการวิจัยทางจิตวิทยานั้น มนุษย์ไม่ได้พยายามที่จะอธิบายลักษณะและคุณสมบัติของบุคคลตามชื่อของคุณสมบัติแห่งพฤติกรรมที่ตั้งไว้ก่อน แต่นักวิชาการได้พยายามที่จะอธิบายลักษณะและคุณสมบัติต่าง ๆ แต่ละอย่างของบุคคล แล้วจึงตั้งชื่อลักษณะนั้น ๆ ให้เป็นคุณสมบัติของพฤติกรรมหนึ่ง ๆ หน้าที่ของนักวิจัยและนักวิชาการในงานวัดผลในการที่จะวิเคราะห์ให้ได้ว่าบุคคลแตกต่างกันอย่างไรก็คือ จะต้องพยายามนิยามและชี้ให้เห็นความแตกต่างของตัวแปรต่าง ๆ เพื่อจะค้นหาวิธีการและเครื่องมือที่จะวัดค่าความแตกต่างนั้นออกมาเป็นตัวเลขให้ได้

ในการบริหารงานบุคคลโดยทั่วไปนั้น ก็พยายามที่จะบรรจุ แต่งตั้งหรือมอบหมายให้บุคคลที่มีความสามารถสูงที่สุดปฏิบัติงานในหน้าที่ต่าง ๆ แต่ส่วนใหญ่มักจะพิจารณาความสามารถนั้นไปในแง่ความรู้ และวิชาการที่จะต้องใช้ในการปฏิบัติงานเป็นประจำสำคัญ ส่วนตามหลักจิตมิตินั้น ไม่ถือแต่ว่า บุคคลต่างกันในเรื่องความรู้และวิชาการที่มีและไม่ถือแต่เพียงว่า ความรู้และวิชาการที่มีนั้นจะช่วยให้บุคคลปฏิบัติงานได้สำเร็จ ตำราวิชาการทางจิตมิตีและผลการวิจัยทางด้านนี้ ได้ชี้ให้เห็นว่า บุคคลแต่ละคนมีคุณสมบัติของพฤติกรรมต่างกัน และคุณสมบัติของพฤติกรรมที่ต่างกันนี้เอง จะเป็นเครื่องประกอบในการพิจารณาว่า บุคคลจะปฏิบัติงานได้ดีและสำเร็จผลหรือไม่ กล่าวอีกนัยหนึ่งก็คือ วิชาการวัดผลแสดงว่าหากประสงค์จะทราบว่าจะปฏิบัติงานได้ดี และสำเร็จผลหรือไม่ จะต้องพิจารณาทั้งความรู้และวิชาการที่จะใช้ในการปฏิบัติงานร่วมกับคุณสมบัติของพฤติกรรมที่บุคคลแต่ละคนมี

ในการวัดผลและประเมินบุคคลแต่ละคนในแง่จิตวิทยานั้น มีเป็น 2 นัย คือ นัยแรกวัดว่าบุคคลแต่ละคน สามารถ ทำอะไร ได้บ้าง และอีกนัยหนึ่งวัดว่าบุคคลแต่ละคน จะทำ

อะไรได้บ้าง¹⁴ การวัดตามนัยแรกเรียกตามสำนวนทางจิตวิทยาว่า เป็นการวัดความสามารถ ในการวัดความสามารถนี้อาจจะวัดความถนัดทางธรรมชาติ หรือความสัมฤทธิ์ผลก็ได้ การวัดตามนัยที่สองเรียกว่า เป็นการวัดบุคลิกภาพ

การวัดความถนัดทางธรรมชาติ เป็นการวัดสิ่งที่บุคคลแต่ละคนสามารถจะเรียนรู้เพื่อนำไปปฏิบัติได้ แต่การวัดความสัมฤทธิ์ผลนั้นเป็นการวัดสิ่งที่ได้เรียนรู้ไปแล้ว กล่าวอีกอย่างหนึ่งก็คือ การวัดความถนัดทางธรรมชาติ เป็นการวัดเพื่อทำนายว่าบุคคลแต่ละคนสามารถจะเรียนรู้หรือพัฒนาตนเองด้วยอะไรบ้าง¹⁵ ส่วนการวัดความสัมฤทธิ์ผลเป็นการวัดเพื่อให้ทราบว่าบุคคลแต่ละคนได้เรียนรู้อะไรไปแล้วในอดีต ในการวัดบุคลิกภาพนั้นเป็นการวัดเพื่อที่จะให้ทราบว่าบุคคลแต่ละคนจะทำอะไร ในการวัดบุคลิกภาพนี้จะมีการวัดความสนใจ ความต้องการพิเศษ และความขัดแย้งภายในตนเองของแต่ละคน เพื่อให้สามารถคาดคะเนได้ว่า บุคคลหนึ่ง ๆ ที่ได้รับมอบหมายให้ปฏิบัติงานอย่างหนึ่ง ซึ่งอาจจะเป็นงานสารบรรณ หรืองานต้อนรับก็ดี จะประพฤติตนในลักษณะการที่เป็นเครื่องส่งเสริมให้ผลการปฏิบัติงานนั้น ๆ มีคุณค่าอยู่ในระดับสูง¹⁶ การวัดบุคลิกภาพนี้ ไม่ใช่เป็นการวัดว่าบุคคลหนึ่ง ๆ ที่เป็นเจ้าหน้าที่สารบรรณหรือพนักงานต้อนรับสามารถปฏิบัติงานร่วมกับคนอื่นได้ดี หรือสามารถแสดงความกระตือรือร้นในการปฏิบัติงานได้ดีเป็นอย่างไร แต่การวัดบุคลิกภาพนี้ต้องการที่จะทราบว่า โดยทั่วไป บุคคลนั้น ๆ มีความกระตือรือร้น หรือความสามารถปฏิบัติงานร่วมกับผู้อื่นในฐานะเจ้าหน้าที่สารบรรณหรือพนักงานต้อนรับ ในการวัดบุคลิกภาพนี้อาจสรุปได้ว่าเป็นการวัดลักษณะของบุคคลแต่ละคน 5 ประการใหญ่ ๆ คือ ลักษณะนิสัย และอุปนิสัย ความสามารถในการปรับตัว อารมณ์ (ร่าเริง กระตือรือร้น เศร้า) ความสนใจและทัศนคติ

ในการวัดลักษณะทางจิตวิทยาของบุคคลแต่ละคน ผู้วัดและประเมินบุคคลอาจพิจารณาลักษณะทางจิตวิทยาของแต่ละคนในแง่การวัดผลแทนที่จะพิจารณาในแง่ของคุณสมบัติทางจิตวิทยาก็ได้ กล่าวคือ ลักษณะทางจิตวิทยาบางอย่างต้องวัดเพื่อให้ทราบค่าปริมาณที่เต็มที่ และบางอย่างต้องวัดให้ทราบค่าที่เป็นตัวแทน

¹⁴ Thorndike, R.L. and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*, pp. 21—22.

¹⁵ Guion, *op. cit.*, pp. 239—276.

¹⁶ Rethlingshafer, Dorothy, *Motivation as Related to Personality*, pp. 254—278. Guilford, J.P., *Personality*, pp. 33—37.

ค่าปริมาณที่เพิ่มที่ที่อาจวัดได้ก็คือ ความสามารถของแต่ละคน ความสามารถนี้แยกได้เป็น 2 อย่าง คือ ความสามารถปกติ (ability) และความสามารถสูงสุด (capacity)¹⁷ ความสามารถปกติ คือ ความสามารถของบุคคลแต่ละคนที่จะปฏิบัติงานหนึ่ง ๆ ในระหว่างที่มีความพยายามในการปฏิบัติงานนั้น ๆ อย่างเต็มที่ ส่วนความสามารถสูงสุดของแต่ละบุคคล หมายถึงความสามารถปกติที่ได้รับการเสริมด้วยการฝึกอบรม ความสามารถปกติย่อมไม่มีทางสูงกว่าความสามารถสูงสุด แต่ความสามารถปกตินี้ จะมีระดับสูงขึ้นจนใกล้เคียงความสามารถสูงสุดหรือไม่ขึ้น ก็ขึ้นอยู่กับว่าบุคคลแต่ละคนนั้น ๆ ได้พัฒนาศักยภาพของตนเองขึ้นอย่างไร ในบางครั้งความสามารถปกติก็ลดหย่อนลง เพราะมีข้อขัดขวางทางอารมณ์ เช่น นักแสดงอาจตื้นก้นเมื่อขึ้นบนเวที ความสามารถที่ต้องการวัดค่าปริมาณเพิ่มที่นี้ อาจจะเป็นความสามารถในการใช้ศัพท์คำนวณ และถ้อยคำ ความสามารถในการเข้าใจงานเข้าใจลักษณะทางวิชาการของงาน ความสามารถที่จะปฏิบัติตนให้สอดคล้องกับภาวะการฉุกเฉิน (ความสามารถนี้ใช้ทดสอบกับพนักงานขับรถ และเจ้าหน้าที่ที่ใช้เครื่องจักร) ในการวัดความสามารถเพื่อให้ทราบค่าปริมาณเพิ่มที่นี้ ทำให้ผู้วัดทราบว่า บุคคลแต่ละคนที่ปฏิบัติงานอย่างหนึ่ง ๆ อย่างดีที่สุดนั้น จะปฏิบัติได้แค่ใดเพียงใด และควรจะได้รับฝึกอบรมต่อไปหรือไม่เพียงใด

ค่าตัวแทนที่ต้องการวัด ไม่ใช่เป็นค่าเกี่ยวกับคนแต่ละคนว่าดีที่สุดหรือสามารถมากที่สุดเป็นอย่างไร ในการวัดค่าตัวแทนทางจิตวิทยา ผู้วัดต้องการทราบว่าโดยทั่ว ๆ ไป บุคคลหนึ่ง ๆ มีความมั่นคงทางจิตใจอย่างไร มีใจกว้างขวางอย่างไร มีความร่าเริงอย่างไร ฯลฯ การที่ผู้วัดเช่นนี้ก็เพราะการวัดค่าปริมาณที่เพิ่มที่นั้น ยังไม่ได้แสดงว่า บุคคลแต่ละคนสามารถปฏิบัติงานได้ดีจริง ๆ ในกรณีพนักงานต้อนรับ ผู้วัดค่าปริมาณเพิ่มที่ที่อาจวัดได้ว่า ผู้ที่จะปฏิบัติงานในหน้าที่พนักงานผู้นี้จะสามารถแสดงความสุขภาพได้ดีที่สุดอย่างไร แต่การวัดความสามารถแสดงความสุขภาพนี้ ทำให้ทราบได้แต่เพียงว่า บุคคลผู้นี้เมื่อประสงค์จะแสดงความสุขภาพ จะสามารถสุขภาพได้มากเพียงใด ตามธรรมดา บุคคลทั่ว ๆ ไปที่ได้ผ่านการศึกษามาตามสมควรก็มีความสามารถแสดงความสุขภาพได้ แต่การที่พนักงานต้อนรับจะปฏิบัติงานด้วยความสุขภาพให้เป็นผลดีแก่หน้าที่นั้น หมายความว่า จะต้องปฏิบัติงานด้วยความสุขภาพตลอดเวลา ตลอดเวลา ไม่ใช่เฉพาะเวลาประสงค์จะสุขภาพเท่านั้น ถ้าจะลองพิจารณาอีกตัวอย่างหนึ่งอาจเห็นได้ชัดขึ้น หัวหน้านางานที่ติดตามธรรมดา จะต้องสามารถรู้ข้อบกพร่องและจุดอ่อนของผู้อยู่ใต้บังคับบัญชาในการวัดค่าบริหารงานอย่างไร แต่ความสามารถที่วัดได้นั้น เป็นความสามารถ

¹⁷ Cronbach, L.J., *Essentials of Psychological Testing*, pp. 13—14.

ที่บุคคลผู้นั้นแสดงในระยะเวลาที่ประสงค์จะแสดงความสามารถ ผู้วัดไม่ทราบว่าจะโดยทั่วๆ ไป เมื่อบุคคลผู้นั้นปฏิบัติงานในฐานะหัวหน้างานในวันหนึ่งๆ เวลาต่างๆ กัน จะมีความสามารถเช่นนั้นตลอดเวลาหรือเปล่า เพราะความสามารถที่แสดงนั้นเป็นความสามารถสูงสุด ฉะนั้นการวัดที่ทำให้ทราบได้ว่า บุคคลผู้นั้นมีความสามารถสูงสุด อยู่ในระดับดีพอที่จะเป็นหัวหน้างานได้ ไม่ได้ทำให้ทราบว่าบุคคลผู้นั้นเป็นหัวหน้างานที่ดีได้

กล่าวโดยสรุป ลักษณะทางจิตวิทยาของบุคคลแต่ละคนที่อาจวัดผล และประเมินได้ มีดังนี้¹⁸

1. ความสามารถ หมายถึง หลักฐานที่แสดงว่าบุคคลแต่ละคนสามารถทำอะไรได้บ้าง ถ้าบุคคลผู้นั้นพยายามทำ แยกเป็น 2 ประเภทใหญ่ๆ คือ

ก. ความถนัดทางธรรมชาติ หมายถึง การปฏิบัติงานซึ่งถือเป็นดัชนีวัดว่าบุคคลผู้นั้นสามารถเรียนรู้อะไรได้บ้าง เพื่อไปปฏิบัติงานในระยะต่อไป

ข. ความสำเร็จผล หมายถึง การปฏิบัติงานซึ่งชี้ให้เห็นว่าบุคคลผู้นั้นได้เรียนรู้อะไรไปแล้วบ้างในการปฏิบัติงาน

2. ตัวแปรทางบุคลิกภาพ หมายถึง สิ่งซึ่งว่าบุคคลหนึ่งๆ จะทำอะไรและมีปฏิกริยาสนองตอบต่อสภาพการณ์ และลักษณะต่างๆ ในการครองชีวิตอย่างไร แยกเป็น 5 ประเภทใหญ่ๆ คือ

ก. ลักษณะนิสัยและอุปนิสัย หมายถึง คุณสมบัติบางประการที่สังคมยกย่องว่าเป็นส่วนดีหรือว่าเป็นส่วนเสีย

ข. ความสามารถในการปรับตัว หมายถึง ระดับความสามารถของแต่ละคนที่ จะปฏิบัติตนและชีวิตอย่างมีความสุข โดยสอดคล้องกับวัฒนธรรมและสังคม

ค. อารมณ์ หมายถึง คุณสมบัติที่เกี่ยวกับ กำลังใจ ความรู้สึกโต้ตอบ และแนวการครองชีวิต

ง. ความสนใจ หมายถึง กิจกรรมต่างๆ ที่บุคคลแต่ละคนประสงค์จะทำหรืออยากจะทำ หรือไม่อยากจะทำ

จ. ทักษะคติ หมายถึง ปฏิกริยาของบุคคลแต่ละคนที่มีต่อ หรือสนองตอบบุคคลอื่น ๆ หรือปรากฏการณ์ต่างๆ หรือแนวคิดต่างๆ ในสังคม

¹⁸ Thorndike, *op. cit.*, pp. 22—25.

วิธีการต่าง ๆ ในการวัดผลบุคคล

การวัดผลบุคคลทำได้ 2 วิธี คือ วิธีทดสอบและวิธีสังเกต การวัดผลด้วยการทดสอบคือการจัดสภาพการณ์พิเศษเพื่อการวัดผล ส่วนการวัดผลด้วยการสังเกตนั้นเป็นการวัดผลโดยสังเกตพฤติกรรมของบุคคลในสภาพการณ์ที่เป็นจริงของชีวิต¹⁹

ลักษณะพิเศษในการวัดผลด้วยการทดสอบมี 3 ประการคือ ประการแรก เป็นการวัดในสถานที่และเวลาที่กำหนดไว้ ประการที่สอง เป็นการวัดด้วยการจัดงาน หรือกลุ่มของงานให้บุคคลแต่ละคนทำเหมือนกัน ประการสุดท้าย เป็นการวัดที่บุคคลซึ่งถูกวัดรู้สึกตัวว่ากำลังถูกวัด

ส่วนวัดผลด้วยการสังเกตนั้น มีลักษณะตรงข้ามกันกับการวัดด้วยการทดสอบ กล่าวคือ ประการแรก เป็นการวัดในระยะเวลาไม่จำกัด ประการที่สอง สภาพการณ์ที่เกี่ยวข้องกับบุคคลแต่ละคนที่ถูกวัดมีลักษณะต่าง ๆ กันออกไป ประการสุดท้าย บุคคลที่ถูกวัดไม่รู้ตัวว่าตนเองกำลังถูกวัด

อย่างไรก็ดีการวัดผลบุคคลที่สมบูรณ์นั้น ย่อมต้องใช้วิธีการทั้ง 2 อย่างประกอบกัน ผู้วัดไม่ควรจะคำนึงแต่เพียงว่า ถ้าหากใช้แบบทดสอบมาตรฐานในเรื่องใดเรื่องหนึ่งแล้ว (เช่นแบบทดสอบความสามารถในการใช้ภาษา) ก็จะได้ข้อเท็จจริงในเรื่องนั้นอย่างสมบูรณ์ เพราะแบบทดสอบไม่ทำให้ผู้วัดสามารถทราบข้อเท็จจริงเกี่ยวกับบุคคลแต่ละคนได้ทั้งหมด ความสามารถต่างๆ ตลอดจนบุคลิกภาพในด้านต่างๆ ซึ่งวัดได้ในแบบทดสอบนั้น เป็นเพียงส่วนหนึ่งของความสามารถ และบุคลิกภาพทั้งหมดของบุคคลหนึ่ง ๆ จริงอยู่ในการวัดด้วยแบบทดสอบทำให้ผู้วัดได้ทราบผลการวัดที่แน่นอนจากสภาพการณ์ที่ได้เตรียมไว้ และได้มีการเปรียบเทียบความสามารถและบุคลิกภาพของแต่ละบุคคลได้อย่างแน่ชัดขึ้น

การทดสอบมี 2 แบบ การทดสอบแบบแรก คือ การทดสอบที่บังคับให้ผู้ถูกวัดต้องบันทึกหลักฐานเกี่ยวกับพฤติกรรมของตน การทดสอบอีกแบบหนึ่งก็คือการวัดโดยผู้วัดเป็นผู้สังเกตผู้ถูกวัด และผู้ถูกวัดไม่จำเป็นต้องมีบันทึกอะไรเกี่ยวกับตนเอง²⁰ ในการวัดแบบแรกก็คือการทำให้ผู้ถูกวัดเขียนคำตอบ ซึ่งอาจจะเป็นเรื่องเกี่ยวกับภาษา ความรู้ในเรื่องเครื่องจักร ความรู้เกี่ยวกับการพิมพ์ดีด ฯลฯ เมื่อผู้วัดได้เห็นคำตอบแล้วก็พิจารณาให้คะแนนเก็บไว้เป็นหลักฐานเกี่ยวกับผู้ถูกวัด ส่วนการวัดแบบที่สองนั้น

¹⁹ Kerlinger, F.N., *Foundations of Behavioral Research*, pp. 481—499, 503—599.

²⁰ Thorndike, *op. cit.*, pp. 17—18.

อาจทำได้โดยผู้ถูกวัดอ่านหนังสือหรือพูดให้ฟัง และผู้ถูกวัดก็ให้คะแนนไปตามดุลพินิจของตน

ลักษณะต่าง ๆ ของการวัดด้วยการสังเกต²¹ ถ้าหากแยกประเภทการสังเกตตามลักษณะของผู้ที่ทำการสังเกต ก็อาจแยกวิธีการวัดด้วยการสังเกตออกได้เป็น 2 วิธี คือ การสังเกตโดยบุคคลอื่น และการสังเกตโดยตนเอง คนอื่นที่อาจจะสังเกตเพื่อวัตถุประสงค์ที่มีหลายประเภท เช่น นายจ้าง เพื่อนร่วมงาน บุคคลในครอบครัวเดียวกัน เป็นต้น การสังเกตโดยบุคคลอื่นนั้นย่อมเป็นการวัดแต่เพียงด้านเดียว คือ วัดในแง่ดุลพินิจของผู้สังเกต ซึ่งอาจมีอคติได้ เช่น ผู้บังคับบัญชาอาจจะเห็นแต่อุปนิสัยใจคอของเจ้าหน้าที่ที่ถูกวัดแต่เพียงเมื่อผู้ถูกวัดอยู่ในสำนักงานเท่านั้น ผู้บังคับบัญชาอาจจะมองแต่เพียงว่าผู้ถูกสังเกตนั้นเป็นคนนอบน้อมหรือไม่ ร่วมมือกับเพื่อนร่วมงานหรือไม่ ลักษณะที่สังเกตเหล่านี้ อาจจะต่างไปจากคุณสมบัติที่จริง ๆ ของผู้ถูกวัดนั้น ๆ กล่าวคือ ถ้าหากสภาพการณ์การปฏิบัติงานเปลี่ยนไป คุณสมบัติที่ผู้บังคับบัญชาสังเกตไว้ก็อาจจะเปลี่ยนไป ส่วนการวัดโดยการสังเกตด้วยตนเองนั้น ก็นับได้ว่ามีข้อจำกัดหลายอย่าง เช่น เจ้าตัวอาจจะไม่เข้าใจตัวเอง หรืออาจจะไม่กล้าที่จะวิเคราะห์ตนเอง จุดอ่อนของคนบางคน เช่น ความเห็นแก่ตัว เรื่องบางเรื่องที่รู้สึกสะเทือนใจง่าย เป็นเรื่องที่บุคคลแต่ละคนไม่ค่อยกล้าจะยอมรับกับตนเองนัก

ถ้าหากแยกประเภทการสังเกตตามลักษณะของเวลาที่สังเกต ก็อาจแยกวัดด้วยวิธีการสังเกตได้เป็น 2 วิธีเช่นเดียวกัน วิธีแรก คือสังเกตโดยการวางแผนการว่าจะสังเกตอะไรบ้าง เช่น สมมติว่าจะสังเกตบุคคลหนึ่ง ๆ ว่ามีแนวโน้มที่จะรู้สึกโกรธในเรื่องใดบ้าง ผู้วัดซึ่งจะเป็นตนเองหรือบุคคลอื่นก็ตาม อาจจะกำหนดเวลาไว้หนึ่งสัปดาห์ หรือหนึ่งเดือน เพื่อยกยบันที่กว่าบุคคลผู้ถูกวัดจะรู้สึกโกรธในเรื่องอะไรบ้าง ในเวลาใด อีกวิธีหนึ่งคือวิธีรำลึกย้อนหลัง ผู้ถูกวัดจะต้องพยายามนึกย้อนหลังไปว่าเคยโกรธกับใครบ้างด้วยเรื่องอะไร ในเวลาใด วิธีการทั้งสองอย่างนี้ ต่างก็มีผลดีและผลเสีย วิธีการแรกคือการสังเกตโดยการวางแผนการ เป็นวิธีการที่ต้องใช้เวลา แต่จะได้ข้อเท็จจริงที่ตรงความเป็นจริงมากกว่าวิธีที่ 2 ส่วนวิธีการที่ 2 เปลืองเวลาน้อยกว่า แต่ข้อเท็จจริงที่ได้มาก็อาจจะมีอคติ

อย่างไรก็ดี การวัดด้วยการสังเกตไม่ว่าด้วยวิธีใดก็ตามย่อมไม่สามารถให้ข้อเท็จจริงเกี่ยวกับตัวบุคคลได้ถูกต้องทั้งหมด ได้กล่าวมาในตอนต้นว่าในการวัดบุคคล ควรจะ

²¹ Kerlinger, *op. cit.*, pp. 592—596.
Thorndike, *op. cit.*, pp. 18—21.

พยายามใช้ทั้งวิธีการทดสอบ และวิธีการสังเกตร่วมกัน แต่เนื่องจากผลของการวัดด้วยการสังเกตยังไม่เป็นที่น่าพอใจ จึงได้มีผู้คิดวิธีการใหม่เรียกว่าวิธีการทดสอบแบบจัดสภาพการณ์ (situational test) ในการวัดด้วยวิธีนี้ ผู้วัดจะต้องพยายามสร้างสภาพการณ์ที่เจาะจงอย่างหนึ่ง แล้วพยายามที่จะวัดผลของพฤติกรรมที่เกี่ยวกับความสามารถหรือบุคลิกภาพของบุคคลในสภาพการณ์เดียวกันนั้น เช่น สมมติว่าต้องการวัดบุคคลคนหนึ่งว่าอารมณ์ฉุนเฉียวมาก หรือไม่ อย่างไร ในการบังคับบัญชางาน ผู้วัดจะต้องสร้างสภาพการณ์ที่กำหนดไว้ว่าจะก่อให้เกิดอารมณ์ฉุนเฉียวเป็นมาตรฐานไว้ การสร้างสภาพการณ์นี้อาจทำได้หลายอย่าง ผู้วัดอาจจะมอบหมายงานชิ้นหนึ่งให้บุคคลที่จะถูกวัดปฏิบัติ พร้อมกับกำหนดผู้ช่วยที่โง่เขลาพูดจาไม่รู้เรื่อง ให้ปฏิบัติร่วมด้วย บุคคลที่ถูกวัดซึ่งถูกจัดให้อยู่ในสภาพการณ์ที่เหมือน ๆ กันเช่นนี้จะแสดงความโมโหมั่งที่จะเกิดความฉุนเฉียวในระดับต่าง ๆ กัน ทำให้ผู้วัดสามารถวัดบุคลิกภาพในด้านนี้ได้ อย่างไรก็ตาม การวัดด้วยวิธีนี้ทำได้ยาก เพราะผู้วัดก็ประสบความลำบากในการจัดสภาพการณ์ที่ใกล้เคียงต่อชีวิตจริงเพื่อเป็นมาตรฐานในการวัด วิธีการวัดด้วยการทดสอบแบบสังเกตจึงไม่แพร่หลาย

แบบทดสอบสองประเภท ในการวัดผลและประเมินบุคคลด้วยการสอบแบบบันทึกหลักฐาน เกี่ยวกับพฤติกรรมของผู้ทดสอบนั้น อาจจะใช้แบบทดสอบ 2 ประเภท คือแบบทดสอบอัตนัย และแบบทดสอบปรนัย แบบทดสอบอัตนัยคือข้อทดสอบที่มีลักษณะสำคัญ 4 ประการ ได้แก่ ประการแรก ผู้ถูกวัดเขียนและเรียบเรียงคำตอบตามความคิดเห็นของตน โดยมีข้อจำกัดในเรื่องการตอบน้อยที่สุด ประการที่สอง ใช้ถ้อยคำสำนวนของตนเอง ประการที่สาม ตอบคำถามที่จำนวนน้อย และประการที่สี่ คำตอบของผู้ถูกวัดมีความสมบูรณ์ในเนื้อหา และมีความถูกต้องในระดับต่าง ๆ กัน ส่วนแบบทดสอบปรนัยนั้น เป็นแบบทดสอบที่ผู้ถูกวัดมักจะต้องให้คำตอบที่ถูก ซึ่งมีที่จะเป็นคำตอบเดียว ลักษณะสำคัญของแบบทดสอบปรนัยมี 4 ประการ คือ ประการแรกคำถามในแบบทดสอบได้จัดเรียงไว้ โดยไม่ให้โอกาสผู้ถูกวัดมาเรียบเรียงความคิดเห็นของตน ประการที่สอง ผู้ถูกวัดจะต้องเลือกคำตอบ ซึ่งได้กำหนดไว้ ประการที่สามคำถามมีเป็นจำนวนมากเมื่อเทียบกับแบบแรก และประการที่สี่ ผู้วัดได้กำหนดคำตอบไว้แน่นอนแล้ว²² อาจสรุปได้ว่าแบบทดสอบปรนัยมีลักษณะเด่นที่ต่างจากแบบอัตนัยก็คือ ความเป็นอันหนึ่งอันเดียวกันของคำตอบ ซึ่งหมายความว่าผู้ถูกวัดคนหนึ่งจะต้องได้คะแนนเท่ากันอยู่เสมอไม่ว่าผู้ให้คะแนนจะเป็นใครก็ตาม

²² Lyman, H.B., *Test Scores and What They Mean*, pp. 18—22.

แบบทดสอบอัตนัย และแบบปรนัยต่างก็มีทั้งข้อดีและเสีย ในที่นี้จะไม่กล่าวถึงสาระสำคัญในเรื่องนี้ เพราะเป็นรายละเอียดเกี่ยวกับทดสอบทั่วไป แต่ก็สรุปย่อได้ว่าในการพิจารณาประโยชน์ของแบบทดสอบปรนัย ผู้วัดไม่ควรจะมองข้ามประโยชน์ของแบบทดสอบแบบอัตนัยไปเสียทีเดียว

เนื่องจากแบบทดสอบแบบมาตรฐานที่ใช้วัดบุคคล ส่วนใหญ่สร้างเป็นแบบทดสอบแบบปรนัย ฉะนั้นการที่จะเข้าใจว่าแบบทดสอบมาตรฐานเหล่านั้นดีหรือไม่ก็ขึ้นอยู่กับความเข้าใจในลักษณะที่ดีของแบบทดสอบแบบปรนัย

ลักษณะที่ดีของแบบทดสอบแบบปรนัยอาจสรุปได้ว่า มี 8 ประการ คือ²³

1. จะต้องไม่ใช่ภาษาที่เข้าใจยาก ยกเว้นแต่ในแบบที่ทดสอบที่จะวัดความสามารถในการอ่าน หรือความสามารถในการใช้ภาษา
2. จะต้องไม่ลอกข้อความจากตำรามาตั้งเป็นคำถามคำตอบ เพราะจะไม่ทำให้สามารถทดสอบความสามารถในการนำความรู้ไปใช้ได้
3. เมื่อมีคำถามที่เกี่ยวข้องกับความเห็นของผู้ทรงคุณวุฒิ ควรระบุแหล่งที่มาด้วย เพื่อจะได้ไม่เป็นการถามความจำ
4. คำถามแต่ละข้อไม่ควรมีเนื้อหาที่จะทำให้สามารถช่วยตอบซึ่งกันและกัน
5. คำถามแต่ละข้อควรเป็นอิสระต่อกัน ไม่ควรจะเป็นคำถามที่จะต้องตอบคำถามอื่น ๆ ก่อนจึงจะตอบคำถามนี้ได้
6. การเรียงคำถามควรจะเรียงด้วยวิธีสุ่ม คือ ไม่เรียงคำถามให้มีคำตอบที่สามารถกำหนดแบบที่จะตอบได้
7. ไม่ควรใช้คำถามที่ล่อให้ผู้ถูกวัดตกหลุมพราง เพราะคำถามทุกข้อมีไว้เพื่อจะทดสอบระดับความรู้ และความเข้าใจโดยทั่วไป
8. ไม่ใช่ข้อความที่มีความหมายหลายแง่

ในการสร้างแบบทดสอบปรนัย อาจสร้างได้ 4 แบบ คือ

1. แบบถูกผิด เป็นแบบที่กำหนดคำถามให้ผู้ถูกวัด ภาเครื่องหมายว่าถูกหรือผิด
2. แบบเติมคำหรือข้อความ เป็นแบบที่กำหนดช่องว่างไว้ให้ผู้ถูกวัด เติมคำหรือข้อความ
3. แบบเลือก เป็นแบบที่กำหนดข้อความ หรือถ้อยคำหลาย ๆ อันไว้ให้ผู้ถูกวัด เลือกเฉพาะข้อความหรือคำที่ถูกต้อง

²³ Travers, *op. cit.*, pp. 181—183.

4. แบบจับคู่ เป็นแบบที่แปลงมาจากแบบเลือก โดยกำหนดให้ผู้ถูกวัดเลือกข้อความ หรือคำที่มีความหมายสัมพันธ์กัน หรือเหมือนกันมาจับคู่เข้าด้วยกัน

ขั้นตอนในการสร้างแบบทดสอบมาตรฐาน แบบทดสอบมาตรฐานโดยทั่วไปเป็นแบบทดสอบปรนัยที่สร้างขึ้นโดยกำหนดเนื้อหา และวิธีการทดสอบแน่นอนตายตัว เพื่อให้สามารถวัดประเด็นหนึ่งประเด็นใดโดยเฉพาะ ให้ได้ผลของการวัดออกมาเท่ากันเสมอ ไม่ว่าจะวัดในที่ไหน ใด เวลาใด แบบทดสอบมาตรฐานนี้ตามปกติจะกำหนดเกณฑ์การวัด (norms) แบบทดสอบปรนัยที่สร้างกันโดยทั่วไปยังไม่นับเป็นแบบทดสอบมาตรฐานจนกว่าจะมีคุณสมบัติต่างๆ ทั้งหมด

เมื่อได้กำหนดโครงการว่าจะมีการทดสอบงานบุคคล โดยใช้แบบทดสอบมาตรฐาน เฉพาะโครงการนั้นๆ ผู้กำหนดโครงการก็ต้องปฏิบัติงานเป็นขั้นๆ ในการสร้างแบบทดสอบ วิเคราะห์แบบทดสอบปรับปรุงแก้ไข และจัดให้มีความเที่ยงตรง ในการปฏิบัติงานเหล่านี้จะต้องดำเนินการเป็นขั้นๆ ดังต่อไปนี้ คือ²⁴

1. กำหนดแนวคิดหลักที่จะทดสอบ ผู้วัดจะต้องมีความเข้าใจในลักษณะของงาน แต่ละขั้นที่จะเกี่ยวข้องกับการวัดนั้นอย่างแจ่มแจ้ง เพื่อที่จะวิเคราะห์ออกมาว่าต้องการจะวัดอะไร และในขณะเดียวกันจะต้องศึกษาลักษณะทางจิตวิทยาของมนุษย์ที่มีอิทธิพลปฏิบัติต่องานนั้นๆ เพื่อที่จะสร้างแบบทดสอบที่สัมพันธ์กับลักษณะทางจิตวิทยา

2. กำหนดรายละเอียดเกี่ยวกับการใช้แบบทดสอบในการดำเนินงาน ในขั้นแรกผู้วัดจะต้องกำหนดให้เห็นภาพรวมๆ ของแบบทดสอบนั้นว่าจะใช้ทำอะไรบ้าง เมื่อไร และทดสอบกับใคร รายละเอียดที่กำหนดนี้อาจแยกเป็นข้อๆ ได้ดังนี้

- ก. หน้าที่ของงานที่ต้องการใช้แบบทดสอบนั้นวัด
- ข. คำถามควรมีลักษณะอันใดบ้างและจำนวนเท่าใด
- ค. เนื้อหาที่ต้องการทดสอบและเวลาที่จะใช้ทดสอบ
- ง. ลักษณะของประชากรที่ต้องการทดสอบ
- จ. ระดับความยากง่ายของแบบทดสอบ
- ฉ. วิธีการทางสถิติที่จะใช้ในการเลือกและปรับปรุงคำถาม

3. วิเคราะห์ร่างแบบทดสอบด้วยวิธีการทางสถิติ เมื่อได้กำหนดแผนการตามข้อ 2 แล้ว และได้ปฏิบัติงานไปถึงขั้นที่ร่างแบบทดสอบแล้ว ผู้วัดก็ต้องนำหลักสถิติมา

²⁴ Thorndike, R.L., *Personnel Selection*, pp. 54—60.

Lawsh, C.H., and Balma, M.J., *Principles of Personnel Testing*, pp. 323—328.

วิเคราะห์ว่า ร่างแบบทดสอบนั้นจะต้องปรับปรุงแก้ไขเปลี่ยนแปลงอย่างไรบ้าง ร่างแบบทดสอบนั้นควรประกอบด้วยคำถามที่จะเป็นตัวแทนของประเด็นที่จะวัด พร้อมกับมีคำอธิบายวิธีการทดสอบเพื่อให้ทุกคนที่จะถูกวัดปฏิบัติการเหมือนกันในการทดสอบ ในการวิเคราะห์ร่างแบบทดสอบด้วยวิธีทางสถิตินี้ ผู้สร้างแบบทดสอบอาจจะรวบรวมบุคคลบางคน ซึ่งมีลักษณะเดียวกับบุคคลที่จะถูกวัดมาลองทดสอบเพื่อที่จะทำให้ทราบได้ว่า ผู้ถูกวัดนั้นได้ตอบแบบทดสอบในประเด็นและตามแนวทางที่ผู้สร้างแบบทดสอบประสงค์หรือไม่ เมื่อผู้สร้างแบบทดสอบได้ทราบข้อเท็จจริงเหล่านี้แล้ว ก็อาจปรับปรุงร่างนี้เสียใหม่ แล้วนำร่างใหม่ไปทดลองกับกลุ่มของบุคคลที่อาจจะถูกทดสอบเพื่อหาข้อมูลทางสถิติ ถ้าหากผู้สร้างแบบทดสอบไม่สามารถหาข้อมูลของบุคคลที่อาจจะถูกวัดได้ ก็อาจจะพยายามหาบุคคลที่มีลักษณะใกล้เคียงกับกลุ่มบุคคลที่จะถูกวัดให้มากที่สุด ข้อมูลทางสถิติที่ควรจะต้องวิเคราะห์ในขั้นนี้ก็คือค่าสหสัมพันธ์ของความเที่ยงตรงระหว่างคำถามแต่ละข้อกับประเด็นที่จะวัด สิ่งที่จะต้องปรับปรุงที่สำคัญที่สุดก็คือคำถามแต่ละข้อ ซึ่งต้องอาศัยหลักการวิเคราะห์ว่าคำถามแต่ละข้อนั้น มีค่าความเที่ยงตรงกับลักษณะต่าง ๆ ของงานที่จะวัดอย่างไร ในการวิเคราะห์นี้อาจทำให้ต้องเพิ่มเติมหรือตัดทอนคำถามบางข้อ ค่าสถิติที่วัดได้ว่าคำถามแต่ละข้อมีค่าสหสัมพันธ์กับสิ่งที่จะวัดเป็นปริมาณเท่าใด จะเป็นหลักในการวิเคราะห์เรื่องนี้ นอกจากนี้ผู้วัดจะต้องหาค่าความเชื่อถือได้ของแบบทดสอบเพื่อทำให้ทราบว่าแบบทดสอบนั้น ๆ มีความยาวเหมาะสมที่จะใช้เป็นเครื่องวัดบุคคลหรือไม่ ถ้าหากว่าแบบทดสอบนั้นต้องการวัดความสามารถในเรื่องเนื้อหา แบบทดสอบนั้นก็ควรจะต้องมีความยาวพอที่จะทำให้บุคคลที่จะถูกวัดส่วนมากสามารถตอบคำถามได้หมดหรือเกือบหมด ถ้าหากว่าเป็นแบบทดสอบที่วัดความเร็วในการปฏิบัติงานก็ควรจะยาวมากกว่าที่บุคคลซึ่งปฏิบัติงานเร็วที่สุดสามารถจะทำเสร็จ ในการวิเคราะห์ร่างแบบทดสอบด้วยวิธีการทางสถิตินี้ ยังอาจทำให้ผู้วัดสามารถทราบอีกด้วยว่า แนวการใช้แบบทดสอบนั้นควรจะปรับปรุงแก้ไขอย่างไร หลังจากที่ได้ปรับปรุงความยาว และคำถามแต่ละข้อแล้ว ผู้สร้างแบบทดสอบก็จะได้แบบทดสอบที่เกือบจะเป็นฉบับสมบูรณ์

4. นำแบบทดสอบมาหาค่าความเที่ยงตรง (validity) และวิเคราะห์ค่าเที่ยงตรงที่ได้ นั้นเกี่ยวกับมาตรการที่จะวัดอื่น ๆ ในการดำเนินงานขั้นสุดท้ายนี้ ผู้สร้างแบบทดสอบจะต้องกำหนดหาค่าความเที่ยงตรงของแบบทดสอบทั้งหมดกับตัวงานที่บุคคลจะปฏิบัติ ในการหาค่าความเที่ยงตรงของแบบทดสอบนี้ ผู้สร้างแบบทดสอบจะต้องวิเคราะห์เป็น 3 นัยประกอบกัน นัยแรกคือค่าคะแนนของบุคคลซึ่งถูกใช้เป็นกลุ่มทดลองของร่างแบบ

ทดสอบ นัยที่สองคือ มาตรการวัดความสำเร็จของงานซึ่งปฏิบัติโดยบุคคลกลุ่มเดียวกับนัยแรก และนัยสุดท้าย คือค่าความสัมพันธ์ทางสถิติระหว่างคะแนนที่ทดลองวัดได้กับมาตรการวัดความสำเร็จของงาน ประเด็นสำคัญของการที่จะทำให้แบบทดสอบมีความเที่ยงตรงได้ก็คือ การที่ได้พิจารณาอย่างรอบคอบแล้วว่า ร่างแบบทดสอบที่สร้างขึ้นนั้นควรจะต้องทดลองวัดผู้ใด เมื่อใด และเป็นจำนวนเท่าใด ร่างแบบทดสอบนั้นควรจะต้องใช้กับผู้สมัครงานใหม่ หรือควรจะต้องใช้กับผู้ปฏิบัติงานมาจนมีประสบการณ์พอสมควรแล้ว การที่จะตัดสินใจได้ถูกในเรื่องเหล่านี้ ก็ขึ้นอยู่กับสภาพการณ์และโอกาสอันสมควรซึ่งจะได้กล่าวต่อไป

ประเด็นแรกที่จะต้องพิจารณาในปัญหาเรื่องการสร้างความเที่ยงตรงนี้ก็คือ กลุ่มบุคคลและเวลาที่ควรใช้ในการทดลองร่างแบบทดสอบ ในการพิจารณาประเด็นนี้ ผู้สร้างแบบทดสอบจะต้องคำนึงว่าหากจะตัดสินใจเลือกกลุ่มบุคคล และเวลาที่ทดลองให้เหมาะสม ควรจะมีข้อแม้สองอย่าง อย่างแรกคือระยะเวลาที่ได้ผลคะแนนจากร่างแบบทดสอบที่ทดลองกับเวลาที่ได้มาตรการวัดความสำเร็จของงานควรจะต้องห่างกันน้อยที่สุด และอย่างที่สอง กลุ่มบุคคลที่จะถูกทดลองควรมีลักษณะ และอยู่ภายใต้สภาพการณ์ที่จะทำให้เกิดแรงกระตุ้น ซึ่งคล้ายหรือเหมือนกันกับบุคคลที่จะถูกวัดจริง ๆ ในโอกาสต่อไป ถ้าพิจารณาข้อแม้สองประการนี้แล้วก็แปลว่า ในการสร้างแบบทดสอบเพื่อคัดเลือกคนที่มาสมัครเข้าทำงาน ร่างแบบทดสอบนั้น ควรจะต้องกับบุคคลที่มาสมัครงานจริง ๆ เพื่อให้ได้กลุ่มบุคคลที่ถูกทดลองเหมือนกันกับบุคคลที่จะถูกวัด แต่ในขณะที่เดียวกันตามข้อแม้ดังกล่าว ผู้สร้างแบบทดสอบก็ควรจะต้องร่างแบบทดสอบเพื่อคัดเลือกคนสมัครเข้าทำงานกับคนที่เคยทำงานมาแล้ว และมีประสบการณ์พอสมควร เพื่อจะให้ได้ผลคะแนนของการทดลองกับมาตรการวัดความสำเร็จของงานได้โดยไม่ทิ้งระยะห่างกัน ฉะนั้นการปฏิบัติตามข้อแม้ดังกล่าวจึงทำทำจะขัดกัน

ถ้าหากว่าจะสร้างแบบทดสอบเพื่อวัดคนสมัครเข้าทำงานโดยทดลองร่างกับกลุ่มบุคคลที่สมัครเข้าทำงานจริง ๆ อาจกล่าวโดยย่อได้ว่าโดยความเป็นจริงนั้น จะทำให้เกิดปัญหาที่เห็นได้ชัดสองประการ ประการแรก การที่จะติดตามเพื่อหามาตรการวัดความสำเร็จของงาน มักจะต้องทำให้เสียเวลามาก เพราะบุคคลที่ถูกทดลองนั้น อาจจะต้องปฏิบัติงานไปในระยะเวลาหนึ่งซึ่งอาจจะนานเป็นปี ๆ จึงจะทำให้ทราบได้ว่าจะกำหนดวัดความสำเร็จของงานนั้นได้อย่างไร และประการที่สอง จำนวนบุคคลที่ถูกทดลอง

อาจจะกระจายหรือสูญหายไป เพราะบางคนอาจจะถูกไล่ออก หรืออาจจะย้ายงาน หรืออาจจะถูกมอบหมายให้ปฏิบัติหน้าที่ที่ต่างออกไป

ฉะนั้นเมื่อพิจารณาในแง่ความสะดวกแล้ว ถ้าหากตัวการสร้างแบบทดสอบเพื่อวัดคนที่สมัครคัดเลือกเข้าทำงาน ก็อาจจะใช้กลุ่มคนที่ปฏิบัติงานแล้วและมีประสบการณ์พอสมควรเป็นกลุ่มทดลอง อย่างไรก็ตาม การปฏิบัติเช่นนี้ย่อมทำให้เกิดปัญหาสองประการคือ ประการแรก คนที่ปฏิบัติงานแล้วเหล่านี้จะมีสภาพทางจิตใจเหมือนกับคนที่มาสมัครเข้าทำงานจริง ๆ หรือไม่ เพราะถ้าหากมีลักษณะไม่เหมือนกัน ค่าที่วัดได้จากการทดลองก็ย่อมจะบิดเบือนไป ไม่เป็นค่าที่แทนความสามารถและความพยายามของผู้สมัครงานจริง ๆ ปัญหาอีกประการหนึ่งก็คือว่า ค่าคะแนนที่วัดได้จากบุคคลเหล่านี้จะได้รับผลกระทบจากระยะเวลาที่ได้ปฏิบัติงานและประสบการณ์ในการทำงานของบุคคลเหล่านั้นหรือไม่

นอกจากเรื่องกลุ่มบุคคลและเวลาที่ควรใช้ในการทดลองแล้ว หากผู้สร้างแบบทดสอบประสงค์จะให้แบบทดสอบมีความเที่ยงตรง ยังควรจะต้องคำนึงถึงประเด็นต่อไป คือจำนวนของบุคคลที่ควรจะต้องใช้ในการทดลองงานบางอย่าง อาจมีบุคคลเข้ามาปฏิบัติงานปีละเป็นพัน ๆ หรือหมื่น ๆ คน แต่งานบางอย่างอาจมีการหมุนเวียนของบุคคลที่เข้ามาทำงานเพียงปีละเป็นร้อย ๆ คน หรือต่ำกว่านั้น ในงานที่มีการหมุนเวียนของบุคคลที่เข้ามาทำงานน้อย ผู้สร้างแบบทดสอบย่อมประสบปัญหาในการสร้างความเที่ยงตรง ยกตัวอย่างเช่น ในกรณีที่มีบุคคลเพียงร้อยคน และผู้สร้างแบบทดสอบวัดค่าสหสัมพันธ์ความเที่ยงตรงได้ 0.05 ในกรณีเช่นนี้แปลว่าค่าสหสัมพันธ์สำหรับรายหนึ่ง ๆ ในทุก ๆ 10 ราย อาจจะมีสูงได้ถึง 0.61 หรืออาจต่ำลงมาได้ถึง 0.3 ฉะนั้นการทำงานหนึ่งงานได้มีการหมุนเวียนของบุคคลที่เข้ามาทำงานมาเป็นพันหรือเป็นหมื่น ย่อมทำให้สามารถสร้างความเที่ยงตรงได้ดียิ่งขึ้น

แนวคิดในการสร้างแบบทดสอบผสม งานในตำแหน่งหน้าที่หนึ่ง ย่อมต้องการผู้ปฏิบัติงานที่มีคุณสมบัติหลาย ๆ อย่างประกอบกัน เช่น ต้องการสติปัญญาผสมกับความชำนาญและความตั้งใจพยายาม หรือต้องการความชำนาญงานผสมกับความพยายามและความสามารถปรับตัวได้ การที่จะวัดคุณสมบัติหลาย ๆ อย่างนี้ได้ ต้องใช้แบบทดสอบหลาย ๆ แบบผสมกัน (battery of tests) การที่นำแบบทดสอบมาผสมนี้มีปัญหาถึงสองชั้น คือ ควรจะเลือกแบบทดสอบอย่างใดมาผสมกัน และค่าที่วัดได้จากแต่ละแบบนั้นมีความหมายสัมพันธ์กันอย่างไร กล่าวโดยย่อได้ว่า ในการเลือกแบบมา

ผสมกันนั้น ผู้วัดจะต้องคำนึงถึงประเด็นสำคัญสองประเด็น คือ ประเด็นแรกได้แก่ค่าความเที่ยงตรงของแบบทดสอบแต่ละแบบจะต้องสูง และประเด็นที่สอง ค่าสหสัมพันธ์ของความเที่ยงตรงระหว่างแบบทดสอบต่าง ๆ จะต้องต่ำ

แนวคิดที่เป็นพื้นฐานในการเลือกแบบทดสอบต่าง ๆ มาผสมกันนี้ อาจกล่าวได้ว่ามีสองอย่าง²⁵ อย่างแรกคือการพิจารณาในแง่งาน โดยพยายามศึกษาให้ทราบว่าในการปฏิบัติงานนั้น ๆ ให้สำเร็จลุล่วงด้วยดีนั้นจะต้องทำการอะไรบ้าง แนวคิดอย่างที่สองก็คือ การพิจารณาในแง่ลักษณะของบุคคล โดยพยายามวิเคราะห์ว่าลักษณะอันใดบ้างของบุคคลที่จะช่วยให้สามารถปฏิบัติงานนั้น ๆ สำเร็จลุล่วงไปได้ด้วยดี

การสร้างแบบทดสอบผสมโดยพิจารณาในแง่งานนั้น ทำได้โดยพยายามสร้างแบบทดสอบที่จะมีลักษณะเป็นการวัดประเด็นต่าง ๆ ของงาน เช่น งานขับรถราง อาจวิเคราะห์ได้ว่าเป็นงานที่ต้องจับปุ่มมกลไกต่าง ๆ เพื่อให้รถสามารถเคลื่อนไปตามถนนได้โดยไม่ปะทะกับสิ่งต่าง ๆ หรืองานวิจัย อาจจะเป็นงานที่ต้องวิเคราะห์แบบการวิจัยเพื่อเลือกแบบที่เหมาะสมมาใช้ในการรวบรวมวิเคราะห์และแปลความหมายของข้อมูล ผู้วัดก็พยายามสร้างแบบทดสอบที่จะวัดสิ่งที่จะต้องทำเหล่านี้

ส่วนการสร้างแบบทดสอบผสมโดยพิจารณาในแง่ลักษณะต่าง ๆ ของบุคคลนั้น ทำได้โดยศึกษาว่า ลักษณะทางจิตวิทยาและพฤติกรรมของมนุษย์นั้นมีอะไรบ้าง และพยายามแยกออกไม่ให้ปะปนกัน เช่น แยกได้ว่าจะต้องมีความสามารถในการใช้ตัวเลข ความคล่องในการใช้ภาษา ความเร็วในการคิดหาเหตุผล เมื่อได้แยกแล้วก็พยายามสร้างแบบทดสอบที่จะวัดลักษณะทางจิตวิทยาและพฤติกรรมเหล่านี้ แล้วนำไปทดลองเพื่อหาข้อเท็จจริงว่า ลักษณะทางจิตวิทยาและพฤติกรรมนั้น ๆ มีส่วนทำให้สามารถในการปฏิบัติงานหนึ่ง ๆ ได้ดีอย่างไร

กล่าวโดยทั่วไปแล้ว แบบทดสอบผสมที่สร้างขึ้นโดยพิจารณาในแง่งานนั้น มีโอกาสที่จะมีความเที่ยงตรงมากกว่าแบบทดสอบผสมที่สร้างโดยพิจารณาในแง่บุคคล เพราะการสร้างโดยวิธีแรกนั้น ได้ถือเอาประเด็นต่าง ๆ ของงานเป็นสิ่งที่มีความสำคัญในการพิจารณา อย่างไรก็ตาม ในทางปฏิบัติ ในการสร้างแบบทดสอบโดยทั่วไป มักจะไม่ได้ยึดแบบหนึ่งแบบใดแต่เพียงอย่างเดียว ผู้สร้างแบบทดสอบพยายามใช้วิธีการต่าง ๆ ระหว่างทั้ง 2 วิธี กล่าวคือ มีการเน้นความสำคัญของการพิจารณาประเด็นต่าง ๆ ของงาน ร่วมไปกับการวิเคราะห์พฤติกรรมและลักษณะทางจิตวิทยาของมนุษย์

²⁵ Tiffin, J., and McCormick, E.J., *Industrial Psychology*, pp. 105—108. Thorndike, *op. cit.*, pp. 36—39.

เกณฑ์การวัด

ค่าคะแนนที่วัดได้เป็นสิ่งที่ไร้ความหมาย หากไม่มีเกณฑ์การวัด (norms) เป็นเครื่องประกอบการพิจารณา ในการสอบแข่งขัน หรือสอบเลื่อนชั้นข้าราชการก็ดี บุคคลทั่ว ๆ ไปอาจจะนึกว่า ผู้ที่ได้คะแนนเต็ม 100 คะแนน ย่อมเป็นผู้ที่ดีที่สุด และผู้ที่ได้คะแนนไม่ต่ำกว่า 50 หรือ 60 คะแนน เป็นผู้ที่ใช้การได้ สมควรได้รับการบรรจุแต่งตั้ง หรือเลื่อนชั้น การคิดเช่นนี้เป็นการที่กักอย่างขาดหลักเกณฑ์ การที่จะพิจารณาได้ว่า ค่าคะแนนที่วัดได้นั้น แสดงความสามารถที่ใช้ได้หรือใช้ไม่ได้ จะต้องถือเกณฑ์การวัดประกอบ หรือกล่าวอย่างง่าย ๆ ก็คือว่า จะต้องพิจารณาหลักการที่จะเทียบผลของการแสดงความสามารถ หรือพฤติกรรมตามเครื่องมือวัดนั้น ๆ ของบุคคลหนึ่ง ๆ ที่ถูกวัดกับบุคคลอื่น ๆ ในประเภทเดียวกัน กล่าวโดยทั่ว ๆ ไปแล้ว ค่าคะแนนที่วัดได้ครั้งหนึ่ง ๆ ไม่สามารถแสดงความสามารถของบุคคลที่สูงหรือต่ำ ที่ดีหรือเลว ค่าคะแนนที่วัดได้แต่ละครั้งแสดงได้แต่เพียงว่า บุคคลหนึ่ง ๆ ที่ถูกวัด มีความสามารถสูงกว่า หรือต่ำกว่าบุคคลอื่น ๆ ในกลุ่มเดียวกัน หรือแสดงว่าบุคคลนั้น มีความสามารถดีขึ้นหรือเลวลงเท่านั้น

เกณฑ์การวัดแบ่งได้เป็น 4 ประเภท²⁶ คือ เกณฑ์วัดตามอายุ (age norms) เกณฑ์วัดตามชั้นเรียน (grade norms) เกณฑ์เปอร์เซ็นต์ไคล์ (percentile norms) และค่าคะแนนมาตรฐาน (standard scores) เกณฑ์วัดตามอายุคือวิธีการหาค่าเฉลี่ยของคุณสมบัติของพฤติกรรมของบุคคลในอายุหนึ่ง ๆ ส่วนเกณฑ์วัดตามชั้นเรียนเป็นค่าเฉลี่ยของคุณสมบัติของพฤติกรรมของบุคคลที่อยู่ในชั้นเรียนหนึ่ง ๆ เกณฑ์การวัดทั้งสองนี้จำเป็นมากในการวัดเด็กและนักเรียน แต่ในการวัดบุคคลในงานบริหารบุคคลนั้น เป็นการวัดที่ต่างไปจากการวัดผลทางการศึกษา ฉะนั้น ในที่นี้จึงจะอธิบายรายละเอียดเฉพาะเกณฑ์การวัดสองประเภทหลังเท่านั้น

เกณฑ์เปอร์เซ็นต์ไคล์ การใช้เกณฑ์เปอร์เซ็นต์ไคล์ เป็นการนำค่าคะแนนที่วัดได้จากแต่ละบุคคลมาเทียบว่า ค่าคะแนนนั้นมีความสัมพันธ์ประการใดกับค่าคะแนนของบุคคลอื่น ๆ ที่อยู่ในประเภทเดียวกันกับบุคคลนั้น วิธีการเทียบค่าคะแนนโดยใช้เกณฑ์เปอร์เซ็นต์ไคล์นี้ มีอธิบายในตำราสถิติเบื้องต้นทั่วไป นอกจากนี้ ยังได้มีผู้คิดสร้างตารางเทียบค่าคะแนนตามเกณฑ์เปอร์เซ็นต์ไคล์นี้ เพื่อให้สะดวกแก่การพิจารณาอีกด้วย อย่าง

²⁶ Thorndike and Hagen, *op. cit.*, pp. 153—172.
Lyman, *op. cit.*, pp. 74—83.

ไรท์ ผู้ที่ศึกษาวิชาวัดผลควรจะทราบวิธีการพิจารณาค่าคะแนนตามเกณฑ์เปอร์เซ็นต์ไค์ ฉะนั้นในที่นี้จะได้อธิบายวิธีพิจารณาค่าคะแนนตามเกณฑ์นี้อย่างง่าย ๆ

สมมติว่า ในการวัดบุคคล 76 คน ค่าคะแนนที่วัดได้มีตั้งแต่ 40 ไปจนถึง 60 โดยบุคคลต่าง ๆ ที่ถูกวัดจะได้ค่าคะแนนเป็นลำดับดังต่อไปนี้ คะแนน 40 มี 2 คน คะแนน 42 มี 1 คน คะแนน 44 มี 2 คน คะแนน 45 มี 2 คน คะแนน 46 มี 2 คน คะแนน 47 มี 3 คน คะแนน 48 มี 3 คน คะแนน 49 มี 7 คน คะแนน 50 มี 6 คน คะแนน 51 มี 7 คน คะแนน 52 มี 5 คน คะแนน 53 มี 7 คน คะแนน 54 มี 6 คน คะแนน 55 มี 7 คน คะแนน 56 มี 5 คน คะแนน 57 มี 5 คน คะแนน 58 มี 3 คน คะแนน 59 มี 1 คน และคะแนน 60 มี 2 คน เมื่อจะหาค่าคะแนนไค์เป็นกัเปอร์เซ็นต์ไค์ก็นับความถี่สะสม (cumulative frequency) ของค่าคะแนนนั้น ๆ แล้วนำความถี่สะสมทั้งหมดมาหาร ผลลัพธ์ที่ได้คือค่าคะแนนตามความหมายของเกณฑ์เปอร์เซ็นต์ไค์ เช่น ต้องการทราบว่า ผู้ที่ได้คะแนน 45 ก็นับความถี่สะสมของค่าคะแนนนี้ ซึ่งได้เท่ากับ 7 แล้วนำความถี่สะสมทั้งหมดซึ่งมีเท่ากับ 76 ไปหาร ฉะนั้น ก็บอกได้ว่า คะแนน 45 = 9 เปอร์เซนต์ไค์ หรือคะแนนที่ 57 ก็ $= \frac{70}{76} = 92$ เปอร์เซนต์ไค์

ค่าคะแนน 9 เปอร์เซนต์ไค์ และ 92 เปอร์เซนต์ไค์นี้ หมายความว่าค่าที่วัดได้นั้น ๆ เมื่อเทียบกับบุคคลในกลุ่มเดียวกันจำนวน 100 ราย ที่แสดงคุณสมบัติของพฤติกรรมอย่างเดียวกัน จะมีผู้ที่มีคุณสมบัติที่วัดได้ต่ำกว่า 9 ราย และ 92 รายตามลำดับ ส่วนค่าคะแนนดิบ (raw scores) คือ 45 และ 57 นั้นไม่มีความหมาย ค่าคะแนนที่วัดด้วยเกณฑ์เปอร์เซ็นต์ไค์นี้ จึงมีประโยชน์และนิยมใช้อย่างกว้างขวางในวงการวัดผลบุคคลแบบทดสอบมาตรฐานหลายแบบ จะแสดงค่าคะแนนที่วัดได้เป็นเกณฑ์เปอร์เซ็นต์ไค์ เช่น ในตารางที่ 1

จากตารางที่ 1 จะเห็นได้ว่า เมื่อผู้ใดก็ตามนำแบบทดสอบนี้ไปใช้วัดบุคคลเพื่อให้ทราบความถนัดทางธรรมชาติพิเศษต่าง ๆ ของบุคคล ก็จะสามารถทราบความหมายของค่าคะแนนที่วัดได้ เช่น ถ้าวัดความมีเหตุผลในเรื่องภาษาได้คะแนน 17 คะแนน เมื่อคิดเป็นเปอร์เซ็นต์ไค์จะเท่ากับ 50 เปอร์เซนต์ไค์ แปลว่า เมื่อเทียบกับบุคคลประเภทเดียวกันกับผู้ถูกวัดเป็นจำนวน 100 ราย จะมีผู้ที่อยู่ต่ำกว่าผู้ถูกวัด 50 ราย หรือถ้าวัดความมีเหตุผลในเรื่องนามธรรมได้คะแนน 70 ก็เท่ากับ 90 เปอร์เซนต์ไค์ แปลว่า มีบุคคลประเภทเดียวกันกับผู้ถูกวัดใน 100 ราย ที่ต่ำกว่าผู้ถูกวัดถึง 90 ราย ฉะนั้น เครื่องมือวัดนี้ชี้ให้เห็นว่า ผู้ใดจะมีความสามารถในเรื่องใดมากกว่าผู้อื่นอย่างไร การที่ได้ถึง 90

ตารางที่ 1

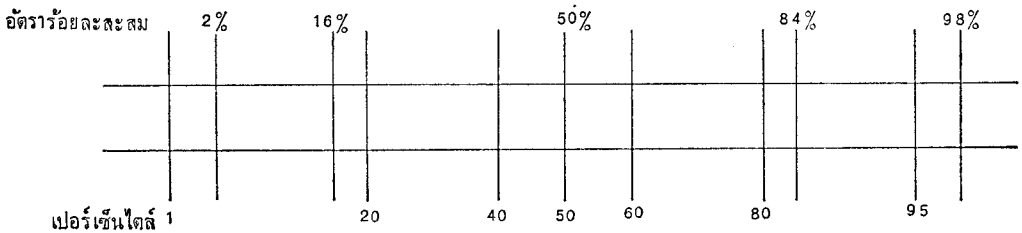
ค่าคะแนนตามเกณฑ์เปอร์เซ็นต์ไต่ของแบบทดสอบชุด DAT สำหรับเพศชาย*

คะแนนเต็ม เปอร์เซ็นต์ไต่	ความมีเหตุผล ในเรื่องภาษา	ความสามารถ ทางตัวเลข	ความมีเหตุผล ในเรื่องนามธรรม	มิติสัมพันธ์	ความมีเหตุผล เกี่ยวกับเครื่อง จักรกล
99	45+	35+	44+	87+	60+
97	36-44	32-34	41-43	81-86	56-59
95	33-35	30-31	39-40	75-80	53-55
90	30-32	27-29	37-38	69-74	50-52
85	27-29	25-26	35-36	64-68	48-49
80	25-26	23-24	34	60-63	46-47
75	24	22	32-33	56-59	44-45
70	22-23	21	31	53-55	42-43
65	21	19-20	30	49-52	41
60	19-20	18	29	45-48	39-40
55	18	17	27-28	41-44	37-38
50	17	16	26	37-40	35-36
45	16	15	24-25	33-36	34

เปอร์เซ็นต์ไต่นี้ ย่อมแสดงว่า บุคคลนั้น ๆ มีความสามารถ หรือมีคุณสมบัติสูงในเรื่องที่วัด อย่างไรก็ดี ในการพิจารณาค่าคะแนนที่วัดได้ตามเกณฑ์เปอร์เซ็นต์ไต่ที่มีข้อพึงระวังอยู่สองประการ ประการแรกคือ กลุ่มของบุคคลที่ถือว่าเป็นประเภทเดียวกับผู้ถูกวัด ค่าคะแนนเปอร์เซ็นต์ไต่ที่ดีจะต้องคำนวณจากกลุ่มบุคคลประเภทเดียวกันจริง ๆ เช่น เครื่องมือวัดที่ใช้คัดเลือกบุคคลเข้าทำงานในตำแหน่งเสมียนพนักงาน ก็ต้องใช้กลุ่มบุคคลที่สมัครเข้าทำงานในตำแหน่งนี้เป็นเกณฑ์ในการหาเปอร์เซ็นต์ไต่ ฉะนั้น เครื่องมือวัดหนึ่ง ๆ จะต้องแสดงค่าเปอร์เซ็นต์ไต่ของกลุ่มบุคคลไว้หลาย ๆ ประเภท เพื่อให้สะดวกแก่การใช้ และในบางโอกาสผู้ใช้ก็อาจจะต้องหาเกณฑ์เปอร์เซ็นต์ไต่จาก

* แหล่งที่มา : Thorndike, R.L., and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*, p. 163.

บุคคลที่ปฏิบัติงานกับหน่วยงานของตนโดยเฉพาะ เพื่อให้ค่าคะแนนที่คำนวณเป็นเปอร์เซ็นต์นี้วัดบุคคลที่จะปฏิบัติงานในหน่วยงานของตนได้อย่างมีความหมายโดยแท้จริง ประการที่สอง ผู้พิจารณาค่าคะแนนที่วัดได้โดยเกณฑ์เปอร์เซ็นต์จะต้องระลึกเสมอว่า ค่าเปอร์เซ็นต์ไม่ได้แสดงความถี่ระหว่างแต่ละเปอร์เซ็นต์เท่า ๆ กัน เช่น เปอร์เซ็นต์ที่ 1 ไม่ได้ห่างจากเปอร์เซ็นต์ที่ 3 เท่ากับเปอร์เซ็นต์ที่ 50 ห่างจากเปอร์เซ็นต์ที่ 52 ในการพิจารณาความถี่ของแต่ละเปอร์เซ็นต์ ผู้วัดจะต้องกลับไปนึกถึงความหมายของส่วนโค้งปกติ ในแผนภาพที่ 1 แสดงความถี่ของเปอร์เซ็นต์ต่าง ๆ



แผนภาพที่ 1 เปอร์เซ็นต์ไต่เทียบกับอัตราร้อยละของผู้ถูกวัด

จากแผนภาพที่ 1 จะเห็นว่า ความถี่ระหว่างเปอร์เซ็นต์ที่ 40-50 และ 50-60 นั้น ต่างจากความถี่ระหว่างเปอร์เซ็นต์ที่ 20-40 และเปอร์เซ็นต์ที่ 60-80 จากรูปแบบของเส้นโค้งปกติ ซึ่งมีแสดงไว้ในตารางทั่วไปจะทำให้สามารถคำนวณได้ว่า เมื่อห่างจากค่าเฉลี่ยออกไปทั้งทางข้างบวกหรือลบแล้ว จำนวนอัตราร้อยละของบุคคลหรือสิ่งที่ถูกวัด จะมีความถี่น้อยลง ในที่นี้ จะเพียงแต่ย้ำว่า ผู้แปลความหมายของค่าคะแนนที่วัดได้ตามเกณฑ์เปอร์เซ็นต์นั้นจะต้องไม่ถือว่า แต่ละเปอร์เซ็นต์มีความถี่เท่า ๆ กัน หรือกล่าวอีกนัยหนึ่งก็คือ จะไม่ผลิผลตามแปลความหมายว่า ผู้ที่ได้ 70 เปอร์เซ็นต์ไต่เด่นกว่าผู้ที่ได้ 60 เปอร์เซ็นต์ไต่ เหมือนกับผู้ที่ได้ 20 เปอร์เซ็นต์ไต่ เก่งกว่าผู้ที่ได้ 10 เปอร์เซ็นต์ไต่ และยิ่งไปกว่านั้น ผู้วัดจะไม่ที่กักว่า ควรจะบรรจุหรือแต่งตั้งผู้ที่ได้เปอร์เซ็นต์สูงเข้าปฏิบัติงานได้ทันที เพราะถึงแม้ว่า คะแนนเปอร์เซ็นต์นี้จะเป็นการแปลความหมายค่าคะแนนอย่างมีเกณฑ์ก็จริง แต่ค่าคะแนนนี้ยังไม่ได้พิจารณาถึงความบกพร่องของการวัดในด้านอื่น ๆ คือ ความเชื่อมั่น (reliability)

และความเที่ยงตรง โดยเฉพาะในกรณีที่เครื่องมือวัดขาดความเที่ยงตรงผู้ที่ได้เปอร์เซ็นต์ไต่ 90 ก็ยังแปลไม่ได้ว่าจะมีบุคคลที่มีคุณสมบัติที่วัดได้ต่ำกว่าผู้นั้นถึงร้อยละ 90 ราย

ค่าคะแนนมาตรฐาน ผู้วัดอาจนำค่าเบี่ยงเบนมาตรฐาน (standard deviation) มาประกอบการพิจารณาค่าคะแนนที่วัดได้ เพื่อให้การแปลความหมายของคะแนนที่วัดได้นั้นมีหลักเกณฑ์ ค่าเบี่ยงเบนมาตรฐานมีอธิบายในตำราสถิติเบื้องต้นทั่วไป วิธีหาค่าเบี่ยงเบนมาตรฐานที่ง่ายที่สุดก็คือ นำค่าที่วัดได้แต่ละจำนวนหักจากค่าเฉลี่ย แล้วเอาผลลัพธ์แต่ละตัวยกกำลังสอง แล้วรวมกันจึงหารด้วยจำนวนของจำนวนทั้งหมด และนำผลลัพธ์ไปถอดครณฑ์สอง เมื่อทราบค่าเบี่ยงเบนมาตรฐานแล้ว ก็นำค่านี้มาเทียบกับค่าที่วัดได้แต่ละครั้ง ทำให้ทราบว่า ค่าที่วัดได้แต่ละครั้งเบี่ยงเบนไปเป็นปริมาณเท่าใด เช่น ถ้าสมมติว่า มีการสอบเลื่อนชั้นข้าราชการจากตำแหน่งหนึ่งไปอีกตำแหน่งหนึ่ง และมีข้อสอบสองฉบับ ฉบับที่หนึ่ง สอบความรู้ทางเศรษฐศาสตร์ อีกฉบับหนึ่งสอบความรู้เฉพาะตำแหน่ง ผลการสอบปรากฏว่า นางสาว ก. สอบได้คะแนนฉบับแรก 77 คะแนน และฉบับที่สอง 55 คะแนน ส่วนนาย ข. สอบฉบับแรกได้ 87 คะแนน และฉบับหลังได้ 48 คะแนน ค่าเฉลี่ยของคะแนนของผู้เข้าสอบทั้งหมดในฉบับแรกเท่ากับ 65 คะแนน และฉบับหลังเท่ากับ 40 คะแนน ส่วนค่าเบี่ยงเบนมาตรฐานของฉบับแรกเท่ากับ 15 และฉบับหลังเท่ากับ 10 ในกรณีเช่นนี้ก็อาจคำนวณค่าคะแนนที่วัดได้ของนางสาว ก. และนาย ข. เทียบกับค่าเบี่ยงเบนมาตรฐานได้ดังนี้

$$\text{นางสาว ก. ได้คะแนนสูงกว่าค่าเฉลี่ยในฉบับแรก} = 77 - 65$$

$$= 12$$

$$\text{ฉะนั้น คะแนนในฉบับแรกของนางสาว ก. เบี่ยงเบน} = \frac{12}{15}$$

$$= 0.8$$

$$\text{นางสาว ก. ได้คะแนนสูงกว่าค่าเฉลี่ยในฉบับที่ 2} = 55 - 40$$

$$= 15$$

$$\text{ฉะนั้น คะแนนนางสาว ก. ในฉบับที่ 2 เบี่ยงเบน} = \frac{15}{10}$$

$$= 1.5$$

$$\text{นาย ข. ได้คะแนนสูงกว่าค่าเฉลี่ยในฉบับแรก} = 87 - 65$$

$$= 22$$

$$\begin{aligned}
 \text{ฉะนั้น ค่าคะแนนในฉบับแรกของนาย ข. เบียงเบน} &= \frac{22}{15} \\
 &= 1.5 \\
 \text{นาย ข. ได้คะแนนสูงกว่าค่าเฉลี่ยในฉบับที่ 2} &= 48-40 \\
 &= 8 \\
 \text{ฉะนั้น ค่าคะแนนในฉบับที่ 2 ของนาย ข. เบียงเบน} &= \frac{8}{10} \\
 &= 0.8
 \end{aligned}$$

การที่จะถือว่านางสาว ก. หรือนาย ข. ได้คะแนนดีหรือไม่ดีจากค่าคะแนนดิบนั้น ไม่ถูกต้อง ต่อเมื่อได้พิจารณาเทียบกับค่าคะแนนมาตรฐานแล้ว จึงจะเห็นได้ชัดว่า นางสาว ก. แสดงพฤติกรรมหรือความสามารถตามข้อสอบวิชาเศรษฐศาสตร์ได้หย่อนกว่าในวิชาความรู้เฉพาะตำแหน่ง และเมื่อเทียบกับนาย ข. แล้ว นางสาว ก. มีความรู้เฉพาะตำแหน่งดีเท่า ๆ กับที่นาย ข. มีความรู้ในวิชาเศรษฐศาสตร์ ต่อเมื่อพิจารณาค่าคะแนนที่วัดได้เช่นนั้น การวินิจฉัยสั่งการในการบริหารงานบุคคลจึงจะเป็นไปตามหลักเกณฑ์

เนื่องจากการเทียบค่าคะแนนที่วัดได้กับค่าเบียงเบนมาตรฐานโดยตรง ทำให้ผลลัพธ์ที่คำนวณได้บางครั้งอาจเป็นทศนิยม หรือบางครั้งติดลบ เป็นการไม่สะดวกแก่การพิจารณาจึงได้มีผู้คิดวิธีแตกแขนงออกไปหลายวิธี เช่น ตั้งเกณฑ์ให้มีค่าเฉลี่ยสมมติเท่ากับ 50 และมีค่าเบียงเบนมาตรฐานเท่ากับ 10 หรือตั้งเกณฑ์ให้มีค่าเฉลี่ยสมมติเท่ากับ 100 และค่าเบียงเบนมาตรฐานเท่ากับ 20 วิธีแรกนั้นใช้ในระบบการบริหารบุคคลของกองทัพเรือสหรัฐ และวิธีที่สองนิยมใช้ในกองทัพบกสหรัฐ²⁷ การคำนวณตามวิธีแรก ทำได้ง่าย ๆ โดยเอาเลข 10 ไปคูณค่าเบียงเบนมาตรฐานที่คำนวณได้ แล้วบวกกับ 50 เช่น ในกรณีของผลการวัดนางสาว ก. ในวิชาเศรษฐศาสตร์ก็จะได้เท่ากับ $(0.8)(10) + (50)$ คือ 58 ส่วนวิธีที่สองก็ทำได้ในทำนองเดียวกัน นอกจากนี้ผู้วัดอาจนำค่าคะแนนที่คำนวณตามเกณฑ์ค่าคะแนนมาตรฐานนี้ไปเทียบกับเส้นโค้งปกติ เพื่อให้ค่าคะแนนนี้มีส่วนละม้ายกับการแจกแจงปกติ (normal distribution) โดยอาศัยหลักคณิตศาสตร์ได้

²⁷ Thorndike and Hagen, *op. cit.*, p. 167.

บทที่ 2

การวิเคราะห์งานและการประเมินผลงาน

เครื่องมือวัดบุคคลที่จะใช้ได้อย่างมีประสิทธิภาพในการบริหารงานนั้น จะต้องเป็นเครื่องมือวัดที่มีคุณค่าในการคาดคะเนได้ว่า ผลของการวัดตามเครื่องมือวัดนั้น ๆ ในระดับต่าง ๆ แสดงว่า บุคคลที่ถูกวัดจะปฏิบัติงานสำเร็จได้ผลดีเพียงใด การที่จะสร้างเครื่องมือวัดได้เช่นนั้น จำต้องมีเกณฑ์เปรียบเทียบในระหว่างการทดลองสร้างเครื่องมือวัด ซึ่งเรียกว่า มาตรฐานวัดความสำเร็จของงาน (criterion measure) การที่จะทราบเกณฑ์นี้ได้ ผู้สร้างเครื่องมือวัดจะต้องมีข้อมูลจากการวิเคราะห์งานและการประเมินผลงาน ฉะนั้น แม้ว่าการวิเคราะห์งานและการประเมินผลงาน เป็นงานอีกด้านหนึ่ง โดยเฉพาะของการบริหารงานบุคคล ผู้ศึกษาเรื่องการวัดผลบุคคลก็จำเป็นที่จะต้องทราบหลักเกณฑ์ในเรื่องนี้ เพื่อประกอบการศึกษาด้วย

ความหมายของการวิเคราะห์งาน

การวิเคราะห์งาน (job analysis) หมายถึง การศึกษาพิจารณางานในตำแหน่งต่าง ๆ ของหน่วยงาน ถือเป็นกระบวนการขั้นต้นในการจัดการงานบุคคล และเป็นขั้นตอนที่ต่อเนื่องจากการจัดหน่วยงาน (organization design) กล่าวคือ เมื่อกำหนดเป้าหมายของงานแล้ว กำหนดหน้าที่ที่จำเป็นต้องจัดทำเพื่อให้งานสำเร็จตามเป้าหมายแล้ว ต่อจากนั้น เป็นการจัดกลุ่มหน้าที่ต่าง ๆ ตามลักษณะงานที่ใกล้เคียงกัน เชื่อมโยงกัน และลดหลั่นกันเป็นระดับต่าง ๆ ด้วยการดำเนินการดังกล่าวโดยสรุปนี้ เป็นเรื่องของการจัดหน่วยงาน ขึ้นต่อไปจึงเป็นการวิเคราะห์งานต่าง ๆ โดยละเอียด เพื่อระบุให้ชัดเจนว่างานแต่ละอย่างมีอะไรบ้าง และจะต้องการคนประเภทไหนมาทำงานเหล่านั้น¹

¹ Amrine, H.T.; Ritchey, J.A.; and Hulley, O.S., *Manufacturing Organization and Management*, pp. 26-35.

ความมุ่งหมายของการวิเคราะห์งาน

การวิเคราะห์งานมีความสำคัญต่อการทำงานส่วนรวมของหน่วยงาน เป็นพื้นฐานในการบริหารงานหลายเรื่อง ดังจำแนกได้ดังต่อไปนี้²

1. การจัดคนให้เพียงพอกับงาน จากการสำรวจ ศึกษางาน ทำให้ทราบลักษณะของงานที่ต้องทำในปัจจุบัน และความต่อเนื่องของงานที่จะต้องทำในระยะเวลาต่อไป ทำให้สามารถจัดสรรคนให้เพียงพอตามหมวดหมู่ของงานได้

2. การสรรหาคนและการเลือกสรรคน เมื่อได้พิจารณางานโดยละเอียด จัดจำแนกเป็นหมวดหมู่ เป็นระดับต่างๆ แล้ว ก็ใช้เป็นแนวในการสรรหาคนได้ถูกต้องและเลือกสรรคนได้ตรงกับงาน

3. การจัดโครงการฝึกอบรม ก่อนการฝึกอบรมจำเป็นต้องรู้ว่างานที่จะต้องทำมีอะไรบ้าง คนที่จะทำงานต้องรู้อะไรบ้าง ขาดความชำนาญชำนาญ หรือต้องเพิ่มเติมด้วยการฝึกอบรมอย่างไร รายละเอียดเหล่านี้จะได้จากการวิเคราะห์งานและข้อกำหนดของคนที่จะทำงานนั้น ๆ

4. การบริหารงานค่าจ้างและเงินเดือน การกำหนดค่าจ้างและเงินเดือนให้แก่ตำแหน่งต่างๆ เป็นการตราค่างานในตำแหน่งเหล่านั้นออกมาเป็นเงิน เป็นการเปรียบเทียบหรือชั่งน้ำหนักระหว่างงานที่คนจะต้องทำในแต่ละตำแหน่งกัน เงินที่จะให้เป็นการตอบแทนการทำงานนั้น ข้อมูลเกี่ยวกับงานที่นำมาประกอบการพิจารณานั้น ได้มาจากการวิเคราะห์งาน

5. การโอนย้าย และการเลื่อนชั้นเลื่อนตำแหน่ง การโอนย้ายบุคคลในตำแหน่งเป็นการพิจารณาคนกับงานในตำแหน่งระดับเดียวกัน แต่โอนย้ายเพื่อเปลี่ยนแปลงให้เหมาะสม ความเหมาะสมดังกล่าวอาจจะเป็นความชำนาญของคนเหมาะกับงาน บุคลิกภาพของคนเหมาะกับสิ่งแวดล้อมในการทำงาน ความรู้และประสบการณ์เหมาะกับงาน เป็นต้น ส่วนการเลื่อนชั้นเลื่อนตำแหน่งเป็นการพิจารณาคนกับงานในตำแหน่งระดับสูงขึ้น คนที่จะได้รับการพิจารณานั้น มีผลงานที่ได้ปฏิบัติแล้วเป็นประโยชน์เป็นผลดีต่อหน่วยงาน และมีศักดิ์พอที่จะทำงานในตำแหน่งสูงขึ้นนั้นได้ งานในตำแหน่งระดับเดียวกันหรือระดับสูงขึ้นที่จะต้องพิจารณาเหล่านี้ ต้องเป็นที่รู้ชัดเจนน้อยก่อนแล้วจึงจะพิจารณาได้ถูกต้องเหมาะสม

² Chruden, H.J., and Sherman, Jr., A.W., *Personnel Management*, pp. 84—87.

6. การพิจารณาคำร้องทุกข์ เมื่อมีการร้องทุกข์เกี่ยวกับปริมาณงาน ความไม่สะดวกในการทำงาน อุปสรรค ปัญหาเกี่ยวกับภาวะแวดล้อมในการทำงาน ปัญหาเกี่ยวกับค่าจ้างเงินเดือน การโอนย้ายและการเลื่อนชั้นเลื่อนตำแหน่ง ข้อมูลสำคัญอย่างหนึ่งที่ใช้ในการพิจารณาคำร้องทุกข์เหล่านี้ คือ ข้อมูลที่ให้ความกระจ่างเกี่ยวกับงานที่ผู้เสนอคำร้องทุกข์ทำอยู่

7. การปรับปรุงแก้ไขสภาพการทำงาน สภาพการทำงานเป็นส่วนหนึ่งของการทำงาน จากการสำรวจวิเคราะห์งาน ก็จะทราบว่าในกระบวนการทำงานแต่ละขั้นตอนดำเนินการอะไร อย่างไร ในสภาพการทำงานเช่นไร ผู้มีหน้าที่จัดหน่วยงานมีโอกาสได้ปรับปรุงแก้ไขสภาพการทำงานให้เหมาะสมได้

8. การกำหนดมาตรฐานการทำงาน ก่อนที่จะกำหนดมาตรฐานการทำงานต้องรู้ลักษณะการทำงานโดยละเอียด แล้วจึงกำหนดมาตรฐานการทำงานโดยการทดลองทำงาน หรือโดยอาศัยประสบการณ์ที่ผ่านมาแล้วเป็นเกณฑ์ประกอบการกำหนด

9. การปรับปรุงผลผลิตของคนทำงาน จากการสำรวจวิเคราะห์งานที่ดำเนินการอยู่แล้ว ทำให้ทราบข้อมูลเกี่ยวกับวิธีทำงาน และข้อขัดข้องเกี่ยวกับคนทำงาน ซึ่งจะช่วยให้หาแนวทางปรับปรุงแก้ไขให้ได้ผลงานดีขึ้น

10. การวางแผนของหน่วยงาน ข้อมูลสรุปจากการวิเคราะห์งานจะใช้ประโยชน์เป็นพื้นฐานในการวางแผนเกี่ยวกับงานที่จะต้องทำ คนที่จะต้องใช่ ตลอดจนปัจจัยประกอบการทำงานอื่น ๆ ยิ่งกว่านั้น ข้อมูลที่ได้จากการวิเคราะห์ยังมีคุณค่าพิเศษเพิ่มเติมจากในเรื่องหลัก ๆ ที่กล่าวแล้ว คือ เมื่อพิจารณาในขอบเขตกว้างกว่าหน่วยงานหนึ่ง หน่วยงานเดียวแล้ว ก็เป็นเครื่องแสดงความต้องการกำลังคนในงานต่าง ๆ โดยส่วนรวม และเป็นข้อมูลที่เป็นประโยชน์ในการแนะแนวอาชีพและการฝึกหัดอาชีพด้วย เมื่อพิจารณาในแง่คุณประโยชน์ของตัวผู้ทำงานเองก็จะมี ความพอใจในงาน และทำงานโดยมีประสิทธิภาพ เมื่อได้รับการบรรจุให้ทำงานที่เหมาะสมกับตน นอกจากนั้น ในเรื่องของการป้องกันภัยจากสภาพการทำงานที่เป็นอันตรายร้ายแรง การศึกษาวิเคราะห์งานก็มีส่วนช่วยกันก่อนแก้ไขได้

ขั้นตอนเบื้องต้นในการวิเคราะห์งาน

ในการวิเคราะห์งาน ควรจะต้องพิจารณาดำเนินการเป็นขั้นตอนดังต่อไปนี้ คือ ขั้นแรก พิจารณาดำเนินการหรือหน่วยงานที่จะต้องรับผิดชอบในการปฏิบัติการ ขั้นที่สอง

พิจารณาว่า โครงการวิเคราะห์งานนั้น ๆ มีคุณสมบัติครบถ้วนหรือไม่ และขั้นสุดท้ายจึงเลือกวิธีที่เหมาะสมในการปฏิบัติการ เนื่องจากวิธีการวิเคราะห์งานนั้นมีเทคนิคและรายละเอียดมากจึงจะกล่าวแยกเป็นอีกตอนหนึ่งต่างหาก ในตอนนี้จะพิจารณาแต่เพียงการกำหนดผู้รับผิดชอบและการวางโครงการเสียก่อน

ผู้รับผิดชอบการวิเคราะห์งาน การวิเคราะห์งานเพื่อประโยชน์ในการวัดผลบุคคลไม่ได้หมายความว่า นักวัดผลบุคคลจะต้องทำหน้าที่วิเคราะห์งาน ฉะนั้น จึงมีปัญหาเกิดขึ้นว่า ควรจะให้ผู้ใด หรือหน่วยงานใดเป็นผู้รับผิดชอบงานนี้ เมื่อกล่าวโดยทั่วไปแล้ว บุคลากร หรือ หน่วยงานบุคลากรควรจะเป็นผู้รับผิดชอบในการดำเนินการวิเคราะห์งาน การที่นักวัดผลเป็นบุคคลคนละคนกับนักวิเคราะห์งานเช่นนี้ จึงเกิดมีประเด็นที่สำคัญอีกประเด็นหนึ่งคือ ปัญหาที่ว่าผู้บริหารหน่วยงานนั้น ๆ เห็นความสำคัญที่จะต้องมีกรวิเคราะห์งานหรือไม่³ ผู้บริหารอาจเห็นว่า จำเป็นต้องมีการวิเคราะห์งานในหน่วยงานของตน เพื่อประโยชน์ในการปรับปรุงแก้ไขอัตราเงินเดือน หรือช่วยในการสรรหาคน หรือในการฝึกอบรมพนักงานและข้าราชการในระหว่างประจำการ ประเด็นที่สำคัญในที่นี้จึงมีอยู่ว่า ผู้บริหารจะต้องเข้าใจความสำคัญของข้อมูลและผลการศึกษาจากการวิเคราะห์งานที่มีต่อการสร้างเครื่องมือวัดผลบุคคล

ความเข้าใจของผู้บริหารในเรื่องนี้จะเป็นประโยชน์ต่อการที่จะเริ่มดำเนินการวิเคราะห์งาน เพราะจะช่วยให้การวิเคราะห์งานซึ่งมีผลกระทบต่อหน่วยงานภายในต่าง ๆ ดำเนินไปด้วยความราบรื่น

การวิเคราะห์งานที่จะเป็นประโยชน์ต่อการสร้างเครื่องมือวัดบุคคล อาจจะต้องมีรายละเอียดในการศึกษามากกว่าการวิเคราะห์งาน เพื่อประโยชน์อื่น ๆ เจตนาของกรวิเคราะห์งานย่อมขึ้นอยู่กับความประสงค์ของผู้บริหารและหน่วยงานแต่ละแห่ง เมื่อผู้บริหารเห็นว่าโครงการฝึกอบรมอาจมีการปรับปรุงได้ ถ้าหากว่าเข้าใจลักษณะของงานที่ผู้เข้ารับการอบรมปฏิบัติดียิ่งขึ้น หรือปัญหาบางอย่างในการคัดเลือกบุคคลเข้าทำงาน อาจแก้ไขได้ หากได้ข้อมูลจากการวิเคราะห์งาน ผู้บริหารก็อาจริเริ่มให้มีการวิเคราะห์งานในหน่วยงานของตนได้ แต่ในกรณีการสร้างเครื่องมือวัดผลบุคคล ผู้บริหารอาจเห็นว่ากรวิเคราะห์งานเพื่อความมุ่งหมายนี้ประการเดียวไม่คุ้มกับค่าใช้จ่ายในการดำเนิน

³ Negro, F.A., *Public Personnel Administration*, pp. 88—89.

การทั้งหมด ฉะนั้นโดยทั่วไป นักสร้างเครื่องมือวัดผลบุคคลอาจจะต้องอาศัยข้อมูล และผลการศึกษาจากการวิเคราะห์งานที่ดำเนินการด้วยวัตถุประสงค์อื่น ๆ ที่นอกเหนือไปจากการสร้างเครื่องมือวัดบุคคล ในกรณีเช่นนี้ผู้สร้างเครื่องมือวัดผลบุคคลควรจะมีส่วนในการพิจารณารายละเอียดของโครงการวิเคราะห์งาน เพื่อให้แน่ใจว่า ตามวัตถุประสงค์ของโครงการวิเคราะห์งานนั้น ๆ จะมีโอกาสได้ข้อมูล และผลการศึกษาคบถวนพอที่จะนำมาประกอบในการสร้างเครื่องมือวัดด้วยหรือไม่ ทั้งนี้ เพราะว่าโครงการวิเคราะห์งานอันจัดวางตามวัตถุประสงค์หนึ่ง ๆ นั้น ไม่จำเป็นจะต้องให้ข้อมูลและผลการศึกษาที่มีลักษณะสมบูรณ์ในการประกอบการสร้างเครื่องมือวัดเสมอไป การวิเคราะห์งานเพื่อประโยชน์ในการปรับปรุงโครงการฝึกอบรมพนักงานและข้าราชการประจำการ อาจขาดรายละเอียดบางอย่างที่จำเป็นต่อการสร้างเครื่องมือวัด แต่โครงการวิเคราะห์งานเพื่อประโยชน์ในการจำแนกตำแหน่ง อาจมีรายละเอียดเกือบครบถ้วนสำหรับนักวิเคราะห์งาน

การวิเคราะห์งานเป็นงานของผู้ที่ได้รับการฝึกฝนในเรื่องนี้โดยเฉพาะ ในทางวิชาการได้มีผู้เสนอว่า การวิเคราะห์งานตามโครงการหนึ่ง ๆ ควรจะค่อย ๆ เริ่มงานแล้วจึงขยายต่อไปตามลำดับ ทั้งนี้ อาจเริ่มต้นโดยมีนักวิเคราะห์งานที่ชำนาญงานเพียงคนเดียวก่อน⁴ ในระยะต่อไป จึงมีการขอยืมตัวเจ้าหน้าที่จากฝ่ายต่าง ๆ มาร่วมงานโดยให้นักวิเคราะห์งานผู้หนึ่งเป็นผู้ฝึกสอน โดยหลักทั่วไป นักวิเคราะห์งานนี้จะต้องเป็นผู้ที่รู้จักงานของหน่วยงานนั้น ๆ เป็นอย่างดี ฉะนั้นหากหน่วยงานหนึ่ง ๆ มอบอำนาจให้บุคคลภายนอกหรือหน่วยงานภายนอกมาดำเนินการวิเคราะห์งาน นักวิเคราะห์งานหรือหน่วยงานภายนอกนี้ ก็ควรจะได้รับความชี้แจงจากบุคคลภายในหน่วยงานที่รู้เรื่องงานในหน่วยงานนี้เป็นอย่างดี⁵

การวิเคราะห์งาน เป็นงานที่ต้องการความร่วมมือจากบุคคลทุกระดับในหน่วยงาน⁶ เจ้าหน้าที่ในระดับสูงควรเข้าใจในคุณค่าของการวิเคราะห์งาน ตลอดจนขอบเขต ค่าใช้จ่าย และเวลาที่จะต้องใช้ในการดำเนินการ เจ้าหน้าที่ในระดับหัวหน้างานควรจะได้รับความชี้แจงเกี่ยวกับการวางโครงการวิเคราะห์งานแต่เนิ่น ๆ และบางครั้งควรขอความร่วมมือ

⁴ Mee, J.F., ed., *Personnel Handbook*, p. 140.

⁵ *Ibid.*, pp. 140—141.

⁶ Negro, *op. cit.*, pp. 89—96.

เมื่อจากเจ้าหน้าที่ในระดับนี้ ให้มีส่วนร่วมในการพิจารณาแนวทางเบื้องต้นของการวางโครงการ เพื่อเป็นการสร้างความเข้าใจอันดีในการที่จะปฏิบัติต่อไป ส่วนเจ้าหน้าที่ในระดับผู้ปฏิบัติงานนั้น ก็อาจสร้างความเข้าใจในเรื่องนี้ได้ด้วยวิธีการต่างๆ เช่น ขอให้เจ้าหน้าที่ในระดับหัวหน้างานไปชี้แจงอธิบาย หรือผู้รับผิดชอบในการดำเนินการวิเคราะห์งานอาจจะชี้แจงเรื่องผ่านทางเจ้าหน้าที่ระดับปฏิบัติงานบางคนก็ได้

หลักเบื้องต้นในการวางโครงการวิเคราะห์งาน โครงการวิเคราะห์งานต่างๆ ไป ควรจะมีคุณสมบัติครบถ้วน 5 ประการดังต่อไปนี้⁷

1. จะต้องมียุทธประสงค์แน่นอน เพื่อให้ทราบว่าจะข้อมูลและผลการศึกษานั้นจะนำไปใช้ประโยชน์อันใด ก่อนที่จะมีการเริ่มโครงการ

2. จะต้องกำหนดวิธีการในการรวบรวมข้อมูลอันเที่ยงตรงไว้ให้แน่ชัด เพื่อให้ทราบว่าจำเป็นต้องใช้นักวิเคราะห์งาน และเจ้าหน้าที่อย่างไร ด้วยเทคนิคอันใด

3. เจ้าหน้าที่ทุกระดับของหน่วยงานจะต้องเข้าใจลักษณะทั่วไปของโครงการ และจะต้องหาทางสร้างความร่วมมือกับบุคคลเหล่านี้

4. จะต้องพยายามดำเนินโครงการให้ต่อเนื่อง เพื่อให้สอดคล้องกับสภาวะความเปลี่ยนแปลงภายในหน่วยงาน และความเปลี่ยนแปลงในเนื้อหาของงาน พิจารณาในการปฏิบัติงาน และสภาพการณ์ปฏิบัติงาน

5. ควรมีคณะกรรมการวิเคราะห์งานโดยเฉพาะ เพื่อที่จะหาทางสร้างความร่วมมือระหว่างฝ่ายต่างๆ ในหน่วยงานให้ดีที่สุด คณะกรรมการนี้จะไม่ปฏิบัติงานซ้อนกับหน่วยบริหารงานบุคคล หรือนักวิเคราะห์งาน ซึ่งจะรับผิดชอบในการปฏิบัติการวิเคราะห์งานโดยตรงเท่านั้น

โครงการวิเคราะห์งานที่สมบูรณ์แบบ ต้องกำหนดหัวข้อที่จะศึกษาให้ครบถ้วน ในการกำหนดหัวข้อนี้ จะต้องพิจารณาว่า การวิเคราะห์งานตามวัตถุประสงค์หนึ่งๆ นั้น ควรจะครอบคลุมข้อมูลอันใด และรายละเอียดเรื่องใดบ้าง ในการที่จะวางโครงการปฏิบัติการให้ครอบคลุมเรื่องที่จะต้องศึกษาทุกเรื่องนี้ คณะกรรมการวิเคราะห์งานอาจมีบทบาทสำคัญยิ่ง หัวข้อที่จะต้องศึกษามีมากเพียงใดคณะกรรมการที่มาจากฝ่ายต่างๆ จะมี

⁷ Mee, *op. cit.*, p. 135.

ส่วนช่วยได้มากขึ้นเท่านั้น⁸ หัวข้อหลักที่จะศึกษาในการวิเคราะห์งานตามโครงการหนึ่ง ๆ นั้น ควรจะครอบคลุมลักษณะของสถานที่ที่ปฏิบัติงาน อำนาจหน้าที่ และความรับผิดชอบ สภาพการณ์ และองค์ประกอบเนื่องเกี่ยวกับการปฏิบัติงาน ลักษณะนิสัยส่วนบุคคล และคุณสมบัติของพฤติกรรม (traits) ที่จำเป็นในการปฏิบัติงาน ตามอำนาจและความรับผิดชอบดังกล่าว หัวข้อของรายละเอียดเหล่านี้อาจแยกเป็นข้อ ๆ ได้ดังนี้⁹

1. ชื่อตำแหน่งงาน หมายความว่าถึงชื่อเฉพาะที่ใช้เรียกในวงงานนั้น ๆ
2. จำนวนเจ้าหน้าที่ที่ปฏิบัติงานในตำแหน่งนั้น ๆ และความสัมพันธ์ของตำแหน่งนั้นกับหน่วยงาน
3. ชื่อของผู้บังคับบัญชาชั้นต้น
4. วัสดุ ครุภัณฑ์ และอุปกรณ์ที่ต้องใช้ในการปฏิบัติงานในตำแหน่งนั้น ๆ
5. คำสั่งและรายละเอียดในการปฏิบัติงานที่ผู้บังคับบัญชามอบหมาย และที่บุคคลในตำแหน่งนั้นจะต้องปฏิบัติ
6. เงินเดือน หรือค่าจ้าง และอัตราชั่วโมงการทำงาน
7. สภาพการณ์ในการทำงาน
8. รายละเอียดของหน้าที่ทั้งหมด แยกเป็นวัน สัปดาห์ และเดือน รวมทั้ง หน้าที่พิเศษจำแนกตามความนานของเวลาที่ต้องปฏิบัติ
9. คุณสมบัติทางการศึกษา และประสบการณ์ที่ต้องมี
10. ทักษะ ความถนัดทางธรรมชาติ และความสามารถที่ต้องมี
11. โอกาสในการเลื่อนตำแหน่งและโยกย้าย
12. ข้อเท็จจริงทั่วไป และความเห็นประกอบ

นอกจากหัวข้อดังกล่าวนี้แล้ว ในการวิเคราะห์งานตามโครงการหนึ่ง ๆ ยังอาจเพิ่มเติมหัวข้อบางประการก็ได้ เช่น ลักษณะความสมบูรณ์ทางร่างกายที่ต้องมี รายละเอียดของการคัดเลือกบุคคลเข้าปฏิบัติงานในตำแหน่งนั้น ๆ ความยากง่ายของงานในแง่ความริเริ่มและคุณพินิจที่ต้องมี เป็นต้น¹⁰ ตัวอย่างของหัวข้อที่ควรจะต้องครอบคลุมแสดงไว้ในแผนภาพที่ 2

⁸ *Ibid.*, pp. 141–143.

⁹ Jucius, M.J., *Personnel Management*, p. 94.

¹⁰ Mee, *op. cit.*, pp. 144–146.

แบบ ก.พ. 0-09 (2)

สำนักงาน ก.พ.

แบบสอบถามหน้าที่และความรับผิดชอบ

<p>แบบสอบถามนี้ ต้องการข้อเท็จจริง และรายละเอียด เกี่ยวกับหน้าที่ และความรับผิดชอบของตำแหน่ง มิใช่เรื่องราวของตัวบุคคล คำตอบของท่าน จะไม่ทำให้ท่านเสียสิทธิ หรือประโยชน์ที่ท่านมีอยู่แล้วหรือที่จะพึงได้รับแต่อย่างใด</p> <p>โปรดอ่านคำอธิบายแบบสอบถามให้เข้าใจก่อนตอบ</p>	<p>1. ชื่อ</p> <p>2. ข้าราชการพลเรือน</p> <p><input type="checkbox"/> สามัญ</p> <p><input type="checkbox"/> วิสามัญ</p> <p>หรือ <input type="checkbox"/> ลูกจ้างประจำ</p> <p>ชั้น อันดับ ชั้น บาท</p> <p>3. ตำแหน่ง</p> <p>4. ตำแหน่งตามลักษณะงานที่กำลังปฏิบัติอยู่ (ถ้าต่างกับข้อ 3.)</p>	<p>(โปรดเว้นช่องว่างไว้)</p>
<p>5. ชื่อและที่ตั้งสถานที่ปฏิบัติราชการ สำหรับผู้ที่ปฏิบัติราชการในส่วนกลาง ให้กรอกเฉพาะข้อ ก. ส่วนผู้ที่ปฏิบัติราชการในส่วนภูมิภาค ให้กรอกทั้งข้อ ก. และข้อ ข. ด้วย</p> <p>ก. กระทรวง กรม ส่วน</p> <p>กอง แผนก โทรศัพท์</p> <p>ข. ชื่อสถานที่ราชการในส่วนภูมิภาค ตำบล</p> <p>อำเภอ จังหวัด เขต</p> <p>6. หน้าที่ของตำแหน่งที่ท่านปฏิบัติอยู่มีอะไรบ้าง และปฏิบัติอย่างไร โปรดอธิบายโดยละเอียด ด้วยถ้อยคำของท่านเอง เรียงตามลำดับความสำคัญเป็นเรื่องๆ ไป พร้อมทั้งระบุด้วยว่า หน้าที่เหล่านั้นท่านได้ปฏิบัติอยู่เป็นประจำ หรือ เป็นครั้งคราว หรือ เป็นฤดูกาล หนึ่ง ขอได้โปรดประมาณเวลาที่ใช้ในการปฏิบัติหน้าที่ แต่ละเรื่องดังกล่าว ว่าเป็นจำนวนกี่เปอร์เซ็นต์ ของเวลาที่ใช้ในการปฏิบัติราชการทั้งหมด</p> <p>(ข้อนี้เป็นคำถามที่สำคัญที่สุด ก่อนตอบโปรดอ่านคำอธิบายแบบสอบถาม ข้อ 6 หน้า 2 ให้เข้าใจโดยละเอียด)</p>		
		<p>เปอร์เซ็นต์</p>

7. ท่านปฏิบัติหน้าที่ดังกล่าว ในข้อ 6. มาเป็นเวลาปีเดือน และท่านได้ปฏิบัติงานใน สายอาชีพนี้มาเป็นเวลาปีเดือน

ก. จำนวนผู้อยู่ใต้บังคับบัญชาที่อยู่ในความรับผิดชอบของท่านรวมทั้งสิ้นคน ในจำนวนนี้เป็นข้าราชการสามัญและวิสามัญคน ลูกจ้างประจำคน ลูกจ้างชั่วคราวและคนงานอื่นๆคน

ข. จำนวนผู้อยู่ใต้บังคับบัญชาที่รายงานโดยตรงต่อท่านมีดังนี้ คือ

ชั้น	ประเภท (สามัญ วิสามัญ ลูกจ้างและคนงานอื่นๆ)	จำนวน	ตำแหน่ง

วิธีการรวบรวมข้อมูล

การรวบรวมข้อมูลเพื่อวิเคราะห์งานอาจทำได้ 3 วิธี คือ วิธีสังเกตงาน วิธีสัมภาษณ์ และวิธีรวบรวมข้อมูลเป็นลายลักษณ์อักษร การสังเกตและการสัมภาษณ์นั้น มีความหมายชัดเจนตามชื่ออยู่แล้ว ส่วนการรวบรวมข้อมูลเป็นลายลักษณ์อักษรนั้น หมายถึง การทอดแบบสอบถามเป็นวิธีที่สำคัญ และยังอาจรวบรวมข้อมูลที่เป็นลายลักษณ์อักษรได้ในเรื่องเกี่ยวกับหลักฐาน และเอกสารที่ผู้รู้จักงานในตำแหน่งนั้นดีมีอยู่ เนื่องจากวิธีการทั้งสามนี้เป็นวิธีที่ต้องใช้เทคนิคพิเศษทางการวิจัยประกอบ ในที่นี้จึงจะกล่าวถึงเทคนิคของแต่ละวิธีโดยสังเขปเสียก่อน

การสังเกต ในการรวบรวมข้อมูล ตามวิธีนี้นักวิเคราะห์งานไปสังเกตงานด้วยตนเอง การสังเกตย่อมมีลักษณะต่างกับการที่เพียงแต่ดู หรือพิจารณาสิ่งต่างๆ ทั่วไป การสังเกตที่จะช่วยให้สามารถรวบรวมข้อมูลได้ควรมีลักษณะ 5 ประการ ดังต่อไปนี้¹¹

1. การสังเกตต้องมีความมุ่งหมายที่แน่นอน และชัดเจนว่าประสงค์จะทราบสิ่งใดหรือเรื่องใด

2. การสังเกตจะต้องมีการกำหนดระยะเวลาที่จะสังเกต เวลาที่กำหนดนี้จะต้องพิจารณาให้เหมาะสมสอดคล้องกับองค์ประกอบหรือปัจจัยของสิ่ง หรือปรากฏการณ์ที่จะสังเกต การที่เพียงแต่ไปหาข้อเท็จจริงในเวลาใดก็ได้ โดยมิได้พิจารณาเสียก่อนย่อมไม่เป็นการสังเกต

3. การสังเกตจะต้องทำให้สามารถทราบความสัมพันธ์ของปริมาณสถิติต่าง ๆ ที่เกี่ยวข้องกับสิ่งหรือปรากฏการณ์ที่สังเกต การที่เพียงแต่พิจารณาหรือมองดูเฉย ๆ ย่อมไม่เป็นการสังเกต

4. การสังเกตจะต้องมีการบันทึก จะต้องดำเนินการบันทึกรายละเอียดข้อเท็จจริงในทันทีในระหว่างการสังเกต หรือโดยเร็วที่สุดเท่าที่จะทำได้ การสังเกตย่อมไม่อาศัยความทรงจำเป็นหลัก

5. การสังเกตจะต้องมีการทดสอบความเชื่อถือได้ (validity) นักวิเคราะห์งานจะต้องมีความชำนาญ หรือได้รับการฝึกฝนอบรมในการสังเกต จึงจะสามารถสังเกตข้อเท็จจริงที่เชื่อถือได้ ผลของการสังเกตจะต้องสามารถทดสอบซ้ำได้อีก

ถึงแม้ว่าการสังเกตจะทำได้หลายอย่างหลายแนว แต่ก็มีหลักกว้าง ๆ ที่จะยึดถือได้

¹¹ Good, C.V and Scates, D.E., *Methods of Research*, pp. 648-650.

ว่า การสังเกตที่ดีนั้นจะต้องทำอย่างไร ไม่ว่าจะการสังเกตนั้น จะทำโดยอย่างไรหรือโดย
แนวใด หลักของการสังเกตที่ดีมี 7 ประการ คือ¹²

1. การหาความรู้เกี่ยวกับสิ่งที่สังเกตก่อนการสังเกต นักวิเคราะห์ จะต้องพยายาม
ศึกษาให้ทราบล่วงหน้าก่อนว่า ปรากฏการณ์ชนิดใด หรือเรื่องใดบ้างที่มีค่าควรแก่
การสังเกต สภาพการณ์และสิ่งแวดล้อมที่เกี่ยวข้องกับปรากฏการณ์นั้น ๆ จะช่วยให้
นักวิเคราะห์ตัดสินใจได้ว่า ควรจะไปสังเกตข้อเท็จจริงในประเด็นใดบ้างที่เกี่ยวกับ
ปรากฏการณ์นั้น ๆ และในเวลาใด

2. ความสอดคล้องระหว่างรายละเอียดของข้อมูลที่สังเกตกับประเด็นของงาน
ที่จะวิเคราะห์ปัญหาเกี่ยวกับงานที่ได้ตั้งขึ้น ย่อมเป็นสิ่งสำคัญที่จะช่วยให้กำหนดได้ว่า
ควรจะต้องใช้ข้อมูลประเภทใดในเรื่องใด ฉะนั้นนักวิเคราะห์จะต้องตั้งประเด็นที่จะทำ
ให้สามารถได้ข้อมูลประเภทนั้น ๆ หรือเรื่องนั้น ๆ มา การที่เพียงแต่ตั้งใจไว้ว่าจะไป
สังเกตขบวนการ และวิธีการต่าง ๆ ทั้งหมดที่เกี่ยวข้องกับปรากฏการณ์นั้น ๆ ย่อมเป็น
การตัดสินใจที่ผิด เพราะอาจจะเป็นการเสียเวลาไปโดยเปล่าประโยชน์

3. วิธีบันทึกข้อสังเกต การสังเกตจะต้องมีการบันทึกรายละเอียดข้อเท็จจริงที่สังเกต
ได้โดยทันทีหรือโดยเร็วที่สุดเท่าที่จะทำได้ การที่จะบันทึกข้อสังเกตให้ตรงกับข้อเท็จ
จริงที่สังเกตได้นั้น นักวิเคราะห์ควรจะกำหนดวิธีที่จะบันทึกไว้ล่วงหน้า เช่น อาจจะ
เตรียมสารบัญหัวข้อปัญหาที่จะสังเกตไว้

4. ประเภทข้อมูลที่จะได้จากการสังเกต พฤติกรรมต่าง ๆ ที่นักวิเคราะห์หัดหรือ
สังเกตนั้น ควรจะได้มีการตีความหมายแยกประเภท ลักษณะออกไปอีกด้วย นัก
วิเคราะห์อาจจะใช้ตารางสามช่องที่มีรายชื่อหรือรายการของพฤติกรรม หรือขบวนการของ
สิ่งที่สังเกต และอีกสองช่องเพื่อบันทึกปริมาณ และลักษณะของพฤติกรรมที่เกี่ยวข้อง
กับปรากฏการณ์ที่สังเกต นักวิเคราะห์ควรจะพยายามที่จะศึกษาปรากฏการณ์นั้น ๆ
ให้เข้าใจชัดเจนว่า ข้อเท็จจริงที่สังเกตได้นั้นมีลักษณะต่าง ๆ กัน และเกี่ยวข้องกับ
ปรากฏการณ์อย่างไร เพื่อที่จะให้สามารถแยกประเภทของข้อเท็จจริงที่ได้มา

5. ความระมัดระวังและความพินิจพิเคราะห์ นักวิเคราะห์จะต้องระวังว่า ตนได้
เตรียมการที่จะสังเกตมาพร้อมแล้ว และในระหว่างที่ทำการสังเกตก็ตั้งใจที่จะ
พิจารณาประเด็นสำคัญ ๆ ที่ได้กำหนดไว้ล่วงหน้า ความระมัดระวัง และความ

¹² Rummel, J.F., and Ballaine, W.C., *Research Methodology in Business*, pp. 81—89.

พินิจพิจารณาที่นี้จะช่วยทำให้นักวิเคราะห์ได้ข้อเท็จจริงที่ไม่บิดเบือนไปจากความเป็นจริง และมีความแน่นอนเชื่อถือได้ ไม่ใช่เป็นข้อเท็จจริงที่สักแต่ว่าได้เห็นและบันทึกเอาไว้เท่านั้น

6. การแยกข้อเท็จจริงที่ไม่เกี่ยวข้องกันออกจากกัน ในการสังเกตข้อเท็จจริงในเชิงคุณภาพ (qualitative) นักวิเคราะห์จะต้องพยายามพิจารณาข้อเท็จจริงประเภทหนึ่งให้แยกออกจากข้อเท็จจริงอีกประเภทหนึ่งโดยเด็ดขาด นักวิเคราะห์จะต้องพยายามไม่ให้ปริมาณคุณภาพของข้อเท็จจริงประเภทหนึ่งมาทำให้การพิจารณาปริมาณคุณภาพของข้อเท็จจริงอีกประเภทหนึ่งบิดเบือนไป ในการสอบ และวัดข้อเท็จจริงในเชิงคุณภาพนั้น นักวิเคราะห์อาจจะได้สถิติข้อเท็จจริงที่บิดเบือนไป เพราะถูกอิทธิพลของข้อเท็จจริงประเภทอื่น ๆ มาครอบงำ

7. ทักษะในการใช้เครื่องมือบันทึกผล การสังเกตที่ดีจะต้องมีอุปกรณ์และเครื่องมือในการบันทึกผล ฉะนั้น นักวิเคราะห์จึงควรพยายามฝึกฝนอบรมให้เกิดความชำนาญในการใช้เครื่องมือบันทึกผล การสังเกต เพื่อไม่ให้ผลของการสังเกตหรือข้อเท็จจริงที่ได้สังเกตมาบิดเบือนไปเพราะความล่าช้าในการบันทึก โดยเฉพาะอย่างยิ่งควรพยายามฝึกฝนใช้เครื่องมือบันทึกผลให้คล่องแคล่วจริงๆ เสียก่อน จึงจะเริ่มลงมือสังเกต ถ้าหากว่าจะต้องมีการสังเกตกันหลาย ๆ คนพร้อม ๆ กัน ก็ควรจะมีการสัมมนาอภิปราย และทดลองใช้ให้แน่ใจว่าทุกคนจะมีแนวและวิธีการเดียวกันในการบันทึกข้อสังเกต

การสัมภาษณ์ การรวบรวมข้อมูลโดยวิธีสัมภาษณ์หมายถึง การที่นักวิเคราะห์ทำงานไปสัมภาษณ์ผู้รู้จกงานนั้นดี ได้แก่ หัวหน้างาน ผู้บริหาร และผู้ปฏิบัติงานนั้น ๆ โดยตรง ในต่างประเทศบุคคลที่ถือได้ว่ารู้จกงานที่จะวิเคราะห์ดีนอกเหนือไปจากที่กล่าว ยังมีอีก คือ สหบาลกรรมกร ในการสัมภาษณ์ นักวิเคราะห์งานจะต้องตั้งประเด็นความมุ่งหมายที่จะสัมภาษณ์ไว้ให้ชัดเจนแน่นอน ในการสัมภาษณ์ นักวิเคราะห์งานจะต้องมีโครงการที่กำหนดไว้ดีแล้ว โดยพิจารณาถึงประเด็นของข้อมูลที่ประสงค์จะได้ ตลอดจนบุคลิกภาพ และสถานะของบุคคลที่จะถูกสัมภาษณ์¹³

1. พิจารณาตัวผู้ให้สัมภาษณ์ นักวิเคราะห์จะต้องเลือกตัวบุคคลที่จะให้สัมภาษณ์ด้วยความรอบคอบ นักวิเคราะห์จะต้องพิจารณาว่าผู้ใดบ้างที่จะทราบข้อเท็จจริงที่ประสงค์จะได้ แล้วพิจารณาต่อไปอีกว่าผู้ทราบข้อเท็จจริงนั้นมีอำนาจ หรืออยู่ในฐานะที่จะให้ข้อเท็จจริงนั้นหรือเปล่า หรือถ้าอยู่ในฐานะที่ให้ข้อเท็จจริงนั้นจะมีความเต็มใจที่จะให้ข้อเท็จจริงหรือเปล่า ความผิดพลาดที่ร้ายแรงมากสำหรับนักวิเคราะห์งานก็คือ กำหนด

ตัวผู้ที่จะให้สัมภาษณ์ผิด ในการที่จะเลือกตัวผู้ให้สัมภาษณ์ให้ถูกนั้น นักวิเคราะห์จะต้องศึกษาภูมิหลังของผู้ที่จะให้สัมภาษณ์ในแง่ความรับผิดชอบของผู้นั้น ความสัมพันธ์กับหน่วยงานของบุคคลนั้นทั้งในปัจจุบัน และในอดีต ประสบการณ์ในเรื่องนั้น ๆ ของผู้ที่จะให้สัมภาษณ์ อคติของบุคคลนั้นที่อาจจะมีได้ และเรื่องอื่น ๆ ในทำนองนี้ที่จะทำให้น้ำใจว่า ผู้ที่จะให้สัมภาษณ์นั้น จะสามารถ และเต็มใจที่จะให้ข้อเท็จจริงที่ประสงค์ได้แน่นอน นอกจากนี้จะเลือกตัวผู้ให้สัมภาษณ์แล้ว นักวิเคราะห์งานจะต้องกำหนดจำนวนของผู้ที่จะให้สัมภาษณ์ว่า เป็นจำนวนที่เพียงพอที่จะทำให้สามารถทำการลงความเห็น (generalization) เกี่ยวกับข้อเท็จจริงได้

2. เตรียมการล่วงหน้าก่อนการสัมภาษณ์ นักวิเคราะห์งานจะต้องกำหนดเวลาและสถานที่ที่แน่นอนเพื่อการสัมภาษณ์ โดยให้สอดคล้องกับความประสงค์ของผู้ที่จะให้สัมภาษณ์ โดยปกติแล้ว อาจจะทำการสัมภาษณ์ได้ในสถานที่ทำงานของผู้ที่จะให้สัมภาษณ์ เพื่อที่จะให้สภาพแวดล้อมในระหว่างการสัมภาษณ์นั้น มีลักษณะเป็นไปในการทำงานเดียวกับในระหว่างที่มีการปฏิบัติงานจริง ๆ ถ้าหากจะสัมภาษณ์ในที่ทำงาน นักวิเคราะห์ก็จะต้องศึกษาหมายกำหนดการปฏิบัติงานของสถานที่ทำงานนั้น ๆ เพื่อจะได้ไม่ไปสัมภาษณ์ในตอนที่มีงานฉุกเฉิน

3. กำหนดคำถามไว้ล่วงหน้า นักวิเคราะห์งานจะต้องตั้งเป้าหมายว่า จะได้ข้อเท็จจริงในประเด็นใด และในการที่แสวงหาข้อเท็จจริงเหล่านั้นจะต้องสร้างทัศนคติกับผู้ให้สัมภาษณ์ในการทำงานใด ควรจะดำเนินการอย่างไรเพื่อเป็นการกระตุ้นให้ผู้ให้สัมภาษณ์เกิดความกระตือรือร้นที่จะให้สัมภาษณ์ การที่จะทราบเรื่องเหล่านี้ ก็ด้วยเหตุที่นักวิเคราะห์มีความเข้าใจแจ่มแจ้งในประเด็นปลีกย่อยทั้งหมดของงานที่จะศึกษา

4. ทดลองสัมภาษณ์ก่อนสัมภาษณ์จริง นักวิเคราะห์งานอาจจะลองสัมภาษณ์กับเพื่อนร่วมงาน หรือบุคคลอื่นที่ไม่ได้กำหนดไว้ว่าจะเป็นผู้ให้สัมภาษณ์ การทดลองสัมภาษณ์กับบุคคลเหล่านี้ จะต้องเลือกบุคคลที่มีลักษณะคล้ายคลึงกับผู้จะให้สัมภาษณ์ ซึ่งจะช่วยให้นักวิเคราะห์มองเห็นว่าแนวที่จะถามผู้ให้สัมภาษณ์นั้นเหมาะสมหรือควรแก้ไขอย่างไรหรือไม่ ในระหว่างการทดลองการสัมภาษณ์ นักวิเคราะห์อาจจะสังเกตวิธีการตอบของผู้ที่ถูกทดลองสัมภาษณ์ ทำให้นักวิเคราะห์มองเห็นปัญหาหรือความยุ่งยาก หรือความไม่เข้าใจซึ่งอาจจะเกิดขึ้นในการสัมภาษณ์จริง แล้วพยายามหลีกเลี่ยงสิ่งเหล่านั้นเสีย

¹³ Hillway, Tyrus, *Introduction to Research*, pp. 199–201.
Rummel and Ballaine, *op. cit.*, pp. 86–104.

5. หากนักวิเคราะห์งานจำเป็นต้องใช้บุคคลอื่นในขณะช่วยสัมภาษณ์ ก็ควรจะต้องฝึกฝนบุคคลเหล่านั้น ให้มีความคล่องแคล่วในการใช้เทคนิคการสัมภาษณ์ ผู้สัมภาษณ์มีหน้าที่จะต้องแสวงหาข้อเท็จจริงจากผู้ให้สัมภาษณ์ ฉะนั้นถ้าหากว่าผู้ให้สัมภาษณ์ไม่ให้ความร่วมมือ ก็อาจจะทำให้การสัมภาษณ์ที่คาดไว้เสียผลไปก็ได้ ฉะนั้น นักวิเคราะห์จึงควรจะต้องฝึกอบรมผู้ช่วยที่จะสัมภาษณ์ให้มีความสามารถ 3 ประการดังต่อไปนี้ ประการแรก ให้สามารถสร้างบรรยากาศที่เป็นกันเองกับผู้ให้สัมภาษณ์ ประการที่สอง สามารถถามคำถามที่มีความหมายชัดเจน และประการสุดท้ายสามารถเข้าใจคำตอบและกิริยาอาการในระหว่างการตอบของผู้ให้สัมภาษณ์ ในการที่จะทำให้เกิดบรรยากาศที่เป็นกันเองนั้น ผู้สัมภาษณ์จะต้องมีความสามารถที่จะทำให้ผู้ให้สัมภาษณ์เกิดความเชื่อถือในตัวผู้สัมภาษณ์ จนมีความแน่ใจว่าผู้สัมภาษณ์นั้นมีความจริงใจ และมีเหตุมีผลสมควรที่มาสัมภาษณ์ ผู้สัมภาษณ์จะต้องมีศิลปะในการสนทนา ทำให้ผู้ให้สัมภาษณ์ไม่รู้สึกเครียดในระหว่างการตอบคำถาม ในรายที่ผู้ให้สัมภาษณ์ต้องการเวลาในการนึกคิด คำตอบผู้สัมภาษณ์ก็รู้จักผ่อนสั้นผ่อนยาว ไม่รีบเข้เข้าคำตอบผู้สัมภาษณ์ ไม่ทำให้ผู้ให้สัมภาษณ์มีความรู้สึกเป็นจำเลยที่ต้องตอบข้อซักถาม

6. สอบความเที่ยงตรงและความแน่นอน (accuracy and reliability) ของข้อเท็จจริงที่ได้มา นักวิเคราะห์งานจะต้องนี้อยู่เสมอว่าคำตอบที่ได้มานั้นอาจจะผิดพลาดก็ได้ ทั้งนี้เพราะผู้ให้สัมภาษณ์ได้ยินได้ฟังมาผิด ๆ หรือความจำของผู้ให้สัมภาษณ์เลือนไป ผู้ให้สัมภาษณ์บางคนอาจจะอยากเน้นหนักความสุขหรือความทุกข์ที่เคยได้รับมากกว่าความจริง ฉะนั้น นักวิเคราะห์งานจะต้องตั้งอกตั้งใจคอยสังเกตว่าคำตอบมีแนวโน้มไปในทำนองนี้หรือเปล่า หากปฏิกิริยา สีหน้าและน้ำเสียงของผู้ให้สัมภาษณ์อาจจะเป็นสิ่งที่น่าสนใจที่นักวิเคราะห์สังเกตเพื่อคอยจับดูว่า คำตอบนั้นมีความหมายเกินเลยความจริงไปหรือเปล่า

7. บันทึกผลการสัมภาษณ์ในทันทีเท่าที่จะทำได้ คำตอบของผู้ให้สัมภาษณ์ควรจะจดบันทึกไว้ทันทีเมื่อสัมภาษณ์เสร็จ หรือหลังจากนั้นโดยด่วนเท่าที่เวลาจะอำนวย ในระหว่างการสัมภาษณ์ นักวิเคราะห์อาจจะจดบันทึกข้อความสั้น ๆ และนำข้อความที่บันทึกไว้ไปเขียนให้เต็มเมื่อสัมภาษณ์เสร็จ

การทอดแบบสอบถาม นักวิเคราะห์อาจจะออกแบบสอบถามเพื่อส่งให้แก่เจ้าหน้าที่ผู้ปฏิบัติงานโดยตรงในการรวบรวมข้อมูลก็ได้ ในการทอดแบบสอบถามนี้ ควรใช้กับผู้ปฏิบัติงานที่อ่านและเขียนหนังสือคล่องพอสมควร แบบสอบถาม มีลักษณะเป็น

รายการของคำถาม คำถามที่ปรากฏในแบบสอบถามจะต้องกำหนดให้มีแนวที่จะถาม เพื่อให้ได้ความจริงบางอย่าง ความจริงนี้อาจจะเกี่ยวกับสถานการณ์ แนวปฏิบัติ หรือความเห็นของผู้รับแบบสอบถาม ซึ่งนักวิเคราะห์งานคาดว่าผู้รับแบบสอบถามจะรู้เรื่องพอที่จะตอบแบบสอบถามได้

การตั้งคำถาม ตลอดจนการจัดเรียงคำถามและการวางรูปของแบบสอบถาม เป็นสิ่งสำคัญที่จะช่วยให้ผู้รับแบบสอบถามให้ความร่วมมือในการตอบ รูปของแบบสอบถาม ควรจะจัดวางให้มีลักษณะพิเศษเป็นที่ดึงดูดสายตาของผู้รับแบบสอบถาม และให้คู่สะดวกที่จะอ่าน เพื่อที่จะให้ผู้รับแบบสอบถามไม่ลำบากในการที่จะอ่านคำถามหนึ่ง ๆ และตอบคำถามนั้น ๆ แบบสอบถามที่มีรูปร่างดี ย่อมทำให้ผู้รับแบบสอบถามเห็นความสำคัญและความมีคุณค่าทางวิชาการของเรื่องที่สอบถาม

การสร้างคำถามเป็นงานที่สำคัญสำหรับนักวิเคราะห์ เพื่อช่วยให้การประมวลข้อมูลขั้นต่อไปสะดวกขึ้น ในการตั้งคำถาม นักวิเคราะห์ควรจะตั้งคำถามถามตัวเองว่า ทำไมผู้รับแบบสอบถามจึงจะต้องตอบคำถามที่ตนถามไป ผู้รับแบบสอบถามจะทราบข้อเท็จจริงที่ถามหรือไม่ นักวิเคราะห์ได้อธิบายเหตุผลเป็นอย่างดีแล้วหรือไม่ ว่าแบบสอบถามนี้ควรจะได้รับความตอบ คำถามที่ตั้งนั้นจะเป็นคำถามที่ให้ข้อเท็จจริง ซึ่งจะช่วยให้ความหมายของประเด็นต่าง ๆ ของลักษณะงานที่ศึกษาหรือไม่ ถ้าหากว่านักวิเคราะห์ตอบคำถามเหล่านี้ว่าใช่ จึงควรจะถือว่าคำถามที่จะสร้างนั้นเป็นคำถามที่จะเป็นประโยชน์กับการวิเคราะห์งาน

คำถามแบบปรนัย เป็นคำถามที่แพร่หลายมาก เพราะจะช่วยให้ได้คำตอบที่ไม่บิดเบือนจากความเป็นจริง การสร้างคำถามที่จะช่วยให้ได้คำตอบที่ไม่บิดเบือนความจริงนี้เป็นหลักสำคัญมาก การที่จะได้คำถามที่มีคุณสมบัติเช่นนี้ นักวิเคราะห์จะต้องยึดหลัก 15 ประการดังต่อไปนี้¹⁴

1. ข้อความ คำถาม และคำตอบที่จะให้ผู้รับแบบสอบถามเลือก จะต้องมีความหมายชัดเจนและมีใจความสั้น ๆ ไม่ยืดเยื้อ
2. ถ้อยคำที่ใช้ในคำถามจะต้องเป็นถ้อยคำที่ไม่มีความหมายกำกวม
3. ควรหลีกเลี่ยงการเรียบเรียงประโยคที่สลับซับซ้อน
4. คำถามบางข้อจะต้องมีคำอธิบายประกอบเพื่อให้ผู้รับแบบสอบถามมีแนวที่จะพิจารณาตัดสินใจในการตอบ
5. ตัดคำที่ฟุ่มเฟือยหรือไม่จำเป็นทิ้ง

¹⁴ Ferber, Robert, and Verdoorn, P.J., *Research Methods in Economics and Business*, pp. 213–228. Rummel and Ballaine, *op. cit.*, pp. 120–122.

6. จะต้องไม่ตั้งคำถามที่จะเป็นการนำผู้ตอบให้ตอบตามแนวหนึ่งแนวใด
7. ตัดคำถามที่มีความสำคัญน้อยออกเสีย
8. คำตอบที่กำหนดไว้ให้ผู้รับแบบสอบถามเลือกจะต้องเป็นที่เข้าใจได้ง่าย
9. ใจความของคำถามจะต้องทำให้ผู้รับแบบสอบถามรู้สึกว่าจะเกี่ยวข้องกับสถานการณ์หรือข้อเท็จจริงที่จะถาม
10. หลีกเลี่ยงการถามความเห็น เว้นแต่ในกรณีที่ยากจะได้แต่เพียงความเห็นโดยเฉพาะ
11. เลือกคำถามที่จะไม่ทำให้ผู้รับแบบสอบถามต้องเกิดความลำบากใจ หรืออึดอัดใจที่จะตอบ
12. หลีกเลี่ยงคำถามที่จะตอบได้หลาย ๆ แนว โดยนักวิเคราะห์ประสงค์จะได้แต่เพียงคำตอบในแนวเดียวเท่านั้น
13. พยายามที่จะให้ผู้รับแบบสอบถามต้องลำบากน้อยที่สุดในการตอบ เช่น อาจจะเพียงให้ผู้รับแบบสอบถามกาเครื่องหมายลงบนแบบสอบถามเท่านั้น
14. ตั้งคำถามที่จะทำให้ผู้รับแบบสอบถามไม่ต้องคิดมากเกินไป ในกรณีที่จะต้องถามเรื่องที่ทำให้ผู้รับแบบสอบถามต้องคิดมาก นักวิเคราะห์อาจจะแยกออกเป็นคำถามย่อย ๆ หลาย ๆ คำถาม เพื่อให้สะดวกแก่การคิดตอบ
15. หลีกเลี่ยงการใช้คำที่เป็นนามธรรม เช่น ยาก ง่าย ฉลาด โง่ มีศีลธรรม ฯลฯ เพราะจะทำให้คำตอบขาดความแน่นอน เพราะคนหนึ่ง ๆ ย่อมมีแนวความคิดในเรื่องนามธรรมในลักษณะต่าง ๆ กัน

ลักษณะของข้อมูลที่เป็นประโยชน์ต่อเครื่องมือวัดผล

โครงการวิเคราะห์งานหนึ่ง ๆ ที่ดำเนินการรวบรวมข้อมูลตามวิธีใดวิธีหนึ่งข้างต้นนี้ จะให้ข้อมูลที่แยกประเภทใหญ่ ๆ ได้ 4 ประเภท คือ ประเภทแรก ลักษณะทั่วไปเกี่ยวกับงาน ประเภทที่สอง ข้อเท็จจริงเกี่ยวกับคุณสมบัติของพฤติกรรมที่จำเป็นต้องมีการปฏิบัติงาน ประเภทที่สาม รายละเอียดของขั้นตอนในการปฏิบัติ (operation information) และประเภทสุดท้ายข้อเท็จจริงเกี่ยวกับกิจกรรมที่จะต้องทำ (activity information) หลักของการวิเคราะห์งานตามที่อ้างไว้ในบางแห่ง อาจย้ําแต่เพียงข้อมูล 3 ประเภทแรก¹⁵ การวิเคราะห์งานที่มุ่งจะได้ข้อมูลเพียงบางอย่างเช่นนี้จะไม่สนองความต้องการในการสร้างเครื่องมือวัดผล ข้อมูลประเภทที่จะมีประโยชน์มากที่สุดต่อการสร้างเครื่องมือวัด คือ ข้อมูลประเภทที่สี่

¹⁵ Stahl, O.G., *Public Personnel Administration*, p. 154.

ลักษณะทั่วไปเกี่ยวกับงาน ได้แก่ ข้อเท็จจริงทั่ว ๆ ไปที่อธิบายงานที่วิเคราะห์นั้น ๆ เช่น ชื่อตำแหน่ง หรือคำอธิบายงานสั้น ๆ ลักษณะทั่วไปที่ได้จากการวิเคราะห์งานในบางโครงการอาจจะจุดเด่นของงาน แต่ในบางครั้งก็อาจเพียงบอกแต่ว่า เป็นงานในประเภทฝีมือ หรือประเภทวิชาชีพเท่านั้น ข้อมูลเช่นนี้ย่อมไม่ช่วยในการที่จะนำไปเทียบวิเคราะห์กับเครื่องมือวัดที่สร้างขึ้น อาจใช้ได้แต่เพียงประกอบในการคัดเลือกเครื่องมือวัดบางอันว่า จะมีลักษณะใกล้เคียงกับงานในตำแหน่งที่จะวัดหรือไม่เท่านั้น ส่วนข้อเท็จจริงเกี่ยวกับคุณสมบัติของพฤติกรรมที่จำเป็นต้องมีในการปฏิบัติงานนั้น เป็นผลการวิเคราะห์งานที่ชี้ให้เห็นว่า งานในตำแหน่งหนึ่ง ๆ จะต้องมีคุณสมบัติเช่นใดบ้าง คุณสมบัติที่นักวิเคราะห์งานสามารถศึกษาได้ในแง่ของความสมบูรณ์ของร่างกายที่จะต้องมีในการปฏิบัติงาน อาจใช้ประกอบในการทดลองสร้างเครื่องมือวัดได้ แต่เครื่องมือวัดที่ทดลองสร้างขึ้น ยังต้องอาศัยคุณสมบัติและความสามารถทางสมองและนามธรรมอีกด้วย ความสามารถประเภทหลังนี้ ถึงแม้ว่านักวิเคราะห์งานสามารถชี้ให้เห็นได้ในผลการวิเคราะห์ แต่ก็นับได้ว่าเป็นข้อเท็จจริงที่ขึ้นอยู่กับดุลพินิจของนักวิเคราะห์งานเอง¹⁶ ฉะนั้น ในการที่จะนำผลการวิเคราะห์งานในประเด็นนี้มาประกอบในการทดลองสร้างเครื่องมือวัด ก็จำเป็นที่จะต้องถือว่าข้อเท็จจริงนี้ยังนับว่าเป็นที่สุดไม่ได้ ในเรื่องรายละเอียดของขั้นตอนในการปฏิบัติที่นักวิเคราะห์งานสามารถแสดงได้ในผลของการศึกษา ก็อาจถือได้ว่าเป็นรายละเอียดมากเกินไปที่จะเป็นในการนำมาประกอบการทดลองสร้างเครื่องมือวัด เหตุผลในเรื่องนี้จะเห็นได้ เมื่อพิจารณารายละเอียดของขั้นตอนการสร้างเครื่องมือวัดในบทต่อไป ผู้ศึกษาเรื่องการวัดผลจะเห็นได้ว่า รายละเอียดของขั้นตอนการปฏิบัติงานที่แยกแยะออกไปจนถึงขั้นที่ว่า บุคคลต้องนั่งหรือยืนอยู่ที่ใด ต้องเคลื่อนไหวอย่างไรบ้างนั้น เป็นสิ่งที่ไม่จำเป็นต่อการสร้างเครื่องมือวัดผลบุคคล ยกเว้นแต่ในกรณีที่สร้างเครื่องมือวัดบุคคลในตำแหน่งงานที่ต้องใช้ความแคล่วคล่องของอวัยวะต่าง ๆ ของร่างกายโดยเฉพาะเท่านั้น

ข้อเท็จจริงเกี่ยวกับกิจกรรมที่จะต้องทำ เป็นข้อมูลที่จะเป็นประโยชน์ต่อผู้สร้างเครื่องมือวัดผลมาก ข้อเท็จจริงประเภทนี้ หากนักวิเคราะห์งานแสดงได้ในรายงานการศึกษาของตน ก็จะช่วยให้ทราบว่า ในตำแหน่งงานหนึ่ง ๆ ผู้ปฏิบัติจะต้องมีกิจกรรมอันใดบ้าง กิจกรรมที่แสดงไว้ในผลการวิเคราะห์งานนี้ถือได้ว่า มีลักษณะเป็นปรนัยมากกว่าข้อเท็จจริงเกี่ยวกับคุณสมบัติของพฤติกรรมที่นักวิเคราะห์งานระบุว่า จำเป็นต้องมีในการปฏิบัติงาน เช่น นักวิเคราะห์งานอาจชี้ให้เห็นว่า งานในตำแหน่งหนึ่ง จะ

¹⁶ French, Wendell, *The Personnel Management Process*, pp. 262—263.

ต้องมีการร่างหนังสือและบันทึกโต้ตอบตามระเบียบของทางราชการ หรือเขียนอธิบายว่า ในการปฏิบัติงานอีกตำแหน่งหนึ่ง จะต้องมีการเก็บสถิติเกี่ยวกับคนเข้าเมืองที่มีลักษณะตามที่ระบุไว้ในกฎของหน่วยงาน กิจกรรมของตำแหน่งงานที่วิเคราะห์ได้เช่นนี้ ย่อมเป็นประโยชน์โดยตรงต่อการนำไปสร้างมาตรฐานเปรียบเทียบที่เรียกว่า มาตรฐานวัดความสำเร็จของงาน ซึ่งจะมีรายละเอียดกล่าวไว้ในบทต่อไป

ข้อมูลจากการวิเคราะห์งานที่จะใช้ประกอบการทดลองสร้างเครื่องมือวัด จะต้องมีความเชื่อมั่น (reliability) หมายความว่า ผลการวิเคราะห์งานในตำแหน่งหนึ่งๆ หากเปลี่ยนตัวนักวิเคราะห์งาน จะต้องได้ผลการวิเคราะห์เหมือนกัน ข้อมูลบางอย่างจากการวิเคราะห์งาน เมื่อเปลี่ยนตัวนักวิเคราะห์งานแล้ว ไม่แน่นอนพอไปว่าจะได้ผลการวิเคราะห์เหมือนเดิม¹⁷ ฉะนั้น จึงได้กล่าวแต่ข้างต้นว่า ข้อมูลประเภทที่สี่ เป็นประเภทที่ดีที่สุดในการประกอบการทดลองสร้างเครื่องมือวัดผล ในการวิเคราะห์งานเพื่อให้ได้ข้อมูลประเภทที่สี่นี้ ผู้สร้างเครื่องมือวัดและนักวิเคราะห์งานอาจร่วมกันร่างรายการตรวจสอบ (checklist) เพื่อวิเคราะห์ให้ได้ว่า ในการปฏิบัติงานตำแหน่งหนึ่ง ๆ ต้องทำกิจกรรมอันใดบ้าง เฉพาะกิจกรรมที่สามารถสังเกตเห็นได้ย่อมไม่มีปัญหา แต่กิจกรรมที่เป็นนามธรรมนั้น ผู้ร่างรายการตรวจสอบจะต้องมีความระมัดระวังเป็นพิเศษเพื่อให้แน่ใจว่า ผู้ปฏิบัติงานในตำแหน่งนั้น ๆ มีกิจกรรมเช่นนั้นจริง เช่น กิจกรรมในเรื่องเกี่ยวกับการใช้ความคิดหาเหตุผลแก้ปัญหาของผู้ปฏิบัติงานในตำแหน่งหัวหน้างาน ได้มีผู้ทดลองร่างรายการตรวจสอบขึ้น ดังแสดงไว้ในแผนภาพที่ 3

จากแผนภาพที่ 3 จะเห็นได้ว่า ในการวิเคราะห์งานเพื่อให้ทราบกิจกรรมที่ปฏิบัติ โดยไม่ได้ใช้วิธีต่าง ๆ ของร่างกายซึ่งมองเห็นได้ หากสร้างบัญชีรายการเช่นนั้นขึ้น จะทำให้นักวิเคราะห์งานสามารถวิเคราะห์กิจกรรมที่ปฏิบัติในตำแหน่งงานนั้น ๆ ได้โดยมีความเชื่อมั่นสูงขึ้น รายการตรวจสอบที่ยกมาเป็นตัวอย่างในแผนภาพนี้ เป็นผลของการวิจัยซึ่งได้ทดลองสร้างรายการตรวจสอบเพื่อวัดกิจกรรมบางอย่างในตำแหน่งงานต่าง ๆ ผู้วิจัยได้รายงานผลว่า เมื่อทดลองคำนวณความเชื่อมั่นของการที่นักวิเคราะห์งาน 4 คน ใช้รายการตรวจสอบนี้เพื่อวิเคราะห์งานในตำแหน่งเดียวกัน ปรากฏว่า มีความเชื่อมั่นสูงถึง 0.80¹⁸ (ความหมายของค่าความเชื่อมั่นมีอธิบายรายละเอียดไว้ในบทที่ 4)

¹⁷ Lawshe, C.H., and Balma, M.J., *Principles of Personnel Testing*, p. 295.

¹⁸ *Ibid.*, p. 294.

การวิเคราะห์งานกับคำอธิบายงาน และข้อกำหนดของพนักงาน

การวิเคราะห์งานเป็นการรวบรวมข้อมูลเกี่ยวกับตำแหน่งงาน ข้อมูลที่ได้นี้จะอยู่ในกระดาษร่าง ซึ่งอาจเป็นแบบสอบถามที่กรอกแล้ว หรือผลบันทึกการสัมภาษณ์ เมื่อนักวิเคราะห์งานได้ประมวลข้อมูลต่าง ๆ ทั้งหมดนี้ได้แล้ว จึงจะเขียนลงในรายงานการวิเคราะห์งาน (job analysis report of job analysis schedule) ลักษณะของรายงานนี้

	ลักษณะประจำกิจกรรมในระดับต่าง ๆ
1	ปฏิบัติงานในสภาพการณ์ที่เหมือนกัน และไม่มีการเปลี่ยนแปลง และต้องใช้สามัญสำนึกทำความเข้าใจสภาพการณ์ดังกล่าว เพื่อปฏิบัติตนให้เข้ากับสภาพการณ์นั้น ๆ ได้ หรือปฏิบัติงานตามคำสั่งง่าย ๆ ที่ผู้บังคับบัญชาอธิบายแล้วได้
2	ปฏิบัติงานในสภาพการณ์ที่เหมือนกันทั่วไป ซึ่งอาจจะมีสิ่งที่เปลี่ยนแปลงบ้างบางอย่างเป็นครั้งคราว (แต่ไม่ใคร่มี) และต้องใช้สามัญสำนึกทำความเข้าใจสภาพการณ์ดังกล่าว เพื่อปฏิบัติตนให้เข้ากับสภาพการณ์นั้น ๆ ได้ หรือ ปฏิบัติงานตามคำสั่งที่ตนต้องติดต่อจากผู้บังคับบัญชาไปอีกหนึ่งชั้น
3	เผชิญปัญหา ซึ่งอยู่ในสภาพการณ์ที่มีลักษณะเปลี่ยนแปลงบางอย่าง แต่เห็นได้ชัดแล้ว ต้องใช้สามัญสำนึกทำความเข้าใจในการแก้ปัญหาเหล่านั้น ๆ หรือปฏิบัติงานตามคำสั่งที่มีรายละเอียด
4	เผชิญปัญหาซึ่งอยู่ในสภาพการณ์ที่มีความเปลี่ยนแปลงหลายอย่าง แต่เห็นได้ชัดและต้องใช้หลัก หรือวิธีในการหาเหตุผลเพื่อแก้ปัญหาเหล่านั้น ๆ หรือ ต้องตีความของคำสั่งต่าง ๆ ที่เกี่ยวข้องกับเรื่องนั้น
5	เผชิญปัญหาซึ่งอยู่ในสภาพการณ์ที่มีความเปลี่ยนแปลงหลายอย่าง ความเปลี่ยนแปลงนี้อาจเห็นได้ชัดบ้างในบางครั้ง หรือบางครั้งก็เป็นนามธรรม และต้องใช้หลักและวิชาการในการจำกัดความหมายของปัญหา รวบรวมข้อเท็จจริงตลอดจนพิจารณาสรุปหาเหตุผลด้วยตนเอง ทั้งยังต้องตีความหมายของหลัก และระเบียบที่ปรากฏอยู่ในเอกสารการปฏิบัติงานต่าง ๆ

แผนภาพที่ 3 รายการตรวจสอบกิจกรรมการใช้เหตุผลในการแก้ปัญหาเพื่อปฏิบัติงาน *

* แหล่งที่มา : Lawshe, C.H., and Balma, M.J., *Principles of Personnel Testing*, p. 294.

แสดงตัวอย่างไว้ในแผนภาพที่ 4 เมื่อมีรายงานนี้แล้ว นักวิเคราะห์งานอาจนำสาระสำคัญทั้งหมดมาเรียบเรียงเพื่อให้เห็นว่า ตำแหน่งงานนั้นมีลักษณะสำคัญประการใดบ้าง ทั้งนี้เพราะรายงานการวิเคราะห์งานมีรายละเอียดมากมายเกินกว่าที่จะนำไปใช้ได้ ในทันที สาระสำคัญที่นำมาเขียนใหม่นี้ เรียกว่า คำอธิบายงาน (job description)¹⁹ คำอธิบายงานนี้จะมีรายละเอียดอย่างใดนั้นขึ้นอยู่กับความมุ่งหมายของการวิเคราะห์งานที่กำหนดไว้แต่เดิม ถ้าหากว่าหน่วยงานนั้น ๆ มีวัตถุประสงค์ที่จะวิเคราะห์งานเพื่อประโยชน์ในการปรับปรุงการบรรจุ เลื่อนขั้น หรือโยกย้ายตำแหน่ง ก็จำเป็นที่จะต้องพิจารณาองค์ประกอบเฉพาะของงานอีกครั้งหนึ่ง องค์ประกอบเฉพาะของงานที่นำมาเขียนขึ้นใหม่นี้ เรียกว่า ข้อกำหนดของคนทำงาน (job specification)²⁰ ข้อกำหนดของคนทำงานนี้จะช่วยในการที่จะให้ทราบว่าบุคคลที่ควรจะมีปฏิบัติงานในตำแหน่งหนึ่ง ๆ นั้น ต้องมีคุณสมบัติอันใดแน่นอนบ้าง

คำอธิบายงาน เนื่องจากคำอธิบายงานบรรจุสาระสำคัญทั้งหมดของตำแหน่งงานที่วิเคราะห์ไว้แล้ว ฉะนั้น ในการพิจารณาผลการศึกษานักวิเคราะห์งานจึงใช้คำอธิบายงานเป็นหลักมากกว่าที่จะใช้รายงานการวิเคราะห์งาน แต่สาระสำคัญที่แสดงไว้ในคำอธิบายงานนั้นมีความเน้นหนักเฉพาะในเรื่องที่นักวิเคราะห์งานได้ตั้งความมุ่งหมายไว้แต่ต้นเท่านั้น ดังนั้นก่อนที่จะเขียนคำอธิบายงาน จึงควรจะมีการพิจารณาให้แน่ชัดลงไปว่า หัวข้อใดบ้างที่วิเคราะห์มา มีความสำคัญตรงกับเจตนารมณ์ที่มีการวิเคราะห์งาน ลักษณะของคำอธิบายงานมีแสดงไว้ในแผนภาพที่ 5

โดยหลักกว้าง ๆ แล้ว ในการเขียนคำอธิบายงาน ควรจะถือมาตรฐานดังต่อไปนี้²¹

1. คำอธิบายงานนั้นจะบรรจุลักษณะสำคัญ ๆ ของตำแหน่งงานนั้น ๆ พอที่จะนำคำอธิบายงานนั้นไปใช้ประโยชน์ได้หรือไม่
2. หัวข้อที่สำคัญเกี่ยวกับลักษณะของตำแหน่งงานนั้น ๆ มีครบถ้วนหรือไม่
3. จำนวนและถ้อยคำที่ใช้มีลักษณะคงเส้นคงวา อันจะทำให้สื่อของคำอธิบายงานเป็นแบบเดียวกันหรือไม่
4. ถ้อยคำจำนวนที่ใช้ชัดเจนและกะทัดรัดหรือไม่

¹⁹ Amerine, Ritchey, and Hulley, *op. cit.*, pp. 410—411.

²⁰ Jucius, *op. cit.*, pp. 93—94.

Pigors, Paul, and Myers, C.A., *Personnel Administration*, pp. 374—375.

²¹ Mee, *op. cit.*, pp. 152—155.

French, *op. cit.*, p. 89.

กรม, กระทรวง.....(ชื่อหน่วยงานที่ทำการวิเคราะห์)

- | | |
|---|--|
| 1. ชื่อตำแหน่ง..... | 2. ลำดับที่ของงาน..... |
| 3. จำนวนผู้ทำงาน..... | 4. ลำดับที่ของหน่วยงาน..... |
| 5. ชื่อตำแหน่งที่เรียกเป็นอย่างอื่น | 6. วันที่..... |
| | จำนวนหน้าของคำอธิบายงาน.... |
| 7. ชื่อตำแหน่งงานที่อาจเรียกเป็นอย่างอื่น.... | 8. ประเภทงานของหน่วยงาน..... |
| 9. รหัสของงานที่มีในบัญชีสำรวจ (ถ้ามี)..... | 10. สาขาย่อยของประเภทงานของหน่วยงาน..... |
| 11. ชื่อนักวิเคราะห์งาน..... | |

12. รายงานย่อของตำแหน่งงาน

.....

ฯลฯ

คุณสมบัติอย่างต่ำที่ต้องมีก่อนเข้าทำงาน

- | | |
|---|------------|
| 13. เพศ..... | อายุ |
| 14. ความสมบูรณ์ของร่างกาย (รวมทั้งส่วนสูงและน้ำหนัก)..... | |
| 15. วุฒิกการศึกษา..... | |
| 16. ประสบการณ์..... | |

17. ความสัมพันธ์กับตำแหน่งงานอื่น

(เช่น ต้องเลื่อนจากตำแหน่งใด และอาจเลื่อนไปตำแหน่งใดได้)

.....

.....

18. ลักษณะการบังคับบัญชา ปกติ ปานกลาง อย่างไม่ใกล้ชิด
- ตำแหน่งผู้บังคับบัญชา.....

19. ลักษณะของการควบคุมผู้ใต้บังคับบัญชา มี ไม่มี
 จำนวนผู้ผู้ใต้บังคับบัญชา.....
 ตำแหน่งผู้ผู้ใต้บังคับบัญชา.....
20. ระยะเวลาที่มีงานมากเป็นพิเศษ.....

นอกจากนี้ยังต้องมีใบประกอบเพิ่มเติมซึ่งรายงานลักษณะงานที่ต้องทำ วัสดุ อุปกรณ์ที่ใช้ สภาพการณ์ทำงาน การเสี่ยงอันตราย ลักษณะพิเศษของตำแหน่งงานนั้นๆ ถ้ามีคำจำกัดความของคำที่ใช้เฉพาะในตำแหน่งงาน และความเห็นประกอบ

แผนภาพที่ 4 ตัวอย่างแบบฟอร์มรายงานการวิเคราะห์งาน*

* แหล่งที่มา : Mee, J.F., *Personnel Handbook*, pp. 129–131.

5. ตำแหน่งงานที่อธิบายนั้นได้ความหมายและเขียนเป็นแบบเดียวกันหรือไม่
6. การอธิบายลักษณะและประเด็นต่างๆ ของตำแหน่งงานต่างๆ ในเรื่องเดียวกันเขียนเป็นแบบเดียวกันหรือไม่
7. ตำแหน่งงานนั้นอธิบายไว้ชัดเจนพอที่จะทำให้ทราบว่าสัมพันธ์กับตำแหน่งงานอื่นอย่างไรหรือไม่

นักบัญชีระดับ 1

1. ความรับผิดชอบ ตำแหน่งในระดับนี้หมายถึง ผู้ที่รับผิดชอบงานทางด้านบัญชีโดยเฉพาะ ซึ่งต้องอาศัยผู้ที่มีความรู้ทางด้านบัญชีในระดับปริญญา ปฏิบัติงานในฐานะนักวิชาการที่ไม่มีอำนาจในการปกครองบังคับบัญชา และหมายรวมถึงผู้ที่รับผิดชอบงานในการตรวจสอบความถูกต้องทางการบัญชี หรือหัวหน้าหน่วยทางการบัญชี ซึ่งปกครองบังคับบัญชาเสมือน ลูกจ้าง หรือพนักงานการเงินและบัญชีในระดับต่ำกว่าจำนวน 1 – 10 คน
2. หน้าที่ ภายใต้อำนาจการปกครองบังคับบัญชาโดยทั่วไป มีหน้าที่ควบคุมและจัดทำบัญชีทุกหมวดตามที่กระทรวงการคลังกำหนดจัดทำบัญชีของสำนักงาน หรือโครงการต่างๆ ตรวจสอบความถูกต้องของบัญชี รายงานการเงินทั้งทางการลงบัญชีและ

ตัวเลข โดยตรวจสอบเอกสารหลักฐานการลงบัญชี หลักฐานการเบิกจ่ายเงิน ยอดเงินคงเหลือ ฯลฯ ว่าถูกต้องหรือไม่ ให้คำแนะนำแก้ไขและปรับปรุงบัญชีที่ผิดพลาดหรือทักท้วงให้ได้รับการแก้ไขให้ถูกต้อง จัดทำรายงานบัญชี รายงานการเงินหรือรายงานการตรวจสอบ และสรุปผลการตรวจสอบบัญชี เพื่อส่งหน่วยงานที่เกี่ยวข้อง ศึกษา สํารวจ ออกตรวจเพื่อเสนอการตัดแปลงวางรูปบัญชีในส่วนราชการที่สังกัด แนะนำและแก้ไขปัญหาต่าง ๆ เกี่ยวกับบัญชี เพื่อให้เป็นไปตามมาตรฐานการบัญชีและตามกฎระเบียบต่าง ๆ ที่เกี่ยวข้อง และปฏิบัติหน้าที่อื่น ๆ ตามที่ได้รับมอบหมาย

3. คุณสมบัติขั้นต่ำ

3.1 ความรู้และความสามารถ

1. หลักการบัญชีและระบบบัญชีราชการ หรือหน่วยงานที่เกี่ยวข้อง
2. ระเบียบการเงิน การงบประมาณ และกฎหมายที่เกี่ยวข้องกับการเงินการบัญชี
3. การสอบบัญชี
4. ให้คำแนะนำแก้ไข ปรับปรุงและทักท้วงเกี่ยวกับวิธีปฏิบัติทางบัญชี
5. การจัดทำรายงาน

3.2 วุฒิและความชำนาญงาน ปริญญาตรีทางบัญชี

แผนภาพที่ 5 รายงานอธิบายและกำหนดงาน

8. คำอธิบายงานนั้นระยะเวลาของการวิเคราะห์ไว้ชัดเจนหรือไม่
9. คำอธิบายงานนั้น มีรายละเอียดแสดงให้ทราบว่า ผู้ใดเป็นผู้ให้ความเห็นชอบแล้วหรือไม่ เมื่อใด
10. คำอธิบายงานนั้นระบุความแตกต่างระหว่างตำแหน่งงานต่าง ๆ ไว้ชัดเจนหรือไม่

ข้อกำหนดของคนทำงาน เมื่อนักวิเคราะห์งานมีคำอธิบายงานแล้ว ก็สามารถที่จะตีความหมายของคำอธิบายงานออกมาในแง่คุณสมบัติที่บุคคลจำเป็นต้องมี หากจะต้องปฏิบัติงานนั้นด้วยดี ความหมายที่วิเคราะห์ได้ใหม่นี้เรียกว่า ข้อกำหนดของคนทำงาน บางครั้งนักวิเคราะห์งานอาจเขียนข้อกำหนดของคนทำงานต่อจากคำอธิบายงาน เพื่อให้ได้ความสมบูรณ์ในรายงานฉบับเดียวไปเลย ในแผนภาพที่ 5 จะเห็นได้ว่า นอกจากจะมีคำอธิบายงานแล้ว ยังมีข้อกำหนดของคนทำงานต่อท้ายอีกด้วย รายงานเช่นนี้อาจเรียกว่า รายงานอธิบายและกำหนดงาน (job description—specification)

ในการแปลความหมายของคำอธิบายงานออกมาเป็นข้อกำหนดของคนทำงานนั้น ควรจะยึดหลัก 6 ประการ ดังต่อไปนี้²²

1. คำอธิบายงานควรสั้น แต่ต้องได้ความชัดเจน
2. สภาพการณ์ทำงานต้องระบุไว้ทั้งหมด
3. คุณสมบัติที่ต้องมีที่ระบุไว้ ต้องเป็นคุณสมบัติที่มีความตรงประเด็น (validity) กับตำแหน่งงานจริง ๆ กล่าวคือ ต้องเป็นคุณสมบัติที่หากขาดเสียแล้ว จะพิสูจน์ได้ว่า ผู้ปฏิบัติงานไม่สามารถปฏิบัติได้ดีหรือปฏิบัติไม่ได้
4. คุณสมบัติที่ระบุไว้ ต้องเป็นคุณสมบัติที่สามารถวัดได้ โดยมีความเที่ยงตรงแน่นอน
5. คุณสมบัติที่ระบุนี้ ต้องมีคำจำกัดความที่แน่นอนตายตัว
6. คุณสมบัติดังกล่าวจะต้องมีการตรวจสอบเป็นครั้งคราว เพื่อให้ทันสมัยกับความเปลี่ยนแปลงของงานในตำแหน่งนั้น ๆ

จากการที่พิจารณาขั้นตอนของการวิเคราะห์งานทั้งหมด จะเห็นได้ว่า ผลการวิเคราะห์งานเป็นสิ่งที่ขึ้นอยู่กับความแปรปรวนของตัวนักวิเคราะห์งานเองไม่น้อย ผลการวิเคราะห์งานที่จะนำมาใช้ประโยชน์ได้จริง ๆ ในการทดลองสร้างเครื่องมือวัด จำเป็นต้องขึ้นอยู่กับความเที่ยงตรงของการวิเคราะห์งานเป็นอย่างมาก ถึงแม้ว่าในที่นี้จะได้ชี้ให้เห็นว่ามีวิธีการบางอย่างที่อาจทำให้ผลการวิเคราะห์งานมีความเป็นปรนัยสูงขึ้น เช่น วิธีใช้รายการตรวจสอบในการวิเคราะห์กิจกรรมของตำแหน่งงาน แต่การที่จะพยายามทำให้ผลการวิเคราะห์งานมีความเป็นปรนัยสูงขึ้นนี้ หมายถึงค่าใช้จ่ายและความยากลำบากเพิ่มตามขึ้นมาด้วย ฉะนั้น ในการวิเคราะห์งานเพื่อนำผลมาเปรียบเทียบในการทดลองสร้างเครื่องมือวัด จึงจะต้องพิจารณาให้เข้ากันกับความสำคัญของงานในตำแหน่ง

²² French, *op. cit.*, p. 95.
Mee, *op. cit.*, p. 158.

หนึ่ง ๆ ที่จะมีการวัดบุคคล กล่าวคือ ตำแหน่งงานนั้นควรจะมีความสำคัญกับหน่วยงาน และหน่วยงานนั้นเห็นว่า หากได้บุคคลที่ขาดคุณสมบัติจริง ๆ มาปฏิบัติงานแล้ว จะเป็นผลเสียหายต่อหน่วยงาน²³

การประเมินผลการปฏิบัติงาน

ในการปฏิบัติงานปกติประจำวัน ผู้ทำงานในหน่วยงานเดียวกัน หรือในกลุ่มเดียวกัน มีความเห็นในลักษณะประเมินผู้ทำงานด้วยกันอยู่เสมอ จะโดยตั้งใจหรือไม่ตั้งใจก็ตาม ผู้บังคับบัญชา ใช้วิจารณ์คุณภาพผลงานของผู้ใต้บังคับบัญชา และคุณลักษณะที่เกี่ยวข้องกับสัมพันธภาพในการทำงานระหว่างผู้บังคับบัญชาและผู้ใต้บังคับบัญชา และอาจใช้ผลประเมินนั้นประกอบการพิจารณาเรื่องต่าง ๆ เกี่ยวกับบุคคลนั้น ในขณะเดียวกัน ผู้ใต้บังคับบัญชาก็ใช้วิจารณ์ของตนประเมินผู้บังคับบัญชา เพื่อนร่วมงาน งานของตน และหน่วยงานที่ตนทำงานอยู่ วิจารณ์บางเรื่องก็อาจจะเป็นเพียงความคิดเห็นส่วนตัวของแต่ละบุคคลเท่านั้น

เมื่อผู้บังคับบัญชามอบหมายงานให้ผู้ใต้บังคับบัญชาไปปฏิบัติแล้ว ย่อมมีการประเมินผลการปฏิบัติงาน (performance evaluation) ว่าดีไม่ดีเพียงใด ซึ่งมีผลทำให้ผู้บังคับบัญชาประเมินค่าของผู้ใต้บังคับบัญชาที่มีต่อหน่วยงานนั้น หน่วยงานบางแห่งอาจจัดโครงการประเมินผลการปฏิบัติงานเป็นแบบพิธีเป็นระบบ บางแห่งอาจไม่จัดทำเป็นระบบ เพียงแต่ผู้บังคับบัญชาประเมินผลงานของผู้ใต้บังคับบัญชาไปตามที่ตนเห็นสมควรเท่านั้น²⁴

อย่างไรก็ตาม หน่วยงานที่มีการบังคับบัญชาเป็นลำดับขั้นนั้น ย่อมเป็นที่ยอมรับกันว่า หน้าที่ในการประเมินผลงานของผู้ใต้บังคับบัญชาเป็นเรื่องสำคัญ เพราะผลประเมินจะนำไปใช้ประโยชน์ได้หลายประการ ที่ใช้กันแพร่หลาย ได้แก่²⁵

1. การเลื่อนตำแหน่ง ข้อมูลที่ได้จากการประเมินจะเป็นประโยชน์ในการพิจารณาว่าผู้ทำงานผู้ใดสมควรได้เลื่อนตำแหน่งสูงขึ้น หรือได้เลื่อนเป็นหัวหน้างาน
2. การโอนย้าย ผลประเมินจะช่วยชี้ว่า ผู้ทำงานผู้ใดมีความชำนาญ ความรู้ ความสามารถ ความสามารถในการปรับตัว เหมาะสมที่จะโอนย้ายไปทำงานในตำแหน่งที่ว่าง หรือทำงานในหน่วยงานส่วนอื่น

²³ Lawshe, *op. cit.*, p. 295.

²⁴ Chruden and Sherman, Jr., *op. cit.*, p. 247.

²⁵ Tiffin, Joseph, "Merit Rating: Its Validity and Techniques," in Doohar, M.J., and Marquis Vivienne, eds., *Rating Employee and Supervisory Performance*, pp. 11—12.

3. การพัฒนาบุคคล ผลประเมินจะเป็นประโยชน์ในการวิเคราะห์ข้อดี ข้อบกพร่องของพนักงาน เพื่อว่าทั้งฝ่ายบริหารหรือฝ่ายหัวหน้างานและผู้ปฏิบัติงานเอง จะได้พยายามพัฒนาทักษะและคุณลักษณะที่จำเป็นต่องาน เพื่อเพิ่มโอกาสที่จะได้ก้าวหน้ายิ่งขึ้น

4. การศึกษาปรับปรุงวิธีการเลือกสรรข้อมูลจากการประเมิน จะช่วยชี้ให้เห็นความแตกต่างของผู้ที่ทำงานได้ดีและไม่ดี ซึ่งจะเป็นช่องทางให้พิจารณาปรับปรุงวิธีการเลือกสรรและเครื่องมือที่ใช้ในการทดสอบเพื่อเลือกสรรบุคคลเข้าทำงาน

5. การให้ออกจากงาน ผลประเมินจะช่วยในการพิจารณาให้พนักงานออกจากการงานเมื่อปรากฏว่าไม่สามารถทำงานได้ดี และเมื่อต้องการลดจำนวนพนักงานได้เป็นธรรมยิ่งขึ้น

เมื่อพิจารณาการใช้ประโยชน์จากผลประเมินดังกล่าวข้างต้นแล้ว จะเห็นว่า ความก้าวหน้าในการทำงานของผู้ใต้บังคับบัญชาขึ้นอยู่กับ การตัดสินใจของผู้บังคับบัญชาอยู่เป็นอันมาก ยิ่งกว่านั้น การตัดสินใจดังกล่าว มีผลโดยตรงต่อการใช้และการพัฒนากำลังคน ซึ่งเป็นเรื่องที่มีความสำคัญต่อความก้าวหน้าของหน่วยงาน

วิวัฒนาการของแนวความคิดในการประเมิน วิธีการประเมินผลการปฏิบัติงานทางอุตสาหกรรมในสหรัฐอเมริกา ได้วิวัฒนาการมาโดยลำดับ เริ่มตั้งแต่การนำแบบอย่างของทหารมาใช้ ซึ่งเป็นแบบรายงานอุปนิสัยและพฤติกรรมของบุคคลโดยย่อ ต่อมาในระหว่างสงครามโลกครั้งที่หนึ่ง ได้มีการคิดค้นการประเมินเปรียบเทียบบุคคล เพื่อใช้ในการคัดเลือกคน และได้ปรับปรุงขึ้นใช้ในกองทัพก ภายหลังจากนั้นทั้งธุรกิจเอกชน วงการทหารและพลเรือนก็เพิ่มความสนใจในการประเมินผลงานยิ่งขึ้น ในระยะหลังสงครามโลกครั้งที่หนึ่งนั้น วงการอุตสาหกรรมมุ่งเน้นการประเมินค่าผู้ทำงานประเภททำงานในระดับขั้นต้นของหน่วยงาน ต่อมาในระยะหลังสงครามโลกครั้งที่สอง จึงได้เพิ่มความสนใจในการประเมินผลงานของผู้บริหารงานระดับกลางและระดับสูง ตลอดจนผู้ทำงานทางเทคนิคและวิชาชีพ

การเปลี่ยนแปลงจุดสนใจเป็นลำดับมา ประกอบกับพัฒนาการของแนวความคิดใหม่ๆ ในเรื่องการวัดผลบุคคล มีผลทำให้เกิดวิถีทางปฏิบัติในการประเมินแบบใหม่ๆ ขึ้น²⁶ อย่างไรก็ดี ไม่ว่าหน่วยงานจะใช้แบบวิธีการประเมินแบบใดก็ตาม สิ่งสำคัญที่ต้องคำนึงถึงก็คือ เกณฑ์การประเมิน (criteria) ซึ่งเป็นพื้นฐานที่ผู้ประเมินจะนำผล

²⁶ Chruden and Sherman, Jr., *op. cit.*, pp. 252—253.

งานของผู้ถูกประเมินแต่ละบุคคลมาเปรียบเทียบ เกณฑ์ดังกล่าว ก็คือ มาตรฐานของงาน (job or performance standards) หรือเกณฑ์งานที่ถือได้ว่าเป็นที่พอใจ (criteria of satisfactory performance) มาตรฐานหรือเกณฑ์เหล่านี้ จะต้องกำหนดขึ้นโดยอาศัยการศึกษา การทำความเข้าใจงานและข้อกำหนดของงาน (job requirements) เป็นหลัก²⁷

มาตรฐานของงาน เป็นเป้าหมายย่อยของหน่วยงาน (organizational subgoals) ในส่วนที่ผู้ทำงานแต่ละคนจะต้องทำให้สำเร็จ เป็นเป้าหมายที่ฝ่ายบริหารระดับสูงหรือฝ่ายจัดการกำหนดขึ้นไว้เป็นเกณฑ์สำหรับเปรียบเทียบกับการปฏิบัติงานจริง ๆ มาตรฐานของงานสัมพันธ์กับคำอธิบายงาน คือ คำอธิบายงานระบุว่าในงานนั้น ๆ มีอะไร จะต้องทำบ้าง ส่วนมาตรฐานของงานระบุว่างานที่จะต้องทำนั้นควรมี ปริมาณมากน้อยแค่ไหน หรือควรมี คุณภาพดีเพียงใด ในบางกรณีอาจพบว่า ในคำอธิบายนั้นเองระบบเพิ่มเติมเกี่ยวกับเป้าหมายของงานสำหรับระยะเวลาหนึ่งไว้ด้วย หรือมีจะนั้นก็แยกไว้เป็นข้อระบุต่างหากและเป็นส่วนเสริมคำอธิบายงานอีกทีหนึ่ง²⁸

โดยทั่วไป งานที่ทำเสร็จจะได้รับการพิจารณาประเมินผล ทั้งในด้านปริมาณและคุณภาพ งานดังกล่าวมีทั้งที่เป็นหน่วยนับได้ และไม่อาจนับได้เป็นหน่วยเสมอไป เพราะฉะนั้น ในการกำหนดมาตรฐานของงานจึงกำหนดไว้ในรูปของปริมาณและคุณภาพแล้วแต่ลักษณะของงาน ได้มีผู้รู้แนะนำว่า ควรพยายามเขียนเป็นปริมาณให้ได้มากที่สุดที่จะทำได้ นอกจากงานบางลักษณะซึ่งจะระบุเป็นปริมาณได้ยาก ก็จำเป็นต้องระบุคุณภาพของงานที่ควรทำได้ไว้²⁹ สำหรับงานที่นับเป็นหน่วย ได้แก่ งานประเภทผลิตรกรรม ก็สามารถเปรียบเทียบผลผลิตของบุคคลในด้านปริมาณและคุณภาพกับมาตรฐานที่กำหนดไว้ได้ง่าย การกำหนดมาตรฐานของงานเช่นนี้ได้วิวัฒนาการต่อเนื่องมาจากพัฒนาการเกี่ยวกับการจัดการแบบวิทยาศาสตร์ (scientific management) และการศึกษาจับเวลาเคลื่อนไหว (time studies) แนวการทำงานและปรับปรุงงานดังกล่าวนี้เป็นที่แพร่หลายในวงการอุตสาหกรรม ตั้งแต่ประมาณกลางศตวรรษที่ 20 โรงงานอุตสาหกรรมกำหนดจำนวนหน่วยผลิตสำหรับช่วงเวลาหนึ่งขึ้น ก็เพื่อใช้ในการคิดคำนวณเงินรางวัลเพิ่มเติมให้แก่ผู้ทำงานที่สามารถผลิตได้สูงกว่าเกณฑ์มาตรฐานด้วย³⁰

²⁷ *Ibid.*, p. 25—7.

²⁸ French, *op. cit.*, pp. 96—97.

²⁹ Patton, Arch, "How to Appraise Executive Performance," *Harvard Business Review*, 38: 63—73, January—February 1960.

³⁰ French, *op. cit.*, p. 98.

แม้แต่งานบางประเภท เช่น งานขายสินค้า ซึ่งแสดงผลงานส่วนใหญ่ได้ในรูปของจำนวน ก็อาจประเมินผลงานได้ในทำนองเดียวกับงานผลิตภัณฑ์ ส่วนงานประเภทที่ไม่สามารถประเมินในรูปของผลผลิตได้ หรืองานซึ่งอาศัยคุณสมบัติที่เป็นนามธรรม (intangible qualities) เป็นสำคัญ เช่น ความอดุสาหะพากเพียร ความริเริ่ม ความจับใจในการตัดสินใจ เป็นต้น ก็ย่อมต้องใช้วิธีการอื่น เป็นต้นว่า การสังเกต และรายงานของผู้บังคับบัญชามาประกอบการประเมิน

เมื่อพิจารณาในหลักการแล้ว การกำหนดมาตรฐานของงาน หรือการกำหนดเกณฑ์ก็มีแนวความคิดทำนองเดียวกับหลักการบริหาร โดยการกำหนดวัตถุประสงค์ (management by objectives)³¹ กล่าวคือ โดยหลักแล้ว การกำหนดมาตรฐานของงานก็ควรพิจารณาโดยอาศัยพื้นฐานจากวัตถุประสงค์รวมของหน่วยงานและยึดถือหลักนี้ตั้งแต่หน่วยงานระดับสูงตลอดลงไปทุกระดับ

โดยเหตุที่ผลประเมินนำไปใช้ประกอบการพิจารณาตัดสินใจเกี่ยวกับการบริหารงานบุคคลอีกหลายเรื่อง จึงจำเป็นที่จะจัดให้มีการประเมินอย่างเป็นระบบมีลักษณะปรนัยและให้ความเป็นธรรม เมื่อพิจารณาในแง่ของการบริหาร ก่อนดำเนินการประเมินผลงานก็ควรจะได้มีการกำหนดนโยบายให้ชัดเจน กำหนดแบบวิธีดำเนินการให้เหมาะสม และจัดการฝึกอบรมผู้ประเมินให้สามารถประเมินได้ถูกต้องเสียก่อน อาจกล่าวได้ว่าโครงการประเมินผลการปฏิบัติงาน จะเป็นผลสำเร็จตามจุดประสงค์หรือไม่เพียงใดนั้นขึ้นอยู่กับผู้บริหารระดับสูงและผู้บังคับบัญชา ผู้บริหารระดับสูงจะต้องเห็นคุณค่าของการประเมินที่จะมีต่องานและผู้ทำงาน ผู้บังคับบัญชาต้องเข้าใจหลักการ และเมื่อได้รับการฝึกอบรมให้รู้วิธีการประเมินแล้ว จะต้องเป็นผู้ร่วมมือร่วมใจดำเนินการอย่างเต็มที่

เมื่อพิจารณาในแง่ของการวัดผล ก็ควรจะพิจารณาส่งเสริมให้การประเมินมีความเชื่อมั่นและความเที่ยงตรงอยู่ในระดับสูง Wendell French ได้กล่าวถึงคุณลักษณะสำคัญ 2 ประการของการประเมินโดยสรุปไว้ว่า³² ความเชื่อมั่นของผลการประเมินผลการปฏิบัติงาน หมายถึง ความคงเส้นคงวา (consistency) ของผลการประเมินหลายๆ ครั้ง ที่ผู้บังคับบัญชาประเมินผู้ใต้บังคับบัญชา ในกรณีที่ผลงานของผู้ใต้บังคับบัญชาไม่ได้เปลี่ยนแปลงไป

³¹ Drucker, P. F., *The Practice of Management*, pp. 121—136.

³² French, *op. cit.*, pp. 187—189.

ความเที่ยงตรงของการประเมินผลการปฏิบัติงาน หมายถึง ความตรงวัตถุประสงค์ของการประเมิน กล่าวคือ การประเมินมีความมุ่งหมายที่จะส่งเสริมงานของหน่วยงานให้บรรลุวัตถุประสงค์ด้วยดี ถ้าหากผลประเมินมีส่วนสนับสนุนความมุ่งหมายนี้ได้มากเพียงใด ก็เท่ากับมีความเที่ยงตรงมากเพียงนั้น และองค์ประกอบสำคัญของความเที่ยงตรงก็คือ ความเชื่อมั่นของการประเมิน หมายความว่า การประเมินจะก่อให้เกิดประโยชน์เสมอตามวัตถุประสงค์ก็ต่อเมื่อมีคุณลักษณะที่เชื่อถือได้เสียก่อน

โดยเหตุที่การประเมินผลการปฏิบัติงานไม่ว่าจะกระทำให้ได้ผลออกมาเป็นตัวเลขหรือโดยใช้เกณฑ์ที่เป็นปรนัย (objective criteria) ได้เสมอไปทุกเรื่อง มิฉะนั้นแล้ว ก็คงจะทำให้แม่นยำได้ง่ายขึ้น หากแต่การประเมินต้องอาศัยวิจารณ์ญาณ (judgments) เป็นพื้นฐาน เพราะฉะนั้นกระบวนการประเมินจึงอาจมีข้อผิดพลาดได้ ข้อผิดพลาดดังกล่าวอาจมีสาเหตุหลายประการ ดังที่ Dalton E. McFarland ได้ตั้งข้อสังเกตเกี่ยวกับสาเหตุที่เป็นอุปสรรคต่อความสำเร็จในการใช้ระบบประเมินผลไว้ เป็นต้นว่า ความเชื่อที่ไม่ตรงกับความเป็นจริงของฝ่ายผู้บริหารและผู้บังคับบัญชา ปัญหาทางจิตวิทยา และข้อบกพร่องทางเทคนิคการประเมิน³³ จึงได้มีการพยายามที่จะส่งเสริมให้ผลประเมินมีความเชื่อถือได้ และความเที่ยงตรงในระดับสูงโดยวิธีต่าง ๆ วิธีหนึ่งที่สำคัญก็คือ การวิจัย ซึ่งใช้ข้อมูลที่เกี่ยวข้องกับงานภายในหน่วยงานนั้นหลาย ๆ เรื่องประกอบกัน Chruden และ Sherman ได้แนะข้อควรศึกษาวิเคราะห์ไว้บางประการดังนี้³⁴

1. ศึกษาแนวโน้มของผู้ประเมิน วิธีทางหนึ่งที่จะพิจารณาความเที่ยงตรงของการประเมิน คือ การศึกษาแนวโน้มว่าผู้ประเมินหลาย ๆ คนประเมินผู้ใต้บังคับบัญชาของตนไปในแนวใด บางคนอาจจะไม่ใคร่เคร่งครัดนัก บางคนอาจจะเข้มงวด เพราะฉะนั้น ผลประเมินของผู้บังคับบัญชาสองแบบนี้ จะนำมาเปรียบเทียบกันในทันทีไม่ได้ นอกจากจะได้ตั้งข้อสังเกตเกี่ยวกับความแตกต่างนี้ไว้ด้วย

2. พิจารณาความแตกต่างของหน่วยงานและประเภทงาน ถ้าหากได้มีการวิเคราะห์ผลการประเมินผลการปฏิบัติงานในหน่วยงานอย่างรอบคอบ มักจะพบว่า การประเมินในส่วนต่าง ๆ ของหน่วยงาน หรือในงานประเภทต่าง ๆ นั้น ปรากฏผลประเมินแตกต่างกันมาก เพราะฉะนั้น การเปรียบเทียบผลประเมิน จึงควรเปรียบเทียบในหมู่คนที่ทำงานประเภทเดียวกันมิใช่ต่างประเภทกัน

³³ McFarland, D.E., *Personnel Management*, pp. 324—329.

³⁴ Chruden and Sherman, Jr., *op. cit.*, pp. 263—265.

3. จัดระบบคะแนน (scoring system) โดยการคิदन้าหนักให้แก่คุณสมบัตินของพฤติกรรม หรือลักษณะเฉพาะบุคคล (characteristics) ที่ใช้ในการประเมิน กำหนดน้ำหนักต่าง ๆ กันโดยถือเกณฑ์ข้อกำหนดของงานเป็นสำคัญ หมายความว่าให้น้ำหนักมากแก่คุณลักษณะที่ต้องการใช้ในางานมาก ให้น้ำหนักน้อยแก่คุณลักษณะที่ต้องการใช้ในางานน้อยกว่า วิธีการนี้ช่วยให้ผู้ประเมินตระหนักว่าคุณลักษณะต่าง ๆ มีคุณค่าแตกต่างกัน

4. พิจารณาผลประเมินเปรียบเทียบกับข้อมูลอื่นด้วย การวัดผลงานในรูปแบบอื่น เช่น รายงานแสดงปริมาณผลิต ควรนำมาใช้ประกอบการประเมินผลงานของบุคคลให้มากที่สุดเท่าที่จะทำได้ โดยเฉพาะอย่างยิ่งเมื่อผลประเมินจะต้องใช้ในการตัดสินใจเกี่ยวกับบุคคล ยิ่งจำเป็นต้องศึกษาความสัมพันธ์ระหว่างผลประเมินกับข้อมูลอื่น ๆ เกี่ยวกับบุคคลนั้น เป็นต้นว่า การลงโทษทางวินัย การร้องเรียน การลาหยุดงาน ข้อมูลเหล่านี้จะช่วยในการพิจารณาลักษณะปรนัยของการประเมินได้

เมื่อการประเมินมีความเชื่อถือได้ และความเที่ยงตรงสูงก็จะอำนวยความสะดวกแก่หน่วยงานได้จริง และมีส่วนส่งเสริมให้หน่วยงานปฏิบัติงานตามเป้าหมายที่กำหนด

บทที่ 3

ตัวแปรของการวัดผล

ตัวแปรของการวัดผล (predictor variable) หมายถึงการบันทึกค่าของตัวแปรที่จะใช้คาดคะเนค่าของตัวแปรอื่น ๆ¹ เฉพาะในเรื่องการวัดผลและประเมินบุคคลนี้ ค่าที่บันทึกได้เป็นค่าพฤติกรรมที่ให้เป็นตัวเลข และเนื่องจากค่าพฤติกรรมที่บันทึกได้นี้ มีความแปรปรวน (variance) ทั้งส่วนที่รวมกันกับผลการบันทึกพฤติกรรมนั้น ๆ และทั้งส่วนที่รวมกันกับผลการบันทึกพฤติกรรมครั้งต่อไป การที่ผลการบันทึกมีลักษณะความแปรปรวนเช่นนี้ จึงเรียกชื่อว่า ตัวแปรของการวัดผล ถ้าจะเปรียบเทียบเรื่องนี้ให้เห็นได้ชัด ก็อาจกล่าวได้ว่า การพิจารณาตัวแปรของการวัดผล เป็นการพิจารณาความแปรปรวนของผลการบันทึกค่าตัวแปรคาดคะเน (predictive variable) หรือค่าตัวแปรอิสระ (independent variable) หรืออาจจะกล่าวอีกอย่างหนึ่งได้ว่า การวิเคราะห์ตัวแปรของการวัดผล² คือ การวิเคราะห์ความแปรปรวนของเครื่องมือวัดนั่นเอง

ก่อนที่จะพิจารณารายละเอียดว่า เครื่องมือวัดมีความแปรปรวนประการใด จึงจะกล่าวได้ว่าเป็นตัวแปรของการวัดผล ควรจะได้พิจารณาคุณสมบัติของเครื่องมือวัดเสียก่อน

คุณสมบัติของเครื่องมือวัด

ดังได้กล่าวไว้ในบทที่ 1 แล้วว่า คนมีลักษณะพฤติกรรมต่างกัน การที่เป็นเช่นนั้นหมายความว่า กิจกรรมบางอย่างที่บุคคลคนหนึ่งทำได้ด้วยความคล่องแคล่ว อาจเป็นงานที่บุคคลบางคนทำได้ด้วยความยากลำบาก การที่จะทราบลักษณะความแตกต่างของ

¹ Good, C.V., ed., *Dictionary of Education*, p. 596.

² Guion, R.M., *Personnel Testing*, p. 51.

แต่ละบุคคลนั้น เป็นหน้าที่ของผู้บริหารงานบุคคลโดยตรง เครื่องมือวัดเป็นเครื่องช่วยในการวิเคราะห์ความแตกต่างของบุคคล ในเรื่องนี้ อาจแยกคุณสมบัติของเครื่องมือวัดออกได้เป็นสองประการ คือ ความสามารถในการคาดคะเน (prediction) และความสามารถในการตรวจวิเคราะห์ (diagnosis)

ความสามารถในการคาดคะเน การวิจัยในแขนงวิชาการวัดผลบุคคลในรอบศตวรรษนี้ ได้เน้นหนักเรื่องความสามารถของเครื่องมือวัดในการที่จะคาดคะเนความสามารถในการปฏิบัติงาน หรือในการแสดงพฤติกรรมจากผลที่วัดได้โดยเครื่องมือวัดหนึ่ง ๆ โดยตลอดมา นักวิจัยได้พยายามหาคำตอบคำถามที่ว่า เครื่องมือวัดนี้แสดงว่าบุคคลมีความรู้ หรือสามารถปฏิบัติงานในเรื่องนั้น ๆ ได้เพียงใด ผลจากการวัดตามเครื่องมือวัดนี้ จะช่วยให้สามารถจัดบุคคลไปปฏิบัติงานอย่างหนึ่ง ๆ ได้สำเร็จเพียงใด เครื่องมือวัดที่แสดงแต่เพียงว่าบุคคลตั้งแต่สองคนขึ้นไปต่างกันอย่างไร ในขณะที่วัดจะมีค่าน้อยยามาก³ เครื่องมือวัดนั้น ๆ จะมีค่าก็ต่อเมื่อสามารถบอกได้อีกว่า บุคคลตั้งแต่สองคนซึ่งต่างกันนั้น เมื่อไปปฏิบัติงานในระยะต่อไปจะได้ผลต่างกันอย่างไร การพิจารณาผลการวัดผลการวัดกับความสามารถที่จะปฏิบัติงานได้ในระยะต่อไปหลังจากที่วัดก็คือ การพิจารณาค่าสหสัมพันธ์ระหว่างผลการวัดของเครื่องมือวัดหนึ่ง ๆ กับความสามารถในการปฏิบัติงานหนึ่ง ๆ ในการหาค่าสหสัมพันธ์ในเรื่องนี้ นักวิจัยได้พยายามใช้วิธีการวิจัยต่าง ๆ หลักเกณฑ์ง่าย ๆ ในเรื่องนี้ก็คือ พยายามเปรียบเทียบค่าที่เครื่องมือวัดหนึ่ง ๆ วัดได้กับมาตรวัดความสำเร็จของงาน (criterion measure) ที่ทดลองกับบุคคลกลุ่มเดียวกัน เช่น ในการหาคุณสมบัติของการคาดคะเนของแบบทดสอบมาตรฐาน Differential Aptitude Test (DAT) ชูตวิชาช่างเครื่องยนต์ ในปี ค.ศ. 1954 Doppelt, Seashore และ Odgers⁴ ได้ทดลองนำเครื่องมือวัดนี้ไปวัดนักเรียนวิชาช่างกล แผนกเครื่องยนต์ระดับ มศ. 4 ในโรงเรียนอาชีวศึกษา 6 แห่ง ในมลรัฐโอไฮโอ ระหว่างภาคการศึกษาแรกปีการศึกษา 1954 และเมื่อสิ้นภาคการศึกษาได้ทดลองให้ครูผู้สอนในชั้นวัดผลการศึกษาในประเด็น 4 ประเด็น คือ ความสามารถในการเข้าใจลักษณะอาชีพของช่างเครื่องยนต์ ความรู้ทางเทคนิคเครื่องยนต์ คุณภาพของการปฏิบัติงาน และปริมาณงานที่สามารถปฏิบัติ

³ Cronbach, L.J., *Essentials of Psychological Testing*, p. 9.

⁴ Guion, *op. cit.*, pp. 264–266.

ได้ และเมื่อนักเรียนเหล่านั้นไปเรียนในชั้น มศ. 5 ก็ได้มีการวัดเช่นนี้อีก ผลของการวัดได้นำมาวิเคราะห์ความสัมพันธ์ ทำให้ทราบค่าความสามารถในการคาดคะเนของแบบทดสอบมาตรฐาน

ความจำเป็นของการคาดคะเน เห็นได้ชัดในกรณีที่จะต้องมีการคัดเลือกบุคคลเข้าปฏิบัติงาน การคัดเลือกคนเข้าปฏิบัติงานทั้งในวงการรัฐบาลและเอกชนของสหรัฐ ได้อาศัยผลการวัดของแบบทดสอบมาตรฐานประกอบการพิจารณาเป็นอย่างมาก ผลการวัดนี้ไม่เพียงแต่ทำให้ทราบว่า บุคคลหนึ่ง ๆ จะปฏิบัติงานในตำแหน่งที่จะบรรจุได้หรือไม่เท่านั้น แต่ยังทำให้ทราบอีกว่า บุคคลเหล่านั้นควรจะได้รับการฝึกอบรมเพิ่มเติมหรือไม่ เครื่องมือวัดที่ดัดย้อมบอกได้ว่า บุคคลใด ลักษณะใด เป็นบุคคลที่มีโอกาสน้อยมากที่จะปรับตัวให้เข้ากับงาน หรือแม้แต่เพียงจะได้ประโยชน์จากการฝึกอบรม

ก่อนที่ Thurstone จะพิสูจน์ว่า ความสามารถทางสมองของบุคคล (mental ability) ประกอบด้วยความสามารถหลาย ๆ อย่างรวมกัน ในปี 1938⁵ นักจิตวิทยาเชื่อว่า เซวาร์นด์ หรือ ความฉลาด (intelligence) เป็นคุณสมบัติทางจิตวิทยาที่แยกไม่ได้ (unitary factor) ความเชื่อเช่นนี้ทำให้ออกันในวงการวัดผลว่า เมื่อสามารถวัดความฉลาดนี้ได้ ก็จะทำให้ทราบองค์ประกอบหลักในการตัดสินใจว่า บุคคลหนึ่ง ๆ จะปฏิบัติงานได้สำเร็จหรือไม่ ผลจากการพิสูจน์ของ Thurstone ทำให้มีการศึกษาคุณสมบัติของการคาดคะเนเครื่องมือวัดความฉลาดใหม่ ทำให้ทราบว่า แท้ที่จริงแล้วผลการวัดความฉลาดสามารถใช้คาดคะเนความสามารถของบุคคลในการปฏิบัติงานได้ดีมาก จนถึงกับได้มีการสรุปว่า “ไม่มีเครื่องมือวัดใด ๆ ไม่ว่าจะ เป็นเครื่องมือวัดความฉลาด หรือเครื่องมือวัดพฤติกรรมอื่น ๆ ที่จะวัดคุณสมบัติทั่วไปของบุคคล อันจะเป็นหลักให้สามารถคาดคะเนได้ว่า บุคคลหนึ่ง ๆ จะปฏิบัติงานต่าง ๆ ได้ดีเพียงไร” แต่ทั้งนี้ไม่ได้หมายความว่า เครื่องมือวัดบุคคลจะคาดคะเนอะไรไม่ได้เลย เครื่องมือวัดหนึ่ง ๆ เมื่อใช้ให้ถูกต้อง จะสามารถคาดคะเนโอกาสที่บุคคลจะทำงานในระยะต่อไปได้สำเร็จตามสัดส่วนความสัมพันธ์ระหว่างเครื่องมือวัดกับงานนั้น ๆ จากตารางที่ 2 จะเห็นได้ว่า เครื่องมือวัดลักษณะต่าง ๆ ของบุคคล จะมีศักยภาพในการคาดคะเนโอกาสที่จะปฏิบัติงานในระยะต่อไปต่าง ๆ กัน

⁵ *Ibid.*, p. 61.

ตารางที่ 2

ศักยภาพในการคาดคะเนของเครื่องมือวัด*

ชนิดของงานที่จะปฏิบัติ	เครื่องมือวัดแยกตามลักษณะการวัด		
	ความฉลาด	ความสามารถในการ เข้าใจมิติ (spatial)	ความประสานระหว่างจิต และอวัยวะ (psycho-motor)
หัวหน้างาน (foreman)	0.52	0.49	0.46
เสมียน (general clerk)	0.60	0.34	0.48
พนักงานเดินตลาด ผู้ใช้นายนต์ (vehicle operator)	0.32	0.50	0.57
ช่างแก้ไขรถยนต์	0.35	0.50	0.69
ช่างไฟฟ้า	0.50	0.53	0.48
(electrical worker)	0.69	0.65	0.56
ผู้ดูแลเครื่องจักร	0.47	0.39	0.52
(machine tender)	0.49	0.52	0.39
ผู้ตรวจงาน (inspector)			

*แหล่งที่มา : Lawshe, C.H., and Balma, M.J., *Principles of Personnel Testing*, p. 87.

จากตารางข้างต้นนี้ แสดงว่า งานในหน้าที่หัวหน้างานต้องการผู้ที่มีคุณสมบัติต่างไปจากงานในหน้าที่ผู้ตรวจงาน ค่า .52 ในช่องความฉลาดของหัวหน้างาน หมายความว่า เครื่องมือวัดความฉลาดที่เชื่อมั่นได้ (reliable) จะมีสัมประสิทธิ์ของความเที่ยงตรง (validity coefficient) กับการปฏิบัติงานในหน้าที่หัวหน้างาน 0.52 ตัวเลขอื่น ๆ ก็มีความหมายโดยนัยเดียวกันนี้ สัมประสิทธิ์ของความเที่ยงตรงนี้จะอธิบายโดยละเอียดในบทที่ 5

ความสามารถในการตรวจวิเคราะห์ นอกจากจะใช้เครื่องมือวัดเปรียบเทียบความสามารถปฏิบัติได้ และพฤติกรรมระหว่างบุคคลเพื่อเป็นการคาดคะเนดังกล่าวมาแล้ว นักทฤษฎีในการวัดยังได้พยายามใช้เครื่องมือวัดให้เป็นประโยชน์ในแง่การเปรียบเทียบลักษณะนิสัยของบุคคลหนึ่ง ๆ อีกด้วย การเปรียบเทียบลักษณะนิสัยภายในบุคคลหนึ่ง ๆ นี้ เรียกว่าเป็นการตรวจวิเคราะห์ การที่นำเครื่องมือวัดมาตรวจวิเคราะห์ก็เพื่อให้ทราบ

ความเข้มแข็งและความอ่อนแอของบุคคลหนึ่ง ๆ ผลที่วัดได้ทำให้ทราบว่าบุคคลหนึ่ง ๆ มีปัญหาอันใดบ้างในการที่ปรับตัวไม่ได้ (maladjustment) ⁶

ลักษณะการวัดเพื่อการคาดคะเน ต่างกับการวัดเพื่อการตรวจวิเคราะห์ตรงที่วิธีการแปลความหมายของค่าคะแนนที่วัดได้ ถ้าหากจะวัดเพื่อการคาดคะเน เมื่อผู้วัดได้ใช้เครื่องมือวัดหนึ่งแล้ว ปรากฏว่าบุคคลที่ถูกวัดคนหนึ่งได้ค่าคะแนน (test score) ต่ำมาก ผู้วัดก็ไม่จำเป็นจะต้องสนใจว่า เหตุใดบุคคลผู้นั้นจึงได้ค่าคะแนนต่ำ แต่ถ้าหากว่าผู้วัดต้องการวัดในเชิงตรวจวิเคราะห์ ก็มีความจำเป็นอย่างยิ่งที่จะต้องสนใจว่า เหตุใดบุคคลผู้นั้นจึงได้ค่าคะแนนต่ำ ผู้วัดจะต้องศึกษาว่า บุคคลผู้นั้นมีลักษณะนิสัยอันใดบ้างที่มีส่วนทำให้ได้รับค่าคะแนนต่ำ ทั้งนี้เพื่อให้ทราบได้ว่า บุคคลผู้นั้นมีปัญหาทางจิตวิทยาประการใด

การวัดในเชิงตรวจวิเคราะห์ใช้มากในวงการจิตแพทย์และวงการศึกษาศาสตร์ ส่วนในวงการบริหารงานบุคคลในต่างประเทศ แม้จะได้มีการส่งเสริมให้มีการวัดบุคคลทั้งในแง่การคาดคะเนและในแง่การตรวจวิเคราะห์ แต่ก็แทบจะกล่าวได้ว่าหน้าที่สำคัญของการวัดงานบุคคลก็คือการคาดคะเน ส่วนในประเทศไทยนั้น การวัดงานบุคคลพยายามดำเนินการเฉพาะในเรื่องการคาดคะเนเท่านั้น ฉะนั้นในหนังสือเล่มนี้ จึงจะไม่กล่าวถึงรายละเอียดของการวัดในเชิงตรวจวิเคราะห์อีก

ความสัมพันธ์ระหว่างค่าความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์

ในกรณีที่ค่าของข้อมูลสองชุดมีองค์ประกอบ (factor) ซึ่งทำให้ข้อมูลทั้งสองชุดมีคุณสมบัติร่วมกัน (common entity) หรือเป็นเหตุทำให้ค่าของข้อมูลทั้งสองชุดเกิดความแปรปรวนร่วมกัน ข้อมูลทั้งสองชุดนี้กล่าวได้ว่า มีความแปรปรวนขององค์ประกอบร่วมกัน (common factor variance) ความแปรปรวนขององค์ประกอบร่วมกันนี้อาจเป็นคุณสมบัติที่มีได้ระหว่างข้อมูลมากกว่าสองชุดก็ได้ ฉะนั้นเพื่อเป็นการแยกประเภทความแปรปรวนขององค์ประกอบร่วมกันระหว่างข้อมูลเพียงสองชุด จึงมีชื่อเรียกโดยเฉพาะว่า ความแปรปรวนร่วม (covariance)

ค่าของความแปรปรวนสัมพันธ์ อาจคำนวณได้จากสมการข้างล่างนี้⁷

$$\sigma_{xy}^2 = \frac{\Sigma(X_i - \bar{X}_x)(Y_i - \bar{Y}_y)}{n}$$

⁶ Cronbach, *op. cit.*, pp. 10—11.

⁷ Nunnally, J.C., *Psychometric Theory*, pp. 113—114.

- σ_{xy}^2 = ความแปรปรวนสัมพันธ์
 X_i = ค่าของแต่ละจำนวนในข้อมูลชุด x
 \bar{X}_x = ค่าเฉลี่ยของข้อมูลชุด x
 Y_i = ค่าของแต่ละจำนวนในข้อมูลชุด y
 \bar{X}_y = ค่าเฉลี่ยของข้อมูลชุด y
 n = จำนวนทั้งหมดของข้อมูลชุด x หรือชุด y

เมื่อเทียบค่าความแปรปรวนสัมพันธ์กับสัมประสิทธิ์สหสัมพันธ์ (coefficient correlation) แล้ว จะเห็นว่า ค่าทั้งสองตัวเกี่ยวข้องกันอย่างใกล้ชิด ดังสูตรของสัมประสิทธิ์สหสัมพันธ์ข้างล่าง⁸

$$r = \frac{\sigma_{xy}^2}{\sqrt{\sigma_x^2 \cdot \sigma_y^2}}$$

r = สัมประสิทธิ์สหสัมพันธ์
 σ_x^2 = ค่าความแปรปรวนของข้อมูลชุด x
 σ_y^2 = ค่าความแปรปรวนของข้อมูลชุด y

เมื่อใช้วิธีการทางพีชคณิตวิเคราะห์สูตรข้างต้นนี้ จะได้ผลดังต่อไปนี้

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \cdot \sigma_y^2}$$

เนื่องจากได้นิยามไว้แล้วว่า ค่าความแปรปรวนขององค์ประกอบร่วมกันคือค่าความแปรปรวนสัมพันธ์นั่นเอง ต่างกันเพียงแต่ว่า ค่าความแปรปรวนขององค์ประกอบร่วมกัน เป็นค่าที่แสดงความสัมพันธ์ระหว่างข้อมูลตั้งแต่สองชุดขึ้นไป ฉะนั้นถ้าหากว่าผู้วิจัยสามารถคำนวณค่าสหสัมพันธ์ได้ ก็ย่อมสามารถคำนวณค่าความแปรปรวนขององค์ประกอบร่วมกันได้ โดยนำค่าสหสัมพันธ์มายกกำลังสอง เช่น ถ้าค่าสหสัมพันธ์เท่ากับ 0.90 ค่าความแปรปรวนขององค์ประกอบร่วมกันก็จะเท่ากับ $(0.90)^2$ คือ 0.81 คำนี้นแสดงว่า 81 เปอร์เซ็นต์ของความแปรปรวนในข้อมูลชุดหนึ่ง มีคุณสมบัติขององค์ประกอบร่วมกันกับความแปรปรวนของข้อมูลอีกชุดหนึ่งหรือชุดอื่น ๆ

การที่หาค่าความแปรปรวนขององค์ประกอบร่วมกันได้จากค่าสหสัมพันธ์นี้ มีประโยชน์มาก เพราะค่าสหสัมพันธ์นี้เป็นค่าที่หาได้สะดวกจากสูตรง่าย ๆ หรือเครื่อง

⁸ Kerlinger, F.N., *Foundations of Behavioral Research*, p. 111.

คำนวณ ฉะนั้นเมื่อทราบค่าสหสัมพันธ์ของเครื่องมือวัดหนึ่ง ๆ แล้ว ก็อาจวิเคราะห์ลักษณะของตัวแปรของการวัดผลต่อไปได้ เช่น สมมติว่า ถ้ามีข้อสอบมาตรฐานสองชุด ชุดที่หนึ่งวัดความสามารถในการเข้าใจคน ชุดที่สองวัดความสามารถที่จะคิดริเริ่มในการทำงาน ข้อสอบทั้งสองชุดนี้ ใช้สำหรับทดสอบบุคคลที่จะปฏิบัติงานในตำแหน่งที่สมมติตำแหน่งหนึ่ง ข้อสอบทั้งสองชุดมีสัมประสิทธิ์ความเชื่อมั่น (reliability coefficient) 0.90 และ 0.80 ตามลำดับ นอกจากนี้ หลังจากการวิจัยข้อสอบทั้งสองแล้ว ปรากฏว่ามีสัมประสิทธิ์สหสัมพันธ์ 0.72 ระหว่างข้อสอบมาตรฐานทั้งสอง จากตัวอย่างนี้อาจนำค่าสหสัมพันธ์ต่าง ๆ มาพิจารณาในแง่ความแปรปรวนขององค์ประกอบรวมกันได้ กล่าวคือ ค่าความเชื่อมั่นของข้อสอบมาตรฐานที่วัดความสามารถในการเข้าใจคน เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดนี้กับการวัดความสามารถในการเข้าใจคน เหตุผลของการเป็นสัมประสิทธิ์สหสัมพันธ์จะได้กล่าวต่อไปในบทที่ 4 ฉะนั้นเมื่อนำค่านี้นมายกกำลังสอง จะได้เป็นค่าความแปรปรวนขององค์ประกอบรวมกัน ระหว่างเครื่องมือวัดชุดนี้กับลักษณะนิสัยเรื่องความสามารถในการเข้าใจคน และในทำนองเดียวกัน ก็อาจหาค่าความแปรปรวนขององค์ประกอบรวมกันระหว่างเครื่องมือวัดชุดที่ 2 กับลักษณะนิสัยเรื่องความคิดริเริ่ม ดังรายละเอียดต่อไปนี้

$(0.90)^2 = 0.81$ หรือ 81 เปอร์เซ็นต์ ของความแปรปรวนในการวัดของเครื่องมือวัดชุดความสามารถในการเข้าใจคน มีคุณสมบัติขององค์ประกอบรวมกันกับความแปรปรวนของลักษณะนิสัยในเรื่องความสามารถในการเข้าใจคน หรือกล่าวได้ว่า อีก 19 เปอร์เซ็นต์ที่เครื่องมือวัดนี้วัดได้ เกิดจากความผิดพลาด

$(0.80)^2 = 0.64$ หรือ 64 เปอร์เซ็นต์ของความแปรปรวนในการวัดของเครื่องมือวัดชุดความคิดริเริ่มในการทำงาน มีคุณสมบัติขององค์ประกอบรวมกันกับความแปรปรวนของลักษณะนิสัยในเรื่องความสามารถในการคิดริเริ่มในการทำงาน หรือกล่าวได้ว่า อีก 36 เปอร์เซ็นต์ที่เครื่องมือวัดนี้วัดได้ เกิดจากความผิดพลาด

$(0.72)^2 = 0.52$ หรือ 52 เปอร์เซ็นต์ของเครื่องมือวัดทั้งสอง มีความแปรปรวนขององค์ประกอบรวมกัน

$0.81 - 0.52 = 0.29$ หรือ 29 เปอร์เซ็นต์ของเครื่องมือวัดชุดการเข้าใจคน สามารถวัดลักษณะนิสัยในเรื่องการเข้าใจคนได้

$0.64 - 0.52 = 0.12$ หรือ 12 เปอร์เซ็นต์ของเครื่องมือวัดชุดความคิดริเริ่มสามารถวัดลักษณะนิสัยในเรื่องความสามารถที่จะคิดริเริ่มในการทำงานได้จริง ๆ

จะเห็นได้ว่า เครื่องมือวัดทั้งสองชุด วัดลักษณะนิสัยที่เครื่องมืออื่น ๆ ตั้งเกณฑ์ไว้ว่าจะวัดได้เพียง 29 เปอร์เซนต์ และ 12 เปอร์เซนต์ตามลำดับเท่านั้น ทั้ง ๆ ที่เครื่องมือวัดทั้งสองมีสัมประสิทธิ์ความเชื่อมั่นสูงถึง 0.90 และ 0.80 ตามลำดับ ทั้งนี้เพราะเครื่องมือวัดทั้งสองมีค่าสหสัมพันธ์ระหว่างกันสูง

มิติของการวัด

ดังที่ได้กล่าวไว้แล้วในบทที่ 1 ว่า มิติเป็นเรื่องที่สำคัญเรื่องหนึ่งในการวัดบุคคล แต่เนื่องจากลักษณะการวัดบุคคลต่างไปจากการวัดลักษณะของร่างกาย หรือการวัดสิ่งอื่น ๆ ที่มองเห็นด้วยตาได้ ฉะนั้นเพื่อที่จะวิเคราะห์มิติของเครื่องมือวัดบุคคลหนึ่ง ๆ นักทฤษฎีในการวัดจึงต้องมีวิธีการพิสูจน์ เรียกว่า วิเคราะห์องค์ประกอบ (factor analysis) ในกรณีที่เครื่องมือวัดแม้จะมีค่าความเชื่อมั่นสูง การวิเคราะห์มิติก็เป็นสิ่งจำเป็นมาก เช่น สมมติว่า มีเครื่องมือวัดบุคคลเพื่อบรรจุเข้าปฏิบัติงานในตำแหน่งหนึ่ง เครื่องมือนั้นสามารถคาดคะเนได้อย่างแม่นยำพอสมควรว่า บุคคลซึ่งแสดงความสามารถและพฤติกรรมในระดับหนึ่งตามเครื่องมือวัดนั้น จะปฏิบัติงานในตำแหน่งหน้าที่ดังกล่าวได้ เนื้อหาของการวัดของเครื่องมือวัดนั้น ก็เป็นลักษณะรวมของตัวแปรทางเศรษฐกิจและสังคม (socio-economic variables) ซึ่งผู้สร้างเครื่องมือวัดได้กำหนดขึ้นตามแนวคิด (operational concept) ของตนเองเท่านั้น⁹ เครื่องมือนั้นจะต้องวัดอะไรบางอย่างที่เกี่ยวกับความสามารถที่จะไปปฏิบัติงานในตำแหน่งหน้าที่ดังกล่าว แต่ “อะไรบางอย่าง” นี้ ยังกำหนดแน่ชัดลงไปไม่ได้ เพราะฉะนั้นวิธีวิเคราะห์องค์ประกอบจึงช่วยในการวิเคราะห์มิติของเครื่องมือวัด เพื่อให้ทราบแน่ชัดลงไปว่า ลักษณะที่จะวัดตามเครื่องมือวัดนั้นประกอบด้วย “ธาตุ” อะไรบ้าง

โดยหลักการง่าย ๆ วิธีวิเคราะห์องค์ประกอบดำเนินการโดยนำผลการทดลองวัดของเครื่องมือวัดหลาย ๆ อัน มาคำนวณหาสัมประสิทธิ์สหสัมพันธ์และจากค่าสัมประสิทธิ์สหสัมพันธ์นี้ จึงมีการเปรียบเทียบค่าความแปรปรวนขององค์ประกอบร่วมกัน ทฤษฎีการหามิติโดยวิธีวิเคราะห์องค์ประกอบมีหลายทฤษฎี รายละเอียดในเรื่องนี้จะกล่าวโดยเฉพาะไว้อีกบทหนึ่งต่างหาก ในที่นี้จะกล่าวเฉพาะแนวคิดเบื้องต้นพร้อมกับผลงานเพื่อทำให้เกิดความเข้าใจลักษณะต่าง ๆ ของตัวแปรในการวัดผลเท่านั้น

⁹ Shaw, M.E., and Wright, J.M., *Scales for the Measurement of Attitudes*, pp. 563-564. Guion, *op. cit.*, p. 52.

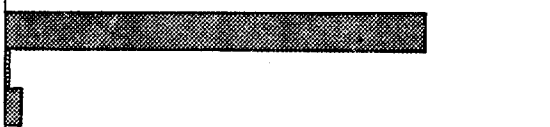
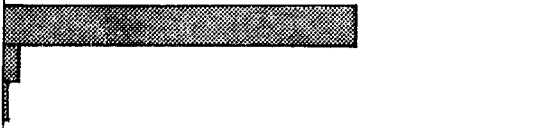
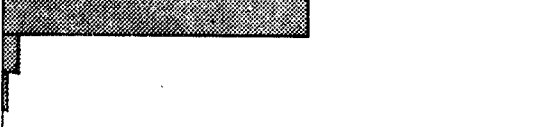

การที่จะเข้าใจแนวคิดเบื้องต้นของการวิเคราะห์องค์ประกอบ อาจทำได้โดยลงสมมติว่า ผู้วัดทราบแล้วว่า เครื่องมือวัดของตนวัดมิติอันใดบ้าง (ซึ่งความจริงไม่มีทางทราบก่อน) ถ้ามีเครื่องมือวัดอยู่ 4 ชุด ชุดที่หนึ่งวัดความสามารถในการเข้าใจคน ชุดที่สองวัดความสามารถในการคิดริเริ่มในการทำงาน ชุดที่ 3 วัดความเป็นตัวของตัวเอง ชุดที่ 4 วัดความสามารถในการพิมพ์ดีด เครื่องมือแต่ละชุดมีค่าสหสัมพันธ์กับองค์ประกอบ ชื่อว่า ความสามารถเข้าใจนามธรรม ความสามารถหาเหตุผล และความสามารถในการใช้ถ้อยคำ ดังตารางที่ 3

ตารางที่ 3

ค่าสหสัมพันธ์สมมติระหว่างเครื่องมือวัดกับองค์ประกอบ

องค์ประกอบ เครื่องมือวัด	ความสามารถ เข้าใจนามธรรม	ความสามารถ หาเหตุผล	ความสามารถ ในการใช้ถ้อยคำ
การเข้าใจคน	0.90	0.00	0.10
ความรู้ริเริ่ม	0.80	0.10	0.05
ความเป็นตัวเอง	0.78	0.10	0.05
ความสามารถ ในการพิมพ์ดีด	0.10	0.15	0.40

จากค่าสหสัมพันธ์ของเครื่องมือวัดกับองค์ประกอบที่สมมตินี้ จะเห็นว่าเครื่องมือวัด 3 ชุดแรก มีค่าสหสัมพันธ์กับองค์ประกอบที่เรียกว่า ความสามารถเข้าใจนามธรรมสูง ส่วนเครื่องมือวัดชุดที่ 4 มีค่าสหสัมพันธ์กับองค์ประกอบนี้น้อยมาก ในขณะเดียวกัน เครื่องมือวัดชุดที่ 2 และที่ 3 มีค่าสหสัมพันธ์กับองค์ประกอบคล้ายคลึงกันมาก กล่าวคือ ค่าสหสัมพันธ์กับองค์ประกอบที่เรียกว่า ความเข้าใจนามธรรมต่างกันเพียง 0.02 ค่าสหสัมพันธ์กับองค์ประกอบที่เรียกว่า ความสามารถในการหาเหตุผลและความสามารถในการใช้ถ้อยคำเท่ากัน คุณสมบัติอันนี้ เครื่องมือวัดชุดอื่นๆ ไม่มี ถ้าจะวิเคราะห์ลักษณะดังกล่าวให้เห็นได้ชัดขึ้น ก็อาจแปลงค่าสหสัมพันธ์เป็นค่าความแปรปรวนขององค์ประกอบร่วมกัน ดังแสดงไว้ในแผนภูมิที่ 1

เครื่องมือวัดชุดที่	องค์ประกอบ	ค่าความแปรปรวนขององค์ประกอบร่วมกัน
1	ความเข้าใจนามธรรม ความสามารถหาเหตุผล ความสามารถในการใช้ถ้อยคำ	
2	ความเข้าใจนามธรรม ความสามารถหาเหตุผล ความสามารถในการใช้ถ้อยคำ	
3	ความเข้าใจนามธรรม ความสามารถหาเหตุผล ความสามารถในการใช้ถ้อยคำ	
4	ความเข้าใจนามธรรม ความสามารถหาเหตุผล ความสามารถในการใช้ถ้อยคำ	

แผนภูมิที่ 1 ค่าความแปรปรวนขององค์ประกอบร่วมกันที่สมมติ

จากแผนภูมิที่ 1 จะสังเกตเห็นได้ชัดขึ้นกว่าตารางที่ 3 อีกว่า ค่าความแปรปรวนขององค์ประกอบร่วมกันระหว่างเครื่องมือวัดความคิดริเริ่ม และความเป็นตัวเองกับองค์ประกอบทั้ง 3 มีลักษณะร่วมกันอยู่มาก ฉะนั้น สัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดทั้งสองนี้ จึงอาจกล่าวได้ว่ามีค่าสูง ส่วนสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดความคิดริเริ่มในการทำงาน กับเครื่องมือวัดความสามารถในการเข้าใจคน นับว่ารองลงไป และที่รองลงไปเป็นลำดับที่ 3 คือ เครื่องมือวัดความสามารถในการเข้าใจคนกับความเป็นตัวเอง เครื่องมือวัดชุดที่ 4 นั้นไม่มีสัมประสิทธิ์สหสัมพันธ์กับเครื่องมือวัดชุดอื่น ๆ เลย เพราะไม่มีสิ่งอันใดที่วัดร่วมกันกับเครื่องมือวัดชุดอื่น ๆ

การที่เครื่องมือวัดต่าง ๆ มีค่าสหสัมพันธ์ร่วมกัน จึงเป็นทางที่ทำให้พิจารณาต่อไปได้ว่า เครื่องมือวัดเหล่านั้น สามารถวัดความสามารถของบุคคล และพฤติกรรมในประเด็นที่ร่วมกัน

ตารางสัมพันธ์แบบเอฟและแบบอาร์

ตารางที่ 4 เป็นตารางที่แสดงความสัมพันธ์ระหว่างเครื่องมือวัดกับองค์ประกอบตารางเช่นนี้เรียกว่า ตารางสัมพันธ์ (matrix) ในกรณีที่แสดงความสัมพันธ์กับองค์ประกอบที่วัดโดยเฉพาะ เรียกว่า ตารางสัมพันธ์แบบเอฟ (F. Matrix)¹⁰ จากตารางดังกล่าวจะเห็นว่าเครื่องมือวัดแต่ละชุดวัดองค์ประกอบ (saturated) ตั้งแต่สองตัวขึ้นไป (factorially complex) การที่เครื่องมือวัดองค์ประกอบหลาย ๆ อย่างเช่นนี้ จึงเป็นปัญหาที่นักสร้างเครื่องมือวัด พยายามสร้างเครื่องมือวัดที่สามารถวัดแยกองค์ประกอบแต่ละตัวได้ (factorially pure) การที่เครื่องมือวัดที่วัดองค์ประกอบหลาย ๆ ตัวพร้อมกันไปไม่เป็นที่พึงปรารถนา ก็เพราะว่า เครื่องมือวัดเช่นนี้ย่อมไม่สามารถแสดงค่าของความสัมพันธ์ระหว่างผลของการปฏิบัติงานหรือพฤติกรรมตามเครื่องมือวัดกับค่าของตัวแปรตาม (dependent variable) เช่น เครื่องมือวัดความสามารถในการพิมพ์ดีดที่ขกไว้เป็นตัวอย่างไม่ไม่มีลักษณะที่จะสามารถแยกวัดองค์ประกอบแต่ละตัวได้ ผลของการวัดโดยเครื่องมือนี้ย่อมทำให้มั่นใจไม่ได้ว่า ความสัมพันธ์ระหว่างค่าคะแนนที่วัดได้จากเครื่องมือนี้กับความสามารถในการพิมพ์ดีดเมื่อถึงเวลาปฏิบัติงาน เป็นความสัมพันธ์ตามแนวที่ผู้ใช้เครื่องมือวัดต้องการ ถ้าเครื่องมือวัดทั้งความสามารถในการใช้ถ้อยคำ และทั้งความสามารถในการใช้เครื่องพิมพ์ ค่าคะแนนของผู้ที่ได้ทำ อาจจะไม่ได้แปลว่า บุคคลผู้นั้นจะไม่สามารถปฏิบัติงานเป็นพนักงานพิมพ์ดีดได้

ค่าสหสัมพันธ์ระหว่างเครื่องมือวัดกับองค์ประกอบต่างๆ เรียกได้ว่าเป็นค่าองค์ประกอบของเครื่องมือวัด (factor loading)¹¹ ค่าองค์ประกอบทุกๆ ตัวของเครื่องมือวัดหนึ่งๆ เมื่อยกกำลังสองแล้วนำมาบวกกัน เรียกว่า ค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบของเครื่องมือวัดอันหนึ่ง ๆ (communality)¹² คำนี้หมายถึงอัตราร้อยละของความแปรปรวนของเครื่องมือวัดนั้นๆ ที่มีส่วนเกิดมาจากองค์ประกอบดังกล่าว อาจพิจารณาความสัมพันธ์ระหว่างค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบของเครื่องมือวัดอันหนึ่ง ๆ กับค่าองค์ประกอบของเครื่องมือวัดได้ในตารางซึ่งสมมติองค์ประกอบสองตัวดังตารางที่ 4

¹⁰ Nunnally, *op. cit.*, pp. 291–296.

¹¹ *Ibid.*, p. 293.

¹² Guilford, J.P., *Psychometric Methods*, p. 355.

ตารางที่ 4

ตารางสัมพันธ์แบบ เอฟ ของชุดเครื่องมือวัดความเป็นผู้นำ (สมมติ)

เครื่องมือวัด องค์ประกอบ	ความสามารถเข้าใจ ในนามธรรม	ความสามารถ หาเหตุผล	h^2
การเข้าใจคน	0.90	0.00	0.81
ความริเริ่ม	0.80	0.10	0.65
ความเป็นตัวเอง	0.70	0.10	0.50
การประสานงาน	0.10	0.70	0.50
การปฏิบัติงานเป็นกลุ่ม	0.10	0.80	0.65

h^2 คือตัวของค่าสัมพันธ์ร่วมระหว่างองค์ประกอบของเครื่องมือวัดอันหนึ่ง ๆ กับค่าสัมพันธ์ร่วม ๆ ของเครื่องมือวัดความสามารถในการเข้าใจคนนี้เท่ากับ

$$(0.90)^2 + (0.00)^2 = 0.81$$

ตารางสัมพันธ์แบบเอฟ มีความสัมพันธ์กับตารางสัมพันธ์แบบอาร์ โดยความเป็นจริงตามทฤษฎีวิเคราะห์องค์ประกอบ ผู้สร้างเครื่องมือวัดจะสร้างตารางสัมพันธ์แบบอาร์ได้ก่อน แล้วจึงนำมาวิเคราะห์เพื่อสร้างตารางสัมพันธ์แบบเอฟ แต่ในขั้นนี้ผู้พิจารณาย้อนกลับโดยสมมติตารางสัมพันธ์แบบเอฟเสียก่อน

ตารางสัมพันธ์แบบอาร์ คือ ตารางที่แสดงความสัมพันธ์ระหว่างเครื่องมือวัดต่าง ๆ ค่าความสัมพันธ์ที่แสดงในตารางประเภทนี้ คือ ค่าสหสัมพันธ์ระหว่างเครื่องมือวัด¹³ โดยปกติเมื่อผู้สร้างเครื่องมือวัดได้ทดลองเครื่องมือวัดของตน เพื่อใช้ในการวัดกับบุคคลกลุ่มต่าง ๆ แล้ว ก็ย่อมสามารถหาสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดที่ตนทดลองสร้างขึ้นได้ ตามทฤษฎีการวิเคราะห์องค์ประกอบ สัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัด มีค่าเท่ากับผลรวมของสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดนั้น ๆ กับองค์ประกอบทั้งหมด ดังสมการในหน้า 72¹⁴

¹³ Nunnally, *op. cit.*, pp. 296–303.

¹⁴ Kerlinger, *op. cit.*, p. 654.

$$R = FF'$$

R = เป็น matrix ของ r สัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัด

F = สัมประสิทธิ์สหสัมพันธ์ของเครื่องมือวัดชุดที่ 1 กับองค์ประกอบแต่ละตัว

F' = สัมประสิทธิ์สหสัมพันธ์ของเครื่องมือวัดชุดที่ 2 กับองค์ประกอบแต่ละตัว

จากตารางที่ 5 ซึ่งแสดงสัมประสิทธิ์สหสัมพันธ์สมมติระหว่างเครื่องมือวัด 5 ชุด กับองค์ประกอบสองตัว อาจคำนวณค่าสหสัมพันธ์ระหว่างเครื่องมือวัดทั้ง 5 ชุดนี้ได้ เช่น สัมประสิทธิ์ระหว่างเครื่องมือวัดการเข้าใจคนกับความริเริ่มจะมีค่าเท่ากับ $(0.90)(0.80) + (0.00)(0.10)$ คือ 0.72 ฉะนั้น เมื่อกำหนดตามวิธีนี้แล้ว ก็อาจแสดงผลเป็นตารางสัมพันธ์แบบอาร์ ดังจะเห็นได้จากตารางที่ 5

ตารางที่ 5

ตารางสัมพันธ์แบบอาร์ ของชุดเครื่องมือวัดความเป็นผู้นำ (สมมติ)

เครื่องมือวัด	การเข้าใจคน	ความริเริ่ม	ความเป็นตนเอง	การประสานงาน	การปฏิบัติงานเป็นกลุ่ม
การเข้าใจคน		0.72	0.63	0.09	0.09
ความริเริ่ม	0.72		0.57	0.15	0.16
ความเป็นตนเอง	0.63	0.57		0.14	0.15
การประสานงาน	0.09	0.15	0.14		0.57
การปฏิบัติงานเป็นกลุ่ม	0.09	0.16	0.15	0.57	

ผลการทดลองสร้างเครื่องมือวัดในศตวรรษนี้ ได้ทำให้ได้ผลการวิเคราะห์ตารางสัมพันธ์แบบอาร์มากมาย ตัวอย่างของจริงของตารางสัมพันธ์แบบอาร์ที่ได้จากการวิจัยเครื่องมือวัดที่ใช้ในกองทัพอากาศสหรัฐ แสดงไว้ในตารางที่ 6

ตารางที่ 6

ตารางสัมพันธ์แบบอาร์ ระหว่างเครื่องมือวัด 9 ชุดของกองทัพอากาศสหรัฐ*

เครื่องมือวัด †	1	2	3	4	5	6	7	8	9
1		0.38	0.55	0.06	-0.04	0.05	0.07	0.05	0.08
2	0.38		0.36	0.40	0.28	0.40	0.11	0.15	0.13
3	0.55	0.36		0.10	0.01	0.18	0.13	0.12	0.10
4	0.06	0.40	0.10		0.32	0.60	0.04	0.06	0.13
5	-0.04	0.28	0.01	0.32		0.35	0.08	0.13	0.11
6	0.05	0.40	0.18	0.60	0.35		0.01	0.06	0.07
7	0.07	0.11	0.13	0.04	0.08	0.01		0.45	0.32
8	0.05	0.15	0.12	0.06	0.13	0.06	0.45		0.32
9	0.08	0.13	0.10	0.13	0.11	0.07	0.32	0.32	

* แหล่งที่มา : Guion, R.M., *Personnel Testing*, p. 57.

† ชื่อของเครื่องมือวัด

1. การใช้คำคล้ายกัน (synonym test)
2. ศัพท์วิชาการ (technical vocabulary for pilot)
3. ความสามารถในการอ่านหนังสือวิชาการ (reading comprehension)
4. ความรู้ในการใช้เครื่องกล (tool functions)
5. ประสพการณ์ (biographical data blank)
6. ความรู้เกี่ยวกับเครื่องยนต์และวิธีแก้ (mechanical information)
7. ความสามารถในการเข้าใจมิติ (spatial orientation)
8. ความเร็วในการรู้จักสิ่งต่างๆ (speed of identification)
9. ความสามารถในการจัดประมวลแบบ (pattern assembly)

จากผลของการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดเช่นตารางที่ 6 นี้ ทำให้นักทฤษฎีการวิเคราะห์องค์ประกอบสามารถหาองค์ประกอบที่ซ่อนอยู่ในตัวเครื่อง

มีอวัตรแต่ละชุดได้ ในระหว่างที่ยังไม่ได้พิจารณารายละเอียดของทฤษฎีการวิเคราะห์องค์ประกอบ ก็อาจจะสร้างความเข้าใจลักษณะของตัวแปรการวัดผลได้โดยลองพิจารณาค่าสหสัมพันธ์ในตารางที่ 6 ด้วยวิธีง่าย ๆ ซึ่งเรียกว่า วิธีการวิเคราะห์เชื่อมโยงเบื้องต้น (elementary linkage analysis)¹⁵ กล่าวคือ ในตารางนี้ ค่าสหสัมพันธ์ที่สูงสุดและลดหลั่นกันลงไปตามลำดับได้แก่ ค่า 0.60 ระหว่างเครื่องมือวัดที่ 4 และที่ 6 ค่า 0.55 ระหว่างเครื่องมือวัดที่ 1 และที่ 3 และ 0.45 ระหว่างเครื่องมือวัดที่ 7 และที่ 8 โดยเฉพาะในแต่ละชุดของกลุ่มที่ 1 และกลุ่มที่ 3 ยังมีค่าสหสัมพันธ์ที่อยู่ในเกณฑ์สูงกับชุดอื่น ๆ อีก คือ ชุดที่ 4 มีค่าสหสัมพันธ์กับชุดที่ 2 เท่ากับ 0.40 ชุดที่ 6 กับชุดที่ 5 เท่ากับ 0.35 ชุดที่ 8 กับชุดที่ 9 เท่ากับ 0.32 ฉะนั้น เครื่องมือวัดที่มีค่าสหสัมพันธ์ระหว่างกันในเกณฑ์สูงอาจจัดได้เป็น 3 กลุ่ม กลุ่มที่หนึ่งได้แก่ชุดที่ 2, 4, 5 และ 6 กลุ่มที่สองชุดที่ 1 และ 3 กลุ่มที่ 3 ชุดที่ 7, 8 และ 9 ค่าสหสัมพันธ์ที่อยู่ในเกณฑ์สูงที่จัดเป็นกลุ่มได้เช่นนี้ ถือได้ว่า มีการวัดมิติเดียวกัน หรือกล่าวตามทฤษฎีการวิเคราะห์องค์ประกอบว่า วัดองค์ประกอบร่วมกัน คุณสมบัติเช่นนี้ทำให้นักทฤษฎีวิเคราะห์องค์ประกอบสามารถวิเคราะห์ต่อไปได้ว่า องค์ประกอบที่เครื่องมือวัดวัดร่วมกันเช่นนี้มีคุณสมบัติประการใด แล้วจึงตั้งชื่อขององค์ประกอบนั้นต่อไป การตั้งชื่อขององค์ประกอบนั้นขึ้นอยู่กับดุลพินิจของนักวิเคราะห์โดยถือจากคุณสมบัติความร่วมมือกันของการวัดนี้

ตัวอย่างของผลการวิเคราะห์มิติของตัวแปรของการวัดผล อาจพิจารณาได้จากผลงานวิจัยของ Rokeach และ Fruchter¹⁶ ซึ่งรายงานไว้ในปี ค.ศ. 1956 นักวิจัยทั้งสองได้ศึกษาความมีทิฐิมานะและความเชื่อที่ไม่ยอมเปลี่ยนแปลง (dogmatism) งานวิจัยเรื่องนี้ต้องการพิสูจน์ว่าความเบียดเบียน (paranoia) และการไม่ยอมรับตัวเอง (self-rejection) เป็นส่วนหนึ่งของความมีทิฐิมานะ จากผลของการคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัด ทำให้ผู้วิจัยสามารถสร้างตารางสัมพันธ์แบบเอฟของเครื่องมือวัดตัวแปรต่าง ๆ ของตนได้ ดังแสดงไว้ในตารางที่ 7

¹⁵ Nunnally, *op. cit.*, pp. 364—365.

¹⁶ Kerlinger, *op. cit.*, pp. 660, 672—674.

ตารางที่ 7

ตารางสัมพัทธ์แบบเอฟของชุดเครื่องมือวัดความมีทิฐิมานะ*

องค์ประกอบ เครื่องมือวัดตัวแปร	1	2	3	h^2
ความกังวล	0.77	-0.25	0.27	0.727
ความเบาเต้ง	0.72	-0.14	0.26	0.597
การไม่ยอมรับตัวเอง	0.69	-0.29	0.37	0.698
ทิฐิมานะ	0.46	0.21	0.62	0.637
ความเชื่อในเผด็จการ	0.27	0.48	0.66	0.737
ความยึดมั่นในระเบียบแบบแผน	0.23	0.32	0.71	0.652
ความถือตัวในเชื้อชาติและวัฒนธรรม	0.21	0.59	0.47	0.614
อนุรักษ์นิยม	-0.07	0.69	0.19	0.523
ความนิยมฝ่ายซ้าย	0.17	-0.63	0.26	0.498
ความนิยมฝ่ายขวา	0.03	0.85	0.19	0.765

*แหล่งที่มา : Kerlinger, F.N., *Foundations of Behavioral Research*, p. 660.

ตารางสัมพัทธ์แบบเอฟที่แสดงไว้นี้ ไม่ได้แสดงชื่อขององค์ประกอบ เพียงแต่ระบุหมายเลขที่ขององค์ประกอบไว้ ตารางนี้ช่วยให้ผู้วิจัยสามารถพิสูจน์ได้ 3 ประการ ประการแรก ทิฐิมานะ มีองค์ประกอบของการวัดต่างไปจากความเชื่อในเผด็จการ ความยึดมั่นในระเบียบแบบแผน และความถือตัวในเชื้อชาติและวัฒนธรรม ประการที่สอง ความกังวล ความเบาเต้ง การไม่ยอมรับตัวเองและทิฐิมานะมีองค์ประกอบร่วมกัน และประการสุดท้าย อนุรักษ์นิยม ความนิยมฝ่ายซ้าย และความนิยมฝ่ายขวา วัดองค์ประกอบต่างกัน

ทฤษฎี Spearman และ Thurstone ในการแยกมิติ

ในปี ค.ศ. 1904 Spearman ได้เสนอทฤษฎีองค์ประกอบสองตัว (two-factor theory) ซึ่งเป็นที่ยกย่องในวงวิชาการเป็นอย่างมาก ทฤษฎีนี้ถือว่า เครื่องมือวัดที่มีค่าสหสัมพันธ์ร่วมกัน แปลว่า เครื่องมือวัดนั้นต้องวัดองค์ประกอบที่ร่วมกัน¹⁷ จากผลของทฤษฎีนี้ และหลักมาตรฐานวัดความเป็นสัดส่วน (criterion of proportionality) ทำให้ Spearman

¹⁷ Guilford, *op. cit.*, pp. 472-474.

สามารถพิสูจน์ทฤษฎีองค์ประกอบทั่วไปของเครื่องมือวัด (general-factor theory) ในปี ค.ศ. 1927 และเป็นแม่บทที่ทำให้เกิดการสร้างทฤษฎีองค์ประกอบอื่นๆ ตามมาอีก ทฤษฎีที่มีชื่อแข่งกับทฤษฎีของ Spearman คือทฤษฎีกลุ่มขององค์ประกอบ (multiple-factor theory) ซึ่ง Thurstone เป็นผู้พิสูจน์ ในปี ค.ศ. 1938

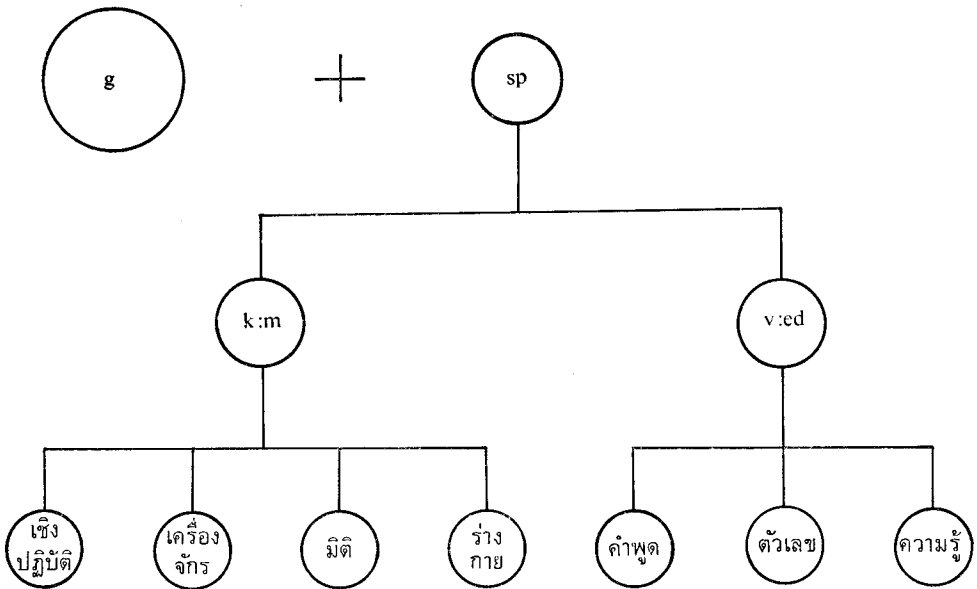
ทฤษฎีองค์ประกอบทั่วไปของเครื่องมือวัดถือว่า ความสามารถทางสมอง (intellectual ability) เป็นองค์ประกอบใหญ่ และเป็นตัวสำคัญที่ทำให้เกิดความแปรปรวนในการแสดงความสามารถหรือพฤติกรรมตามเครื่องมือวัดต่างๆ ไป ความสามารถทางสมองนี้ Spearman สมมติชื่อย่อว่า *g* การที่ Spearman ถูเช่นนี้ก็เพราะว่า ผลของการวิเคราะห์ของ Spearman ชี้ให้เห็นว่าเครื่องมือวัดที่วัดความสามารถทางสมองล้วนมีสัมประสิทธิ์สหสัมพันธ์เชิงปฏิฐานระหว่างกันทั้งสิ้น ผลพิสูจน์ข้อนี้ ตรงกับข้อสมมติ (postulate) ของนักวัดผลในรุ่นก่อน Spearman เช่น Alfred Binet Spearman ได้พิสูจน์ว่า เมื่อสร้างตารางสัมพันธ์แบบอาร์แล้ว ผลการวัดจะแสดงลำดับลดหลั่น (hierarchy) ขององค์ประกอบร่วม (common factor) ที่เครื่องมือวัดต่างๆ วัดได้ Spearman ได้ชี้ต่อไปอีกว่า เครื่องมือวัดต่างๆ จะสามารถวัดความสามารถพิเศษเฉพาะ (specific ability) ได้บ้างเป็นส่วนน้อย แต่เนื่องจากความสามารถพิเศษเฉพาะนี้ มีลักษณะเฉพาะและปริมาณที่วัดได้มีน้อย จึงไม่น่าจะมีความสำคัญที่จะต้องวัด¹⁸ กล่าวอีกนัยหนึ่งก็คือว่าเครื่องมือวัดที่วัดองค์ประกอบร่วม คือ ความสามารถทางสมองเท่านั้นที่จะคาดคะเนความสามารถที่จะปฏิบัติงาน หรือแสดงพฤติกรรมในระยะต่อไปได้ ส่วนเครื่องมือวัดที่วัดความสามารถพิเศษเฉพาะไม่มีคุณค่าในการคาดคะเนเลย

สาระสำคัญของทฤษฎีนี้ตรงกับสามัญสำนึกของบุคคลทั่วไป หลังจากที่ Spearman ได้พิสูจน์ทฤษฎีของตนแล้ว ต่อมาในปี ค.ศ. 1950 Vernon¹⁹ ก็ได้พิสูจน์ต่อไปอีกว่า *g* เป็นองค์ประกอบตัวสำคัญ ที่ทำให้เกิดความแปรปรวนรวม (total variance) ในการแสดงพฤติกรรมตามเครื่องมือวัดที่ใช้วัดทหารเกณฑ์ในกองทัพบกและกองทัพเรือของสหราชอาณาจักร ความแปรปรวนที่เหลืออีกไม่ถึง 50 เปอร์เซ็นต์เกิดจากความสามารถพิเศษเฉพาะสองกลุ่ม กลุ่มแรก ได้แก่ ความสามารถในเรื่องคำพูด (verbal) ตัว

¹⁸ Travers, R.M.W., *Educational Measurement*, p. 351.

¹⁹ Guion, *op. cit.*, pp. 60–61.

เลข (numerical) และความรู้ที่เรียนมา (education) รวมเรียกย่อว่า v:ed กลุ่มที่สอง เป็นความสามารถในเชิงปฏิบัติ (practical) เกี่ยวกับเครื่องจักร (mechanical) ใน การเข้าใจมิติ (spatial) และความสามารถของร่างกาย (physical) รวมเรียกย่อว่า k:m องค์ประกอบตามทฤษฎีนี้แสดงไว้ในแผนภูมิที่ 2



แผนภูมิที่ 2 องค์ประกอบตามทฤษฎี Spearman

จากแผนภูมิที่ 2 จะเห็นได้ว่า กำหนดให้ g ซึ่งแทนองค์ประกอบความสามารถทางสมองมีขนาดใหญ่กว่าความสามารถพิเศษเฉพาะ เพราะได้กล่าวมาแล้วว่า เครื่องมือวัดที่พิสูจน์ตามทฤษฎีนี้ มีความปรวนแปรอันเกิดจากองค์ประกอบร่วมกันกว่า 50 เปอร์เซ็นต์ ส่วนองค์ประกอบในกลุ่ม v:ed และกลุ่ม k:m นั้น แยกออกได้เป็นองค์ประกอบเฉพาะ 7 ตัว

ในปัจจุบันถือว่า ทฤษฎีของ Spearman วิเคราะห์มิติของเครื่องมือวัดหายากเกินไป การที่ Spearman พิสูจน์ได้ผลตามทฤษฎีของตนก็เพราะได้ทดลองแต่เครื่องมือวัดความ

สามารถทางสมองคล้าย ๆ กันเท่านั้น ลักษณะของเครื่องมือวัดก็มีลักษณะแต่ในเรื่องการใช้คำพูดเป็นหลัก ซึ่งผู้ที่จะแสดงความสามารถหรือพฤติกรรมตามเครื่องมือนี้ จะต้องมีมีความสามารถในการคิดหาเหตุผลตามแนวที่ Spearman ได้ตั้งไว้ว่าเป็นหัวใจของเขาวงกตหรือความฉลาด การที่ผลการค้นคว้าของ Spearman จำกัดอยู่แต่ในวงเครื่องมือวัดประเภทเดียวเช่นนี้ จึงทำให้นักทฤษฎีการวัดผลรุ่นหลังเกิดความพยายามที่จะวิเคราะห์มิติของตัวแปรการวัดผลให้ละเอียดลึกซึ้งกว้างขวางออกไปอีก ผลงานที่สำคัญคือ ทฤษฎีของ Thurstone ซึ่งเป็นทฤษฎีที่มีแนวความคิดตรงข้ามกับของ Spearman กล่าวคือ Thurstone ถือว่าการวิเคราะห์มิติของเครื่องมือวัด ของตัวแปรการวัดผลจะต้องพิจารณาจากความแตกต่างระหว่างบุคคลมากกว่าความคล้ายคลึงกันของบุคคลทั้งหลาย การที่ Spearman ได้ตั้งทฤษฎีองค์ประกอบร่วม เป็นการที่พิจารณามิติของตัวแปรของเครื่องมือวัดในแง่ความคล้ายกันของบุคคลทั้งหลาย Thurstone ได้ทดลองใช้เครื่องมือวัดหลาย ๆ ชุด กับนักเรียนที่คัดเลือกไว้แล้วเป็นจำนวน 250 คน เครื่องมือวัดเหล่านี้ได้ตั้งเกณฑ์ที่จะวัดความสามารถในเชิงนามธรรม ในการเข้าใจมิติ ในเชิงหาเหตุผลกับตัวเลข ในเชิงหาเหตุผลกับถ้อยคำ ในเชิงหาเหตุผลกับมิติ และในเชิงความจำ ชุดเครื่องมือวัดที่มีชื่อเสียงของ Thurstone คือ Chicago Tests of Primary Mental Ability ซึ่งวัดองค์ประกอบ 6 ตัว ได้แก่ ความสามารถในการเข้าใจคำพูด ความคล่องในเรื่องภาษา การหาเหตุผล ความจำ ตัวเลข และการเข้าใจมิติ ผลการวิจัยเครื่องมือวัดของ Thurstone พิสูจน์ได้ว่า เครื่องมือวัดความสามารถต่าง ๆ หลายชุดมีค่าสหสัมพันธ์กันถึงระดับที่จะสรุปได้ว่า มีกลุ่มขององค์ประกอบที่ทำให้เครื่องมือวัดเกิดความแปรปรวน Thurstone จึงได้ตั้งทฤษฎีใหม่²⁰ ที่เรียกว่า ทฤษฎีกลุ่มขององค์ประกอบ ทฤษฎีนี้ถือว่ามีองค์ประกอบที่สำคัญอยู่ 9 ตัว ในตัวแปรของการวัดผล ได้แก่ความสามารถในการเข้าใจมิติ (space ability) ความรวดเร็วทางสัญชาตญาณ (perceptual speed) ความคล่องทางตัวเลข (numerical facility) ความสามารถในเรื่องคำพูด (verbal relations) ความคล่องในเรื่องภาษา (word fluency) ความจำ (memory) อุปมาน (induction) การวิเคราะห์เหตุผล (reasoning) และอนุมาน (deduction) รวมเรียกเป็นตัวย่อว่า SPNVWMIRD

²⁰ Travers, *op. cit.*, pp. 351—355.
Guion, *op. cit.*, p. 61.

นักวัดผลและนักจิตวิทยาในระยะ 30 ปีที่ผ่านมาได้ถือทฤษฎีกลุ่มองค์ประกอบเป็นหลักในการวิเคราะห์เครื่องมือวัด และลักษณะนิสัยของบุคคล ถึงแม้ว่าผลการวิจัยในรุ่นหลัง Thurstone ได้พิสูจน์ว่า แท้ที่จริงแล้วตัวแปรของการวัดผลยังสามารถแยกแยะองค์ประกอบได้อีกมากมาย แต่องค์ประกอบตามทฤษฎีของ Thurstone ก็ยังเป็นเค้าโครงใหญ่ที่ยึดถือกันต่อมา ผลการวิจัยที่แยกแยะรายละเอียดขององค์ประกอบต่างๆ เช่น ผลงานของ Guilford, Cattell และ French ซึ่งเป็นหลักในการอธิบายลักษณะนิสัยที่วัดได้ในเรื่องเชาวน์ ความถนัดพิเศษ และบุคลิกภาพ ซึ่งถือเป็นคำอธิบายลักษณะนิสัยของบุคคลที่วัดได้ในปัจจุบัน จะได้แยกกล่าวไว้ในบทความต่างๆ ของเรื่องนั้นโดยเฉพาะ

ค่าความเชื่อมั่นของเครื่องมือวัด

การวัดที่เป็นมาตรฐาน (standardized testing) จะต้องคำนึงว่าพีธีการในการวัดเชื่อมั่น (reliable) ได้เพียงใด การพิจารณาเรื่องนี้หมายถึงการวิเคราะห์ว่า การวัดนั้น ๆ วัดสิ่งที่ต้องการวัดได้อย่างแม่นยำ (accurate) สักเท่าใด ผลของการวัดมีค่าเที่ยงตรง (precision) หรือไม่ ถ้าหากว่ามี การวัดซ้ำอีกหลาย ๆ ครั้งจะได้ผลเหมือนเดิมหรือไม่

การวัดบุคคลและงาน ในแง่ของผลงานที่หน่วยงานหนึ่ง ๆ ต้องการ ยังไม่มีเครื่องมือทางวิทยาศาสตร์ที่เชื่อถือได้ (scientific apparatus) การวัดในเรื่องที่มีเครื่องมือวัดแน่นอน เช่น เทอร์โมมิเตอร์ หรือ เครื่องชั่งตวงก็ยังมีปัญหายุ่งยากในเรื่องความเชื่อมั่น (reliability) ฉะนั้น การวัดบุคคล และงานในหน่วยงาน จึงย่อมมีปัญหาในเรื่องนี้ที่ถุกุณขึ้นไปอีก ในที่นี้จะยกตัวอย่างง่าย ๆ ให้เห็นว่า แม้การวัดที่มีเครื่องมือวัดแน่นอนแล้ว ก็ยังมีปัญหาในเรื่องความเชื่อมั่นมากมายเพียงใด ถ้าลองชั่งน้ำหนักบุคคลคนหนึ่งวันละหนึ่งครั้งเป็นเวลาสองวัน ด้วยเครื่องชั่งที่ไวเครื่องเดียวกันในเวลาตรงกัน จะพบว่าน้ำหนักที่ชั่งทั้งสองวันนั้นไม่ตรงกันทีเดียว ถ้าลองชั่งบุคคลอื่น ๆ ด้วยวิธีเดียวกันนี้ ก็จะพบว่า ความแตกต่างระหว่างน้ำหนักที่ชั่งทั้งสองครั้งของแต่ละคนจะมีมากน้อยต่าง ๆ กันไป ความแตกต่างของผลที่ชั่งทั้งสองครั้งของแต่ละคนมีสาเหตุหลายประการ บางคนอาจเพิ่งรับประทานอาหารอิ่มก่อนชั่งในครั้งแรก และรับประทานไม่อิ่มในการชั่งครั้งที่สอง บางคนอาจใส่เสื้อผ้าหนาในการชั่งครั้งหนึ่ง และใส่เสื้อผ้าบางในการชั่งอีกครั้งหนึ่ง สาเหตุต่าง ๆ ที่ไม่เหมือนกันนี้ เป็นสาเหตุที่ทำให้ผลการวัดน้ำหนักต่างกันออกไป (chance errors) ผลผิดพลาดของการวัดในกรณีเช่นน้ำหนักนี้ จะมีปริมาณน้อยมาก เพราะมีเครื่องมือวัดที่แน่นอน น้ำหนักที่ต่างไปอาจเป็นเพียง

เศษหนึ่งส่วนยี่สิบกิโกรัม หรือเศษหนึ่งส่วนสิบกิโกรัมเท่านั้น แต่ถ้าลองให้เพื่อนฝูงของบุคคลหนึ่ง ๆ ประมาณน้ำหนักของบุคคลนั้นด้วยสายตา ความแตกต่างของผลการประมาณน้ำหนักจะยังมีมากขึ้นเป็นทวีคูณ

วิธีการวิเคราะห์ความเชื่อมั่นมีสองวิธี วิธีแรกวิเคราะห์ความแปรปรวน (variance) ของผลการวัดหลาย ๆ ครั้งในการวัดสิ่งหนึ่ง ๆ วิธีที่สองวิเคราะห์สัมประสิทธิ์ของสหสัมพันธ์ (coefficient correlation) ระหว่างผลการวัดของบุคคลหลาย ๆ คนในกลุ่มเดียวกันในการวัดหลาย ๆ ครั้ง

ค่าผิดพลาดมาตรฐานของการวัด (standard error measurements)

การวิเคราะห์ความเชื่อมั่นตามวิธีแรก ถือว่า การวัดที่เชื่อมั่นได้จะต้องให้ผลที่แม่นยำ แต่ผลของการวัดล้วนแต่ขาดความแม่นยำมาบ้างน้อยบ้าง ดังในตัวอย่างในย่อหน้าต่อไป กล่าวอีกนัยหนึ่งก็คือ ผลของการวัดจะต้องมีข้อผิดพลาด ฉะนั้น ถ้าจะตั้งเป็นสมการ ก็กล่าวได้ว่าผลของการวัดมีค่าเท่ากับผลที่ถูกต้องในการวัด รวมกับผลที่ผิดพลาดในการวัด ดังสมการข้างล่างนี้¹

$$X_t = X_\infty + X_e$$

ในที่นี้

- X_t = ผลที่วัดได้ทั้งหมด
- X_∞ = ผลที่ถูกต้องในการวัด
- X_e = ผลที่ผิดพลาดในการวัด

ผลที่ถูกต้องในการวัด (true component) อาจอธิบายได้เป็นหลายลักษณะ เช่น อาจถือได้ว่า เป็นผลที่จะวัดบุคคลหนึ่ง ๆ ได้ในโอกาสที่มีเครื่องมือวัดที่แม่นยำที่สุด หรืออาจถือได้ว่า เป็นค่าเฉลี่ย (means) ของผลที่วัดได้ทั้งหมดในการวัดคน ๆ หนึ่ง เป็นจำนวนมากมายหลาย ๆ ครั้ง ในความเป็นจริง การวัดบุคคลหนึ่ง ๆ ไม่ว่าในเรื่องใดก็ตาม ก็ไม่สามารถวัดได้เป็นเส้นเป็นลันครั้ง เพราะฉะนั้น ค่าเฉลี่ยของผลที่วัดได้ทั้งหมดจากการวัดมากมายหลาย ๆ ครั้งนั้น ย่อมหาค่าจริง ๆ ออกมาไม่ได้ ส่วนผลที่ผิดพลาดในการวัด เป็นค่าเพิ่มหรือลด (positive and negative increment) ซึ่งขึ้นอยู่กับสภาพการณ์และโอกาสของการวัดบุคคลหนึ่ง ๆ หลาย ๆ ครั้ง สาเหตุที่ทำให้เกิดผลผิดพลาดในการวัดนี้ บางครั้งก็อาจจะทราบ เช่นในตัวอย่างเรื่องการวัดน้ำหนัก แต่บางครั้งก็ไม่มีทางทราบถึงสาเหตุที่ทำให้เกิดผลผิดพลาดในการวัดได้

¹ Kerlinger, F.N., *Foundations of Behavioral Research*, p. 433.

ในการวิเคราะห์ความแม่นยำของผลการวัดเพื่อให้ทราบความเชื่อมั่นนั้น ได้มีผู้ตั้งสมมติฐานซึ่งมีผู้ยอมรับไว้พิจารณาและควรมานำมากล่าวในที่นี้สองประการ คือ² สมมติฐานแรกถือว่า เนื่องจากผลที่ผิดพลาดในการวัดมีทั้งค่าบวกและค่าลบ เพราะฉะนั้นถ้าหากวัดกับประชากรจำนวนมากๆ ค่าเฉลี่ยของผลที่ผิดพลาดในการวัดนี้จะเท่ากับศูนย์ ดังสมการข้างล่างนี้

$$M_e = 0$$

$$M_e = \text{ค่าเฉลี่ยของผลผิดพลาดในการวัด}$$

สมมติฐานที่สองถือว่า ในการวัดประชากรจำนวนมาก ๆ ผลที่ผิดพลาดในการวัดจะไม่มีสัมประสิทธิ์ของสหสัมพันธ์กับผลที่ถูกต้องในการวัด กล่าวอีกนัยหนึ่งก็คือ ผู้ที่มีผลการวัดได้สูง ไม่จำเป็นจะต้องมีค่าของผลที่ผิดพลาดในการวัดสูงหรือต่ำตามไปด้วย ดังสมการต่อไปนี้³

$$r_{oe} = 0$$

$$r_{oe} = \text{ค่าสหสัมพันธ์ระหว่างผลที่ถูกต้องกับผลที่ผิดพลาดในการวัด}$$

จากสมมติฐานทั้งสองนี้ ผู้วิเคราะห์ความเชื่อมั่นของการวัด อาจตั้งสมการให้กว้างขวางต่อไปได้อีก จากกฎที่ว่า ค่าเฉลี่ยของจำนวนรวมของส่วนต่างๆ ที่ยังไม่ได้อธิบาย (sum of unweighted components) มีค่าเท่ากับจำนวนรวมของค่าเฉลี่ยทั้งหมด ดังสมการ

$$M_t = M_\infty + M_e$$

$$M_t = \text{ค่าเฉลี่ยของผลที่วัดได้ทั้งหมด}$$

$$M_\infty = \text{ค่าเฉลี่ยของผลที่ถูกต้องในการวัด}$$

$$M_e = \text{ค่าเฉลี่ยของผลที่ผิดพลาดในการวัด}$$

$$\text{แต่ } M_e = 0 \text{ ตามสมมติฐานแรก}$$

ฉะนั้น เมื่อแทนค่า M_e ลงไป จึงได้สมการ

$$M_t = M_\infty$$

² Guilford, J.P., *Psychometric Methods*, pp. 349–350.

³ *Ibid.*, p. 350.

ในกรณีที่ไม่ได้วัดหลาย ๆ ครั้งจริงๆ ค่าเฉลี่ยของผลผิดพลาดในการวัดอาจไม่เท่ากับศูนย์ก็ได้ ในกรณีนี้ ค่าเฉลี่ยของผลที่วัดได้ทั้งหมดก็ย่อมไม่เท่ากับค่าเฉลี่ยของผลที่ถูกต้องในการวัด

ในทำนองเดียวกัน ค่าความแปรปรวนของจำนวนรวมของส่วนต่าง ๆ ที่ยังไม่ได้อธิบาย ซึ่งเท่ากับจำนวนรวมของค่าเฉลี่ยทั้งหมดดังสมการ⁴

$$\sigma_t^2 = \sigma_\infty^2 + \sigma_e^2 + 2r_{\infty e} \sigma_\infty \sigma_e$$

$$\sigma_t^2 = \text{ค่าความแปรปรวนของผลที่วัดได้ทั้งหมด}$$

$$\sigma_\infty^2 = \text{ค่าความแปรปรวนของผลที่ถูกต้องในการวัด}$$

$$\sigma_e^2 = \text{ค่าความแปรปรวนของผลที่ผิดพลาดในการวัด}$$

$$2r_{\infty e} \sigma_\infty \sigma_e = \text{ค่าความแปรปรวนสัมพันธ์ (covariance)}$$

แต่สมมติฐานที่สองระบุว่า ผลที่วัดได้ทั้งหมดไม่มีค่าสหสัมพันธ์กับผลที่ถูกต้องในการวัด ฉะนั้น เมื่อแทนค่า $r_{\infty e}$ จะเปลี่ยนสมการข้างต้นเป็น

$$\sigma_t^2 = \sigma_\infty^2 + \sigma_e^2$$

$$\text{หรือ} \quad \sigma_\infty^2 = \sigma_t^2 - \sigma_e^2$$

การวิเคราะห์ความเชื่อมั่นโดยถือความแม่นยำของผลการวัด กล่าวได้อีกอย่างหนึ่งว่า เป็นการวิเคราะห์ความเชื่อมั่นโดยถือหลักว่า ถ้าค่าความแปรปรวนของผลที่วัดได้ทั้งหมดใกล้เคียงกับความแปรปรวนของผลที่ถูกต้องในการวัดมากเท่าใด ผลที่วัดได้นั้นก็ยิ่งเชื่อมั่นได้มากขึ้นเท่านั้น⁵ หรือถ้าจะแปลความหมายนี้ออกเป็นสมการจะได้ดังข้างล่างนี้

$$r_{tt} = \frac{\sigma_\infty^2}{\sigma_t^2}$$

$$r_{tt} = \text{สัมประสิทธิ์ของความเชื่อมั่น}$$

$$\sigma_\infty^2 = \text{ค่าความแปรปรวนของผลที่ถูกต้องในการวัด}$$

$$\sigma_t^2 = \text{ค่าความแปรปรวนของผลที่วัดได้ทั้งหมด}$$

⁴ Kerlinger, *op. cit.*, p. 231.

⁵ Guilford, *op. cit.*, p. 350.

Guion, R.M., *Personnel Testing*, pp. 34-40.

แต่เนื่องจากในการวัดจริง ๆ ย่อมไม่มีผู้ใดทราบค่าความแปรปรวนของผลที่ถูกต้องในการวัด ฉะนั้นในการหาค่าความเชื่อมั่น จึงจำเป็นต้องเอาสมการที่ได้ตามสมมติฐานที่ส่งมาแทนค่า

$$\begin{aligned} \text{ฉะนั้น} \quad r_{tt} &= \frac{(\sigma_t^2 - \sigma_e^2)}{\sigma_t^2} \\ &= 1 - \frac{\sigma_e^2}{\sigma_t^2} \end{aligned}$$

$$\text{หรือ} \quad \sigma_e^2 = \sigma_t^2 (1 - r_{tt})$$

$$\text{ฉะนั้น} \quad \sigma_e = \sigma_t \sqrt{1 - r_{tt}}$$

สมการสุดท้ายที่ได้มานี้ จะมีประโยชน์มากในการพิจารณาค่าความเชื่อมั่น เพราะค่าความเชื่อมั่นตามทฤษฎีที่ตั้งไว้นั้น เป็นค่าที่จะหาไม่ได้จริงๆ ค่าของสมการนี้เรียกว่าค่าผิดพลาดมาตรฐานของการวัด ซึ่งหมายถึง ค่าเบี่ยงเบนที่ผิดพลาดของผลที่วัดได้มานั่นเอง ค่านี้จะเป็นเครื่องชี้ให้เห็นว่า ผลที่วัดได้มาแต่ละค่า นั้น มีความผิดพลาดมากน้อยเพียงใด ถ้าจะลองนึกเทียบกับตัวอย่าง อาจจะเปรียบเทียบได้กับการที่ซึ่งน้ำหนักบุคคลผู้หนึ่งหลายร้อยครั้ง ผลน้ำหนักที่ชั่งได้ทุกครั้ง จะแสดงการแจกแจงของความถี่ (frequency distribution) ที่แทนค่าน้ำหนักของบุคคลผู้นั้น ค่าน้ำหนักเหล่านี้จะมีค่าเฉลี่ย ซึ่งแสดงว่า บุคคลนั้นมีน้ำหนักที่แท้จริงประมาณเท่าใด ค่าผิดพลาดมาตรฐานของการวัด ก็คือค่าเบี่ยงเบนมาตรฐาน (standard deviation) ของค่าเฉลี่ยของผลการวัดในเรื่องนี้

ในการวัดครั้งหนึ่ง ๆ ที่ให้ผลออกมาและสามารถประมาณค่าความเชื่อมั่น ด้วยวิธีการซึ่งจะกล่าวต่อไปในเรื่องของสัมประสิทธิ์ของความเชื่อมั่น ผู้วัดจะต้องคำนึงถึงความหมายของผลที่วัดได้ว่าเชื่อมั่นได้เพียงใด โดยอาศัยค่าผิดพลาดมาตรฐานของการวัดประกอบการพิจารณา ยกตัวอย่างเช่น ในการวัดเรื่องหนึ่งปรากฏว่า ได้ผลการวัดของบุคคลผู้หนึ่งออกเป็นค่าคะแนน 120 แต้ม และเครื่องมือที่ใช้วัดนั้นพิสูจน์แล้วว่า มีสัมประสิทธิ์ของความเชื่อมั่น 0.96 และมีค่าเบี่ยงเบนมาตรฐาน 10 ฉะนั้น เอาค่าที่ได้ขึ้นไปแทนในสมการค่าผิดพลาดมาตรฐานของการวัด

$$\begin{aligned}
 \sigma_e &= \sigma_t \sqrt{1 - r_{tt}} \\
 &= 10 \sqrt{1 - 0.96} \\
 &= 2
 \end{aligned}$$

ค่าผิดพลาดมาตรฐานของการวัดที่ได้ 2 นี้ หมายความว่า ผลการวัดบุคคลแต่ละคนที่ถูกวัดด้วยเครื่องมือนี้ จะมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 2 นั่นคือโอกาสที่ผลการวัดของบุคคลนี้จะเบี่ยงเบนไปหนึ่งค่าเบี่ยงเบนมาตรฐานจะเพิ่มหรือลดลงสองแต้ม และตามค่าเบี่ยงเบนมาตรฐานของเส้นโค้งปกติ (normal curve) ซึ่งระบุว่า จะมีกรณี 34.1 เปอร์เซ็นต์ที่ตกอยู่ในหนึ่งค่าเบี่ยงเบนมาตรฐาน ฉะนั้น โอกาสที่ผลการวัดของบุคคลนี้จะเพิ่มหรือลดภายในสองแต้ม จะมีค่าเท่ากับ 68.2 เปอร์เซ็นต์ หรือกล่าวในทำนองกลับกัน เพื่อให้สะดวกแก่การพิจารณา จะมีโอกาส 31.8 เปอร์เซ็นต์ที่ผลการวัดของบุคคลนี้จะผิดพลาดถึงสองแต้ม หรือหมายความว่า จะมีโอกาสประมาณ 1 ใน 3 ที่ผลการวัดของบุคคลนี้ หรือบุคคลคนหนึ่งคนใดก็ตามที่ถูกวัดด้วยเครื่องมือนี้จะผิดพลาดไปถึงสองแต้ม ในทำนองเดียวกันก็อาจจะคิดต่อไปได้ถึงสองค่าเบี่ยงเบนมาตรฐาน (± 47.7 เปอร์เซ็นต์) ซึ่งหมายความว่า จะมีโอกาสประมาณ 1 ใน 20 (4.6 เปอร์เซ็นต์) ที่ผลการวัดของบุคคลที่ถูกวัดด้วยเครื่องมือนี้จะผิดพลาดไป 4 แต้ม ในกรณีของตัวอย่างที่ยกมาว่า วัดได้ 120 แต้ม ก็ตีความหมายได้ว่า จะมีโอกาส 1 ใน 3 ที่บุคคลผู้นั้นจะได้คะแนนผิดพลาดไปจากคะแนนจริงๆ ของตน มากหรือน้อยไปสองคะแนน คะแนน 120 แต้มที่ได้นี้ ถือเป็นค่าเฉลี่ยของผลการวัดบุคคลนี้ ซึ่งนับได้ว่ามีค่าใกล้เคียงกับคะแนนจริงที่ควรจะได้ จึงกล่าวได้ว่า จะมีโอกาสประมาณ 1 ใน 3 ที่คะแนนจริงของบุคคลผู้นั้นควรจะผิดไปจาก 118–122 แต้ม

ค่าผิดพลาดมาตรฐานของการวัด เป็นค่าที่มีประโยชน์มาก เมื่อจะมีการแปลความหมายของผลที่วัดได้ ผู้วัดไม่ควรจะถือแต่เพียงปริมาณความมากน้อยของค่าคะแนนเท่านั้น แต่จะต้องพิจารณาโอกาสที่จะผิดพลาดของค่าคะแนน โดยอาศัยค่าผิดพลาดมาตรฐานของการวัดด้วย แบบทดสอบมาตรฐานที่ใช้กันอยู่ในท้องตลาดจะระบุค่าความเชื่อมั่นไว้ จึงเป็นหน้าที่ของผู้วัดจะต้องนำมาพิจารณาหาค่าผิดพลาดมาตรฐานของการวัดต่อไป ค่าความเชื่อมั่นของเครื่องมือวัดจะมีค่าสหสัมพันธ์เชิงนิเสธกับค่าผิดพลาดมาตรฐาน ดังตารางที่ 8

ตารางที่ 8

ค่าผิดพลาดมาตรฐานของการวัดเทียบกับค่าความเชื่อมั่น*

สัมประสิทธิ์ความเชื่อมั่น	ค่าผิดพลาดมาตรฐาน
.50	.71 σ_t
.60	.63 σ_t
.70	.55 σ_t
.80	.45 σ_t
.85	.38 σ_t
.90	.32 σ_t
.95	.22 σ_t
.98	.14 σ_t

*แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 133.

จากตารางข้างบนนี้ จะเห็นได้ว่า ปริมาณค่าผิดพลาดมาตรฐานจะลดลงเมื่อค่าความเชื่อมั่นเพิ่มขึ้น แต่ถึงแม้ว่า ค่าความเชื่อมั่นสูงถึงขนาด .95 หรือ .98 ค่าผิดพลาดมาตรฐานก็ยังมิได้น้อย ฉะนั้น เมื่อผู้วัดได้ใช้เครื่องมือวัดที่มีค่าความเชื่อมั่นสูง ก็ควรจะพึงระลึกลึกถึงความจริงข้อนี้ด้วย

สัมประสิทธิ์ของความเชื่อมั่น (reliability coefficient)

การวิเคราะห์ความเชื่อมั่นตามวิธีที่สอง ซึ่งถือหลักพิจารณาสัมประสิทธิ์ของสหสัมพันธ์ระหว่างผลการวัดของบุคคลหลายคนในกลุ่มเดียวกัน ในการวัดหลายๆ ครั้ง จะทำให้ได้ค่าสัมประสิทธิ์ของความเชื่อมั่น แนวคิดในเรื่องนี้ถือว่า การวัดที่เชื่อมั่นได้จะต้องแสดงลำดับที่ของบุคคลแต่ละคนในกลุ่มได้ถูกต้อง การคิดเช่นนั้นเป็นการเลี่ยงการหาค่าความเชื่อมั่นจริงๆ เพราะได้กล่าวมาแล้วว่าค่าความเชื่อมั่นตามสูตร $r_{tt} = \frac{\sigma_{\infty}^2}{\sigma_t^2}$ นั้นคำนวณไม่ได้ เพราะไม่มีผู้ใดทราบ σ_{∞}^2 ที่แท้จริง แนวคิดที่ถือว่าการวัดที่สามารถวัด

ลำดับที่ของบุคคลในกลุ่มได้ถูกต้อง เป็นการวัดที่เชื่อมั่นได้ ทำให้สามารถคำนวณค่าความเชื่อมั่นจากสัมประสิทธิ์ของสหสัมพันธ์ในกรณีต่างๆ ได้ สัมประสิทธิ์ของความเชื่อมั่นที่คำนวณได้โดยอาศัยแนวคิดนี้ แสดงว่าเครื่องมือวัดนั้นๆ สามารถวัดซ้ำๆ กันหลายๆ ครั้ง โดยได้ผลในเรื่องลำดับที่ ในกลุ่มของบุคคลที่วัดได้ถูกต้อง กล่าวอีกนัยหนึ่งก็คือ ถ้าค่าสัมประสิทธิ์ของสหสัมพันธ์ยิ่งสูงขึ้นเท่าใด โอกาสที่เครื่องมือวัดนั้นจะแสดงลำดับที่ของบุคคลหนึ่งๆ ในกลุ่ม ก็จะถูกต้องแน่นอนยิ่งขึ้นเท่านั้น

เครื่องมือวัดหนึ่งๆ จะเชื่อมั่นได้ก็ต่อเมื่อวัดให้แสดงตำแหน่งที่ของบุคคลแต่ละคนในกลุ่มอย่างใกล้เคียงกันทุกครั้ง ถ้ายังมีความเชื่อมั่นสูง ค่าของค่าผิดพลาดมาตรฐานก็ยิ่งน้อยลงดังได้กล่าวมาแล้ว การที่จะหาค่าความเชื่อมั่นของเครื่องมือวัดแต่ละอันจึงเป็นเรื่องสำคัญ นักทฤษฎีการวัดได้คิดวิธีที่จะหาสัมประสิทธิ์ของความเชื่อมั่นไว้หลายวิธี วิธีที่สำคัญและใช้กันแพร่หลายที่จะนำมากล่าวในที่นี้มี 4 วิธี คือ

1. วิธีทดลองเครื่องมือวัดซ้ำ (test-retest method)
2. วิธีเปรียบเทียบเครื่องมือวัดนั้นๆ กับเครื่องมือวัดอื่น ๆ (parallel test forms)
3. วิธีแบ่งครึ่งเครื่องมือวัด (split-half method)
4. วิธีวิเคราะห์รายการ (item statistic)

วิธีทดลองเครื่องมือวัดซ้ำ วิธีนี้เป็นวิธีที่จะหาสัมประสิทธิ์ของความคงตัว (coefficient of stability) ถ้าลองนึกเทียบกับตัวอย่างเรื่องการชั่งน้ำหนักที่เคยกกล่าวมา การหาความเชื่อมั่นด้วยวิธีนี้ก็ คือ การนำเครื่องมือวัดชั่งน้ำหนักไปชั่งบุคคลหนึ่งๆ ไม่น้อยกว่าสองครั้ง ในกรณีเช่นนี้ ผู้ที่ทำการชั่งน้ำหนักในสองครั้งนั้นไม่ควรจะเป็นคนๆ เดียวกันเพื่อจะได้ไม่ทำให้ผลการชั่งครั้งแรกที่จำได้ ไปมีอิทธิพลต่อผลการชั่งในครั้งที่สอง การชั่งครั้งที่สองอาจจะทำติดต่อกับครั้งแรกได้ ถ้าหากต้องการทราบแต่เพียงว่า พิจารณาชั่งน้ำหนักนั้นๆ มีความแม่นยำเพียงใด เพราะว่าการชั่งติดๆ กันจะทำให้ทราบว่าผลน้ำหนักที่ชั่งได้ทั้งสองครั้งต่างกันนั้น เนื่องจากพิจารณาในการวัดมากกว่าความเปลี่ยนแปลงในเรื่องน้ำหนักจริง ๆ ของตัวบุคคล แต่ถ้าหากต้องการทราบว่า ผลน้ำหนักของบุคคลในวาระหนึ่งๆ สัมพันธ์กับผลน้ำหนักของบุคคลนั้นในวาระอื่น ๆ อย่างไร การชั่งน้ำหนักครั้งที่สองก็ควรทิ้งระยะจากครั้งที่หนึ่ง เพื่อให้ความเปลี่ยนแปลงของบุคคลมีโอกาสนำมามีส่วนในน้ำหนักที่เปลี่ยนไปในครั้งที่สอง การชั่งหรือวัดเช่นนั้น ทำให้ทราบความผันแปรทั้งส่วนที่เนื่องมาจากตัวบุคคล และทั้งส่วนที่เนื่องมาจากพิจารณาในการวัด ในการวัดเรื่องต่างๆ เกี่ยวกับบุคคลเพื่อประโยชน์ทางจิตวิทยาและการบริหาร จำเป็นที่

จะต้องคำนึงถึงความผันแปรทั้งสองอย่างนี้ จึงได้มีการยอมรับว่า เมื่อจะหาค่าความเชื่อมั่นด้วยวิธีนี้ จะต้องมีการทิ้งระยะการทดลองเครื่องมือวัดในครั้งที่สองพอสมควร ทั้งนี้ เพราะการวัดเช่นนี้ได้มีลักษณะเป็นเช่นเดียวกับการวัดลักษณะต่าง ๆ ของร่างกายตามธรรมดา ในการวัดความสูง ความเตี้ย หรือความเร็วในการแสดงปฏิกิริยาของกล้ามเนื้อ ผู้วัดอาจไม่จำเป็นต้องสนใจว่าบุคคลนั้น ๆ มีความผันแปรในเรื่องที่วัดอย่างไร แต่ถ้าสมมติว่า ผู้วัดต้องการวัดว่า เจ้าหน้าที่ที่ธุรการคนหนึ่งสามารถอ่านหนังสือได้เข้าใจเพียงใด และผู้วัดนั้นนำเครื่องมือวัดที่จะหาค่าความเชื่อมั่นมาทดลองกับกลุ่มเจ้าหน้าที่ธุรการที่คัดเลือกไว้แล้วด้วยวิธีการสุ่มตัวอย่าง โดยให้เจ้าหน้าที่ธุรการเหล่านั้นแสดงความสามารถในการอ่านข้อความที่คัดเลือกไว้แล้วว่าเป็นเครื่องมือวัด เป็นจำนวนสองครั้งต่อเนื่องกัน ผลที่วัดได้ในการอ่านครั้งที่สองจะไม่แสดงความสามารถจริง ๆ ที่เจ้าหน้าที่ธุรการเหล่านั้นมีอยู่ก่อนการใช้เครื่องมือวัดทั้งสองครั้ง เช่น ผลคะแนนที่วัดได้ในครั้งที่สองอาจสูงขึ้นเพราะผู้ถูกวัดจำข้อความที่อ่านได้จากกรวัดครั้งแรก กล่าวอีกนัยหนึ่งคือ เนื้อหาของเครื่องมือวัดที่ใช้ในการวัดครั้งที่สอง โดยไม่ได้ทิ้งระยะกับครั้งที่หนึ่งในกรณีนี้ จะไม่มีคุณสมบัติเหมือนกับเนื้อหาของเครื่องมือวัดที่ใช้ในครั้งแรก

วิธีหาค่าความเชื่อมั่นด้วยการทดลองเครื่องมือวัดซ้ำนี้ เหมาะแก่เครื่องมือวัดชนิดระบบให้คะแนนโดยผู้อื่น (rating procedure) หรือในตำราการบริหารงานบุคคลส่วนใหญ่เรียกว่า การให้คะแนนตามระบบคุณธรรม (merit rating system)⁶ อย่างไรก็ดี เนื่องจากยังไม่ได้มีการพิสูจน์ว่า การทดลองเครื่องมือวัดซ้ำในครั้งที่สองนั้น ควรทิ้งระยะห่างจากครั้งแรกนานเท่าใด จึงจะทำให้เกิดผลเสียตามที่กล่าวมาแล้วน้อยที่สุด ฉะนั้นค่าความเชื่อมั่นที่ทำได้ตามวิธีนี้ จึงถือกันว่ามีคุณภาพน้อยมาก โดยเฉพาะในกรณีที่เครื่องมือวัดเป็นชนิดระบบให้คะแนนโดยผู้อื่น โอกาสที่ผู้ให้คะแนนจะจำผลคะแนนครั้งแรกได้ในระหว่างที่ผู้ให้คะแนนครั้งที่สองนั้นมีมาก

สัมประสิทธิ์ของความเชื่อมั่นที่ทำได้ตามวิธีก็คือ ค่าสหสัมพันธ์ระหว่างผลการวัดครั้งแรกกับครั้งที่สอง ฉะนั้น ค่าสัมประสิทธิ์ของความเชื่อมั่นจึงมีลักษณะเหมือนกับลักษณะทั่วไปของค่าสหสัมพันธ์ทุกประการ คือ มีค่าตั้งแต่ 1.00 ไปจนถึง -1.00

วิธีเปรียบเทียบเครื่องมือวัดนั้น ๆ กับเครื่องมือวัดอื่น ๆ วิธีหาค่าความเชื่อมั่นวิธีที่สองซึ่งคำนวณได้โดยทดลองเครื่องมือวัดนั้น ๆ กับเครื่องมือวัดอื่น ๆ มีลักษณะเป็นการหาสัมประสิทธิ์ของความเสมอเหมือนกัน (coefficient of equivalence) การที่นักทฤษฎี

⁶ *Ibid.*, pp. 40-41.

การวัดได้วางหลักการหาความเชื่อมั่นตามวิธีชนกก็เพราะว่า งานหรือเนื้อหาที่กำหนดไว้ในเครื่องมือวัด เป็นเพียงส่วนเล็กน้อยของงานที่จะต้องปฏิบัติจริงๆ ทั้งหมด หรือพฤติกรรมจริงๆ ทั้งหมดของบุคคลที่ถูกวัด งานหรือพฤติกรรมที่ปรากฏในเนื้อหาของเครื่องมือวัดไม่จำเป็นที่จะต้องแทนลักษณะจริงๆ ของงานทั้งหมด หรือพฤติกรรมทุกอย่างของบุคคลที่จะถูกวัดได้ ฉะนั้น หากจะมีการพิสูจน์ว่า เครื่องมือนั้นมีความเชื่อมั่นอย่างแท้จริง ก็จำเป็นจะต้องนำเครื่องมือวัดนั้นไปเทียบกับเครื่องมือวัดอื่นๆ ที่เชื่อมั่นได้แล้ว เพื่อแสดงให้เห็นว่า เครื่องมือวัดที่สร้างใหม่ สามารถวัดได้โดยมีความเชื่อมั่นตามส่วนของความเชื่อมั่นของเครื่องมือที่ใช้เทียบนั้นๆ แต่โดยความเป็นจริงแล้ว เครื่องมือวัดลักษณะบุคคลที่มีความเชื่อมั่นอันจะใช้เป็นเครื่องเทียบครั้งแรกนั้นไม่มี ฉะนั้น ในการหาความเชื่อมั่นตามวิธีชนก จึงดำเนินการโดยอาศัยสมมติฐานที่ว่า ถ้าเนื้อหาของเครื่องมือวัดสองอย่างเป็นตัวแทนของงานหรือพฤติกรรมของบุคคลมากเท่าใด เนื้อหาของเครื่องมือวัดทั้งสองนั้นจะสัมพันธ์กันมากเท่านั้น⁷ จากสมมติฐานนี้ ผู้สร้างเครื่องมือวัดก็ต้องพยายามหาความสัมพันธ์ของสหสัมพันธ์ระหว่างผลการวัดโดยเครื่องมือทั้งสองนี้ สัมประสิทธิ์ของความเชื่อมั่นที่ได้โดยวิธีชนกเป็นค่าที่ถือว่าแสดงความเป็นตัวแทนของงาน หรือพฤติกรรมของบุคคลทั้งหมดของเนื้อหาในเครื่องมือวัดนั้นๆ โดยทางทฤษฎี เครื่องมือวัดที่จะถือว่ามีความเสมอเหมือนกัน จะต้องมีความสัมพันธ์ 3 ประการ คือ ประการแรก มีปริมาณของรายการที่จะวัดในประเด็นหนึ่งๆ เท่าเทียมกัน ประการที่สอง ค่าเฉลี่ยของผลเครื่องมือวัดจะไม่ผิดเพี้ยนกันโดยมีนัยสำคัญ และประการสุดท้าย ค่าเบี่ยงเบนมาตรฐานจะต้องใกล้เคียงกัน

เครื่องมือวัดสองอันที่นำมาเทียบค่าความเสมอเหมือนกันนั้น จะมีเป้าหมายที่จะวัดเหมือนกัน แต่เนื้อหาที่จะเป็นตัวแทนของงานหรือพฤติกรรมที่จะวัด จะต้องต่างกัน ถ้าหากว่าประสงค์จะสร้างเครื่องมือวัดความรู้ระดับประจำแผนก โดยพยายามให้เครื่องมือวัดนั้นมีความเชื่อมั่นได้ ผู้สร้างก็ต้องมีเครื่องมือวัดสองอันที่บรรจุเนื้อหาที่เป็นตัวแทนของความรู้ในเรื่องนี้ คำถามหรือแนวที่จะวัดในเครื่องมือวัด จะต้องมีความคล้ายคลึงกัน มีความยากง่ายทัดเทียมกัน มีการให้ความสำคัญของประเด็นความรู้เหมือนกัน แต่ข้อความในเนื้อหาของเครื่องมือวัดแต่ละอันจะไม่ซ้ำกัน ผลของการทดสอบวัดโดยใช้เครื่องมือทั้งสองอันนี้ เมื่อนำมาหาค่าสหสัมพันธ์แล้ว จะถือได้ว่าเป็นค่าความเชื่อมั่นแบบสัมประสิทธิ์ของความเสมอเหมือนกัน

⁷ Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, pp. 127-128.

ในทางปฏิบัติ ผู้สร้างเครื่องมือวัดอาจประสบความลำบากที่จะหาเครื่องมือวัดที่มีลักษณะคล้ายคลึงกันตามที่กล่าวมา บางครั้งเมื่อได้สร้างเครื่องมือวัดขึ้นมาแล้ว อาจหาเครื่องมือวัดอีกอันหนึ่งมาทดลองเทียบไม่ได้ บางครั้งเมื่อมีเครื่องมือวัดที่จะเทียบ ผู้สร้างเครื่องมือวัดก็ไม่มีโอกาสที่จะทดลองเครื่องมือวัดทั้งสองอันนี้ให้ทั้งระยะกันได้ เพราะสภาพการณ์ไม่อำนวย เช่น ผู้รับการทดลองไม่สามารถรอเพื่อการทดลองซ้ำได้ ในกรณีที่ไม่ได้ทั้งระยะการทดลองวัดเครื่องมือวัดทั้งสองนี้ ก็จะมีผลเสียเช่นเดียวกับในกรณีการหาค่าสัมประสิทธิ์ของความคงตัว

สัมประสิทธิ์ของความเสมอเหมือนกัน ยากที่จะมีค่าสูง เพราะเครื่องมือวัดทั้งสองอันต่างก็เป็นเพียงตัวแทนของงานที่จะทำ หรือพฤติกรรมของบุคคลทั้งหมด ในกรณีที่สิ่งที่จะวัดนั้นมีขอบเขตแน่นอน เช่นในเรื่องความสัมฤทธิ์ผล (achievement) การหาค่าความเสมอเหมือนกันทำได้ง่ายกว่าการวัดลักษณะของบุคคลบางอย่าง ที่ยังกำหนดขอบเขตไว้ได้ไม่แน่ชัดทีเดียว เช่น ในเรื่องตัวแปรทางบุคลิกภาพเกี่ยวกับแรงกระตุ้น (motivation) หรืออารมณ์ (temperament)

วิธีแบ่งครึ่งเครื่องมือวัด วิธีนี้เป็นวิธีที่แพร่หลายที่สุดในการใช้หาค่าความเชื่อมั่น แนวคิดของวิธีนี้กำหนดว่า เครื่องมือวัดที่ดีจะต้องมีเอกภาพในการวัด (functional unity) กล่าวคือ ส่วนตอนต่างๆ ของเครื่องมือวัดหนึ่งๆ จะต้องมีความสัมพันธ์ซึ่งกันและกัน เพื่อให้เกิดความประสานกันในการที่จะวัดสิ่งๆ เดียวกัน สัมประสิทธิ์ของความเชื่อมั่นที่ทำได้ตามวิธีนี้ บางครั้งก็เรียกว่าสัมประสิทธิ์ของความคงเส้นคงวาภายใน (internal consistency) เครื่องมือวัดที่มีค่าสัมประสิทธิ์นี้ หมายความว่า ข้อความที่ทดสอบแต่ละข้อ หรือส่วนของเครื่องมือวัดแต่ละส่วนจะมีความเสมอเหมือนกันกับข้อความ หรือส่วนอื่นๆ ทุกข้อความหรือทุกส่วน

ผู้สร้างเครื่องมือวัดที่จะหาค่าความเชื่อมั่นตามวิธีนี้ จะต้องนำเครื่องมือวัดของตนมาแบ่งออกเป็นสองส่วนเท่าๆ กัน⁸ การที่จัดข้อความในเครื่องมือวัดออกเป็นส่วนหนึ่งส่วนใดนั้น ขึ้นอยู่กับผลการวิเคราะห์เนื้อหาของข้อความ และความยากง่ายของข้อความแต่ละอัน โดยหลักการกว้างๆ ส่วนสองส่วนของเครื่องมือวัดนี้ จะมีเนื้อหาและความยากง่ายที่ตัดเทียมกัน วิธีที่สะดวกที่สุดก็คือ จัดข้อความที่จะวัดที่มีเลขหมายคี่ และคู่ไว้เป็นคนละพวก เครื่องมือวัดที่จะหาความเชื่อมั่นด้วยวิธีนี้ ควรจะมีข้อความที่จะวัดไม่น้อยกว่า 60 ข้อ เครื่องมือวัดที่แบ่งออกเป็นสองส่วนนี้ มิใช่แบ่งเพื่อจะให้

⁸ Guion, *op. cit.*, pp. 42-43.

การแยกวัดที่ละส่วน แต่ส่วนที่แยกไว้นั้นจะแยกเพื่อแบ่งพวกของผลการวัด สัมประสิทธิ์ของสหสัมพันธ์ระหว่างผลการวัดทั้งสองส่วน จะแสดงค่าความเสมอเหมือนกันระหว่างส่วนทั้งสองของเครื่องมือวัดหนึ่ง ๆ กล่าวอีกนัยหนึ่งคือ แสดงค่าความเชื่อมั่นของเครื่องมือวัดนั้น ๆ แต่ละส่วน ฉะนั้น เพื่อที่จะให้ได้ค่าความเชื่อมั่นของเครื่องมือวัดทั้งอัน จึงได้มีผู้คิดสูตรขึ้น เรียกว่า Spearman-Brown Prophecy Formula⁹

$$r_{11} = \frac{2r_{11}}{1+r_{11}}$$

r_{11} = สัมประสิทธิ์ของความเชื่อมั่น

r_{11} = สัมประสิทธิ์ของสหสัมพันธ์ระหว่างเครื่องมือวัดทั้งสองส่วน

ฉะนั้น หากเครื่องมือวัดหนึ่งที่ถูกแบ่งออกเป็นสองส่วนมีค่าสหสัมพันธ์ระหว่างสองส่วนนั้น = 0.70 ก็อาจหาค่าความเชื่อมั่นของเครื่องมือวัดนี้ได้ โดยแทนค่า

$$\begin{aligned} r_{11} &= \frac{2(0.70)}{1+0.70} \\ &= \frac{1.40}{1.70} \\ &= 0.82 \end{aligned}$$

นอกจากสูตรนี้แล้ว ยังมีผู้คิดสูตรอื่นใช้อีก แต่ไม่แพร่หลายเท่าเช่น Rulon Formula¹⁰

$$r_{11} = 1 - \frac{\sigma_d^2}{\sigma_t^2}$$

r_{11} = สัมประสิทธิ์ของความเชื่อมั่น

d = ความแตกต่างระหว่างผลการวัดบุคคลแต่ละคนระหว่างเครื่องมือวัดสองส่วน

σ_d^2 = ความแปรปรวนของความแตกต่างของผล (d)

σ_t^2 = ความแปรปรวนของผลคะแนนทั้งหมด

การหาค่าความเชื่อมั่นด้วยวิธีแบ่งครึ่งเครื่องมือวัด เป็นวิธีที่สะดวกมาก เพราะไม่จำเป็นต้องอาศัยเครื่องมือวัดอื่น ๆ มาเทียบ หรือไม่จำเป็นต้องรอเวลาเพื่อทดลองวัดซ้ำ

⁹ Thorndike and Hagen, *op. cit.*, p. 129.

¹⁰ Guilford, *op. cit.*, p. 379.

ฉะนั้น จึงเป็นที่นิยมแพร่หลาย แบบทดสอบมาตรฐานส่วนใหญ่ในห้องตลาดมักแสดงค่าความเชื่อมั่นที่หาด้วยวิธีนี้ แต่การหาค่าความเชื่อมั่นด้วยวิธีนี้ก็มักเจออยู่น้อยหลายประการ ซึ่งผู้ใช้แบบทดสอบมาตรฐานเหล่านี้หรือผู้ที่สร้างเครื่องมือวัดควรจะคำนึงถึง ประการแรก ค่าความเชื่อมั่นที่หาได้ด้วยวิธีนี้ไม่ได้วิเคราะห์ความเปลี่ยนแปลงของบุคคลซึ่งจะถูกวัดจริง ๆ เพราะบุคคลที่ถูกทดลองวัดได้แสดงผลงานหรือพฤติกรรมของตน เฉพาะชั่วเวลาที่ทำหรือปฏิบัติตามเครื่องมือวัดเท่านั้น กล่าวอีกอย่างหนึ่งก็คือ ผลของการทดลองเครื่องมือวัดด้วยวิธีนี้ แทนค่าผลงาน หรือพฤติกรรมของบุคคล เฉพาะชั่วระยะเวลาที่ถูกทดลองเท่านั้น มิใช่เป็นผลงานหรือพฤติกรรมจริง ๆ ทั้งหมดที่บุคคลที่ถูกทดลองจะแสดงออก ผลของการทดลองมีค่าเที่ยงตรงเฉพาะชั่วระยะเวลาหนึ่ง ประการที่สอง ในกรณีที่เครื่องมือวัดมีลักษณะเป็นการวัดความเร็วในการทำงานหรือความฉับพลันในการแสดงผลพฤติกรรม (speed test) ค่าความเชื่อมั่นที่คำนวณจากผลการทดลองวัดด้วยวิธีนี้จะไม่แสดงผลการปฏิบัติงาน หรือพฤติกรรมของบุคคลอย่างแท้จริง เพราะความแตกต่างระหว่างผลที่ทดลองวัดได้จากบุคคลต่าง ๆ มีสาเหตุเนื่องมาจากการที่บุคคลที่ถูกทดลองแต่ละคนรีบปฏิบัติงานตามคำสั่งของเครื่องมือวัดมากกว่าที่จะเนื่องมาจากระดับความสามารถในการปฏิบัติงาน หรือระดับความแตกต่างในเรื่องพฤติกรรมของบุคคลเหล่านั้นจริง ๆ แต่อย่างไรก็ดี เครื่องมือวัดบุคคลโดยทั่ว ๆ ไปก็มิใช่จะวัดแต่ความเร็วในการปฏิบัติงานหรือการแสดงผลพฤติกรรมแต่เพียงอย่างเดียว ฉะนั้น จึงอาจกล่าวได้ว่า เครื่องมือวัดที่อาศัยหลักความเร็วในการปฏิบัติงาน หรือความฉับพลันในการแสดงผลพฤติกรรมน้อยเท่าใด ก็จะทำให้ค่าความเชื่อมั่นที่หาได้ด้วยการทดลองวัดตามวิธีนี้ผิดพลาดน้อยลงไปเท่านั้น

วิธีวิเคราะห์รายการ เนื่องจากวิธีแบ่งครึ่งเครื่องมือวัดไม่สามารถคำนวณค่าความคงเส้นคงวาภายในได้อย่างแท้จริง เพราะลักษณะของค่าความเชื่อมั่นที่คำนวณได้โดยวิธีแบ่งครึ่งเครื่องมือวัดมีลักษณะเป็นค่าความเสมอเหมือนกันระหว่างสองส่วนของเครื่องมือวัดมากกว่า ฉะนั้น จึงได้มีผู้คิดวิธีวิเคราะห์ข้อความที่ใช้ทดสอบทุก ๆ อันภายในเครื่องมือวัด เพื่อจะให้ได้ค่าความคงเส้นคงวาภายในได้อย่างแท้จริงเรียกว่า *Kuder-Richardson Formula*¹¹

¹¹ Guion, *op. cit.*, p. 44.

$$r_{11} = \frac{n}{n-1} \left(1 - \frac{\sum pq}{\sigma_t^2} \right)$$

- r_{11} = สัมประสิทธิ์ของความเชื่อมั่น
 n = จำนวนของข้อความที่ใช้ทดสอบทั้งหมดในเครื่องมือวัด
 p = อัตราส่วนของคำตอบที่ถูกต้องของแต่ละข้อความที่ใช้ทดสอบ
 q = อัตราส่วนของคำตอบที่ผิดของแต่ละข้อความที่ใช้ทดสอบ
 σ_t^2 = ค่าความแปรปรวนของผลการวัดทั้งหมด

ถ้าสมมติว่า เครื่องมือวัดหนึ่งมีข้อความที่ใช้ทดสอบเป็นคำถาม 12 คำถาม ต้องการกำหนดค่าความเชื่อมั่นตามสูตร Kuder-Richardson จึงได้นำเครื่องมือวัดนี้ไปทดลองวัดกับบุคคล 10 คน ปรากฏว่า อัตราส่วนของผู้ที่ตอบคำถามถูก และผิดในแต่ละข้อ ตั้งแต่ข้อ 1-12 มีดังนี้ ข้อหนึ่ง ถูกหมด (1.0:0) ข้อสองและข้อสาม ถูก 9 คน ผิด 1 คน (0.9:0.1) ข้อ 4 ถูก 7 คน (0.7:0.3) ข้อ 5 และ 6 ถูก 6 คน (0.6:0.4) ข้อ 7 ถูก 5 คน (0.5:0.5) ข้อ 8 ถูก 4 คน (0.4:0.6) ข้อ 9 และ ข้อ 10 ถูก 3 คน (0.3:0.7) ข้อ 11 ถูก 2 คน (0.2:0.8) ข้อ 12 ถูก 1 คน (0.1:0.9) และปรากฏว่า ผลการวัดทั้งหมดมีค่าความแปรปรวน 9.45

$$\begin{aligned} \text{ฉะนั้น } r_{11} &= \frac{12}{11} \left(1 - \frac{2.03}{9.45} \right) \\ &= 0.857 \end{aligned}$$

ค่าความเชื่อมั่นที่คำนวณโดยวิธีวิเคราะห์รายกรณี ถือว่ามีค่าความคงเส้นคงวาภายในมากกว่าที่จะคำนวณด้วยวิธีอื่น แต่มีข้อแม้ว่า ข้อความที่ใช้ทดสอบทั้งหมดจะต้องมีความคล้ายกันในกลุ่ม กล่าวคือ ทุก ๆ ข้อความที่ใช้ทดสอบจะต้องร่วมประสานกันในการวัดสิ่งใดสิ่งหนึ่งตามความมุ่งหมายของเครื่องมือวัดซึ่งอาจจะเป็นความสัมฤทธิ์ผลหรือบุคลิกภาพก็ได้ อย่างไรก็ตาม ค่าความเชื่อมั่นที่คำนวณได้ด้วยวิธีนี้ก็ยังมิจุดอ่อนในประเด็นที่ว่า ไม่สามารถวัดความสามารถในการทำงาน หรือพฤติกรรมทั้งหมดจริงๆ ของบุคคลได้ และถ้าเป็นการวัดความนับปล้นก็มีโอกาสที่จะทำให้ค่าความเชื่อมั่นที่ได้ผิดพลาดไปเช่นเดียวกับค่าความเชื่อมั่นที่คำนวณได้โดยวิธีแบ่งครึ่งเครื่องมือวัด

เครื่องมือวัดประเภทแบบทดสอบมาตรฐานที่ใช้ในห้องตลาด มีการกำหนดหาค่าความเชื่อมั่นด้วยวิธีต่าง ๆ กัน ดังจะเห็นได้จากตารางดังต่อไปนี้

ตารางที่ 9

วิธีคำนวณสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบมาตรฐานบางแบบ *

ชื่อแบบทดสอบ	วิธี	สัมประ- สิทธิ์
1. CEEB Scholastic Aptitude-Verbal	แบ่งครึ่งเครื่องมือวัด	0.87
2. Kuder Preference Record (interests)	ทดลองเครื่องมือวัดซ้ำ	0.93
3. Adaptability Test (mental ability)	เปรียบเทียบเครื่องมือวัด	0.89
4. Bernreuter, "self-confidence" score	เปรียบเทียบเครื่องมือวัด	0.86
5. Kuhlmann-Anderson (IQ)	เปรียบเทียบเครื่องมือวัด	0.69

* แหล่งที่มา: Cronbach, L.J., *Essentials of Psychological Testing*, p. 72.

ความแปรปรวนที่มีผลกระทบต่อค่าความเชื่อมั่น

ความแปรปรวนในการปฏิบัติงาน หรือในการแสดงพฤติกรรมของบุคคลหนึ่ง ๆ มีผลทำให้ผลที่วัดได้ด้วยเครื่องมือหนึ่ง ๆ แปรปรวนไปด้วย นักทฤษฎีการวัดในปัจจุบันได้เห็นพ้องต้องกันว่า ลักษณะทางจิตวิทยาของบุคคลหลายประการทำให้ผลที่เครื่องมือวัดหนึ่ง ๆ วัดได้แปรปรวนไป ลักษณะที่ทำให้เกิดความแปรปรวนนี้ แยกได้เป็น 5 ประการ¹²

1. ลักษณะนิสัยทั่วไปประจำตัว (lasting and general characteristics)
2. ลักษณะนิสัยเฉพาะตัวประจำตัว (lasting but specific characteristics)
3. ลักษณะนิสัยทั่วไปที่มีเป็นครั้งคราว (temporary but general characteristics)
4. ลักษณะนิสัยเฉพาะตัวที่มีเป็นครั้งคราว (temporary and specific characteristics)
5. ความแปรปรวนอันบังเอิญ (error variance)

ลักษณะนิสัยทั่วไปประจำตัวที่อาจทำให้ผลที่เครื่องมือวัดหนึ่ง ๆ วัดได้แปรปรวนไป ได้แก่ ความ เช่นถ้าจะวัดความสัมฤทธิ์ผลระดับของเขาวัวของแต่ละบุคคลจะทำให้ผลที่วัดได้ไม่ได้แตกต่างกันเพราะความสัมฤทธิ์ผลที่แต่ละคนมีเท่านั้น หรือบางคนอาจจะมีอาการคล่องตัวในการถูกวัด (test-wise) ซึ่งทำให้ผู้นั้นมีผลในการวัดสูงกว่าที่ควรจะเป็น

¹² Lawshe, C.H., and Balma, M.J., *Principles of Personnel Testing*, p. 277.

ลักษณะนิสัยเฉพาะตัวประจำตัวหมายถึง ความสามารถ หรือทักษะที่บุคคลหนึ่ง ๆ มีเฉพาะตัว อันทำให้บุคคลผู้นั้นสามารถปฏิบัติงานหรือแสดงพฤติกรรมที่จะถูกวัดได้ คล่องกว่าผู้อื่น ในกรณีเช่นนั้น ถ้าหากว่าข้อความที่ใช้ในการทดสอบในเครื่องมือวัดมีสาระสำคัญตรงกับลักษณะนิสัยเฉพาะตัวประจำตัวของบุคคลนั้น ก็เป็นที่น่าสงสัยได้ว่า ค่าความเชื่อมั่นที่คำนวณได้ ย่อมผิดไปจากค่าที่ควรจะเป็นจริง

ลักษณะนิสัยทั่วไปที่มีเป็นครั้งคราว เป็นสิ่งที่จะมีผลกระทบต่อเครื่องมือวัดแทบทุกอัน บางครั้งสุขภาพของบุคคลอาจจะไม่อยู่ในสภาพปกติ อาจจะเหนื่อยเกินไป อาจจะเครียดเกินไป อาจจะขาดความรู้สึกพร้อม สภาพแวดล้อมเช่นแสงสว่างและความร้อนอาจทำให้บุคคลมีความรู้สึกต่าง ๆ กันออกไป เป็นผลให้ผลที่วัดได้แปรปรวนไป

ลักษณะนิสัยเฉพาะตัวที่มีเป็นครั้งคราว อาจทำให้ผลการวัดของเครื่องมือวัดบางอันแปรปรวนไป บุคคลบางคนมีระดับความเข้าใจคำสั่งที่จะปฏิบัติตามเครื่องมือวัดต่างกัน บุคคลบางคนมีความพร้อมที่จะถูกวัดในเรื่องบางเรื่องในระหว่างที่บุคคลอื่นไม่มี บางคนอาจปฏิบัติตามข้อความที่ใช้ทดสอบได้ แต่ในขณะที่ถูกวัดไม่สามารถทำได้ บุคคลแต่ละคนมีระดับความตั้งใจ ความระมัดระวัง และความเที่ยงตรงในเรื่องรายละเอียดต่างกัน

ส่วนความแปรปรวนอันบังเอิญนั้นหมายถึง สิ่งบางสิ่งบางอย่างที่ยังกำหนดไม่ได้ว่าเป็นลักษณะพิเศษอันใด เช่น ความบังเอิญในการเลือกคำตอบ หรือความบังเอิญในการเขียนคำตอบ เป็นต้น

ลักษณะความแปรปรวนที่อธิบายมาทั้งหมดนี้ ถ้าจะสรุปเพื่อให้ง่ายก็อาจจะกล่าวได้สั้น ๆ ว่า ความเปลี่ยนแปลงในเรื่องเวลา ย่อมทำให้การปฏิบัติงานหรือการแสดงพฤติกรรมเปลี่ยนแปลงไป ในขณะที่พฤติกรรมจริง ๆ ของแต่ละบุคคลก็เปลี่ยนแปลงไปด้วย เครื่องมือวัดที่มีค่าความเชื่อมั่นอันคำนวณได้จากผลการปฏิบัติงานหรือการแสดงพฤติกรรมของบุคคลเฉพาะชั่วระยะเวลาเดียว จึงยังถือว่ามีค่าความเชื่อมั่นที่มีคุณภาพยังไม่สูงพอ ยิ่งไปกว่านั้น ความสามารถในการปฏิบัติงานและพฤติกรรมทั้งหมดของแต่ละบุคคลยังมีขอบเขตกว้างขวาง เครื่องมือวัดที่มีเนื้อหาอันเป็นตัวแทนเพียงส่วนหนึ่งของการปฏิบัติงานและพฤติกรรมทั้งหมด จึงยังถือไม่ได้ว่าเป็นเครื่องมือวัดที่มีค่าความเชื่อมั่นอันสมบูรณ์ วิธีการต่าง ๆ ในการหาค่าความเชื่อมั่นที่กล่าวมาแล้วต่างก็มีจุดอ่อนในเรื่องเหล่านี้ ดังจะเห็นได้จากตารางที่ 10

ตารางที่ 10

ความแปรปรวนกับวิธีหาค่าความเชื่อมั่นวิธีต่าง ๆ *

วิธีหาค่าความเชื่อมั่น	ความแปรปรวนที่เกิดจาก:			
	ตัวเครื่องมือวัด	ตัวบุคคล	งานและพฤติกรรม	การเปลี่ยนแปลงความฉับพลันของบุคคล
1. ทดลองเครื่องมือวัดซ้ำโดยไม่ทิ้งระยะ	มี	ไม่มี	ไม่มี	มี
2. ทดลองเครื่องมือวัดซ้ำโดยทิ้งระยะ	มี	มี	ไม่มี	มี
3. เปรียบเทียบเครื่องมือวัดโดยไม่ทิ้งระยะ	มี	ไม่มี	มี	มี
4. เปรียบเทียบเครื่องมือวัดโดยทิ้งระยะ	มี	มี	มี	มี
5. แบ่งครึ่งเครื่องมือวัด	มี	ไม่มี	มี	ไม่มี
6. วิเคราะห์รายการ	มี	ไม่มี	มี	ไม่มี

* แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 131.

การแปลความหมายของสัมประสิทธิ์ความเชื่อมั่น

สัมประสิทธิ์ของความเชื่อมั่นของเครื่องมือวัดแต่ละอัน จะต้องพิจารณาให้รอบคอบ โดยไม่คำนึงถึงแต่เพียงค่าผิดพลาดมาตรฐานดังที่ได้กล่าวมาแล้วในตอนต้นบทแล้วนั้น ผู้วิเคราะห์ค่าความเชื่อมั่นยังจะต้องคำนึงอีกด้วยว่า ค่านี้เป็นค่าที่แสดงความคงเส้นคงวาของการวัดลำดับที่ของบุคคลในกลุ่มหนึ่ง ๆ เท่านั้น และยิ่งไปกว่านั้น ความแปรปรวนต่าง ๆ ที่มีผลกระทบต่อการศึกษา และการแสดงพฤติกรรมตามวิธีการของเครื่องมือวัด ยังเป็นสิ่งที่สำคัญยิ่งอีกอย่างหนึ่งที่จะต้องพิจารณาประกอบควบคู่ไปกับสัมประสิทธิ์ของความเชื่อมั่นของเครื่องมือวัดหนึ่ง ๆ อีกด้วย การที่จะกล่าวว่

ค่าความเชื่อมั่น 0.80 หรือ 0.85 เป็นค่าความเชื่อมั่นที่สูง จึงเป็นการกล่าวที่ผิด นอกจากนั้น ค่าความเชื่อมั่นควรจะสูง หรือควรจะต่ำ ก็ขึ้นกับนโยบายของหน่วยงาน หน่วยงานที่มีนโยบายคัดเลือกบุคคลระดับปานกลางเข้าปฏิบัติงาน อาจไม่จำเป็นที่จะต้องใช้เครื่องมือวัดที่มีค่าสูง หรือหน่วยงานที่ประสงค์จะค้นคว้าวิจัยเกี่ยวกับความสัมฤทธิ์หรือพฤติกรรมอื่น ๆ ของเจ้าหน้าที่ ก็อาจไม่จำเป็นต้องใช้เครื่องมือวัดที่มีสัมประสิทธิ์ของความเชื่อมั่นสูง แต่อย่างไรก็ดี โดยทั่วไป เครื่องมือวัดที่มีค่าความเชื่อมั่นสูง ก็ย่อมเป็นที่พึงปรารถนาของหน่วยงานที่จะใช้ นอกจากนี้จะพิจารณาประเด็นต่าง ๆ ที่เกี่ยวข้องกับความหมายของสัมประสิทธิ์ความเชื่อมั่นดังกล่าวมาแล้วนี้ ผู้ทดลองหรือผู้ใช้เครื่องมือวัดยังจะต้องคำนึงถึงปัจจัยอีก 4 ประการประกอบด้วย คือ¹³

1. พิสัย (range) ของความสามารถ และพฤติกรรมของบุคคลในกลุ่มที่ใช้ทดลองหาค่าความเชื่อมั่น
2. ค่าเฉลี่ยของความสามารถ และพฤติกรรมของบุคคลในกลุ่มที่ใช้ทดลองหาค่าความเชื่อมั่น
3. ความยาวของเครื่องมือวัด
4. วิธีปฏิบัติการในการคำนวณค่าสัมประสิทธิ์ความเชื่อมั่น

พิสัยของบุคคลในกลุ่ม สัมประสิทธิ์ของความเชื่อมั่นแสดงได้แต่เพียงว่า เมื่อใช้เครื่องมือวัดนั้น ๆ วัดแล้ว จะสามารถวัดลำดับที่ของบุคคลในกลุ่มหนึ่ง ๆ ได้คงเส้นคงวาเพียงใด ฉะนั้น ถ้าหากเครื่องมือวัดหนึ่ง ๆ ใช้วัดบุคคลที่มีระดับฝีมือในการปฏิบัติงาน หรือมีลักษณะพฤติกรรมต่างกันมาก ๆ ผลการวัดก็จะเห็นได้ชัด แต่ถ้าหากว่าเป็นการวัดบุคคลที่มีระดับฝีมือในการปฏิบัติงาน หรือพฤติกรรมใกล้เคียงกัน หรือเกือบเหมือนกัน ผลการวัดก็จะยอมไม่ชัด

ในกรณีการวัดเชาว์ เช่น แบบทดสอบมาตรฐาน Otis Quick-Scoring Mental Ability Test-Beta ปรากฏว่า ค่าความเชื่อมั่นที่คำนวณได้จากการทดลองนักเรียนแต่ละชั้นมีค่าตั้งแต่ 0.65—0.87 ฉะนั้น ค่าความเชื่อมั่นเฉลี่ยจึงเท่ากับ 0.78 แต่ในกรณีที่ใช้ผลการทดลองนักเรียนทุกชั้นรวมกัน ปรากฏว่าได้ค่าความเชื่อมั่นถึง 0.96 ค่าความเชื่อมั่นเหล่านี้ล้วนแต่เป็นค่าความเชื่อมั่นของเครื่องมือวัดอันเดียวกันทั้งสิ้น แสดงให้เห็นว่า ในกรณีที่ใช้กลุ่มบุคคลที่มีระดับความสามารถ และพฤติกรรมซึ่งมีพิสัยกว้าง ทำให้ได้ค่าความเชื่อมั่นสูง แบบทดสอบมาตรฐานอื่น ๆ ที่ได้มีการเปรียบเทียบ

¹³ Thorndike and Hagen, *op. cit.*, pp. 135—139.

ค่าความเชื่อมั่นซึ่งหาได้จาก การทดลองกับประชากรที่มีพฤติกรรมในพิสัยต่าง ๆ กัน ก็ปรากฏว่าได้ผลเช่นเดียวกันนี้

การที่พิสัยของความสามารถในการปฏิบัติงาน และพฤติกรรมของบุคคลที่ถูกทดลองหาค่าความเชื่อมั่นแคบ ก็อาจกล่าวได้อีกอย่างหนึ่งว่า บุคคลกลุ่มนั้น มีความคล้ายคลึงกันในกลุ่ม (homogeneity) กลุ่มที่มีลักษณะคล้ายคลึงกันภายในกลุ่ม ย่อมมีความแปรปรวนของผลการปฏิบัติงานและพฤติกรรมน้อย ค่าความแปรปรวนของความสามารถในการปฏิบัติงานและพฤติกรรมนี้ เป็นผลทำให้ค่าความเชื่อมั่นน้อยตามลงไปด้วย¹⁴ ข้อเท็จจริงอันนี้ มีความสำคัญในการตีความหมายของสัมประสิทธิ์ความเชื่อมั่น ในกรณีที่ผู้ใช้เครื่องมือวัดทราบค่าความแปรปรวนของกลุ่มบุคคลที่จะถูกวัด ผู้ใช้ก็อาจนำค่าความแปรปรวนของกลุ่มบุคคลที่ถูกทดลองมาเทียบกับความแปรปรวนของกลุ่มบุคคลที่จะถูกวัด เพื่อจะทำให้ทราบค่าความเชื่อมั่นที่แท้จริงในการวัดบุคคลแต่ละครั้งได้ ดังสมการข้างล่าง¹⁵

$$r_{nn} = 1 - \frac{\sigma_o^2 (1-r_{oo})}{\sigma_n^2}$$

r_{nn} = สัมประสิทธิ์ความเชื่อมั่นที่จะถือได้ในการวัดครั้งใหม่

σ_o^2 = ค่าความแปรปรวนของกลุ่มที่ใช้ทดลองเดิม

r_{oo} = สัมประสิทธิ์ความเชื่อมั่นที่ได้จากกลุ่มที่ทดลองเดิม

σ_n^2 = ค่าความแปรปรวนของกลุ่มที่ใช้วัดใหม่

แต่การที่จะมีค่าความแปรปรวนของความสามารถในการปฏิบัติงาน และพฤติกรรมของบุคคลที่จะถูกวัดแต่ละครั้งได้ ก็อาจจะเป็นอุปสรรค การที่จะมีค่าความแปรปรวนนี้ได้ หมายความว่า บุคคลที่จะถูกวัดจะต้องถูกบรรจุ และมอบหมายให้ปฏิบัติงานเป็นระยะเวลาพอสมควร ฉะนั้น ในกรณีที่ไม่มีค่าความแปรปรวนของกลุ่มบุคคลที่จะวัดใหม่มาเทียบ อย่างน้อยผู้ใช้เครื่องมือวัดนั้น ๆ ก็จะต้องคำนึงว่าค่าความเชื่อมั่นที่ระบุไว้นั้น อาจมีค่ามากเกินไปได้

ค่าเฉลี่ยของบุคคลในกลุ่ม ค่าความเชื่อมั่นของเครื่องมือวัดหนึ่ง ๆ จะถือว่าแสดงความเที่ยงตรงโดยไม่ได้คำนึงถึงว่า กลุ่มบุคคลที่ใช้ทดลองนั้นมีระดับความสามารถใน

¹⁴ Guion, *op. cit.* pp. 37-38.

¹⁵ *Ibid.*, p. 47.

เกณฑ์ใดไม่ได้ ในกรณีที่นำเครื่องมือวัดไปทดลองกับบุคคลที่มีระดับความสามารถในการปฏิบัติงานต่ำกว่าปกติ ผลที่วัดได้จากการทดลองก็อาจจะเป็นผลของการที่บุคคลกลุ่มนั้นเดาเป็นส่วนใหญ่ ความแม่นยำของผลที่วัดได้ก็ย่อมต่ำ ในทางกลับกัน ถ้าหากว่ากลุ่มที่ทดลองนั้นมีระดับความสามารถสูง ทุก ๆ คนในกลุ่มทำตามเครื่องมือวัดได้หมดหรือเกือบหมด ผลที่วัดได้จากการทดลองมาได้ก็ย่อมไม่แสดงว่าบุคคลที่ทำได้หรือทำไม่ได้ต่างกันอย่างไร กล่าวอีกนัยหนึ่งก็คือ ถ้าหากว่ากลุ่มที่ใช้ทดลองมีระดับความสามารถต่ำ ความแปรปรวนอันบังเอิญจะมีส่วนเข้ามาเกี่ยวข้องมาก ทำให้ผลการทดลองบิดเบือนไปมาก

ตามธรรมดา ผู้สร้างแบบทดสอบมาตรฐานจะระบุค่าผิดพลาดมาตรฐานของแบบทดสอบของตน เมื่อใช้กับบุคคลที่มีระดับผลการปฏิบัติงานต่าง ๆ กัน ค่าผิดพลาดมาตรฐานนี้ จะเป็นเครื่องช่วยในการประเมินว่า ถ้าจะใช้แบบทดสอบมาตรฐานนี้กับบุคคลในระดับความสามารถต่าง ๆ กัน จะถือได้ว่า ความเชื่อมั่นที่ทดลองหาไว้ได้นั้น มีความแม่นยำเพียงใด ทั้งนี้ ผู้ใช้แบบทดสอบอาจคำนวณค่าที่ควรจะเป็นจริงของค่าที่วัดได้ ดังเช่นในแบบทดสอบมาตรฐานที่ใช้วัดเชาวน์ Revised Stanford-Binet ได้ระบุค่าผิดพลาดมาตรฐานของค่าเชาวน์ในระดับต่าง ๆ ไว้ดังนี้

ตารางที่ 11

ค่าผิดพลาดมาตรฐานของ Revised Stanford-Binet*

ระดับเชาวน์	ค่าผิดพลาดมาตรฐานของแต่ละระดับ
130 ขึ้นไป	5.2
110-129	4.9
90-109	4.5
70-89	3.8
70 ลงมา	2.2

*แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 136.

จากสถิติข้างต้นนี้จะเห็นว่า ถ้าหากใช้แบบทดสอบมาตรฐานกับบุคคลที่มีความฉลาดหรือระดับเขาวนต่ำกว่าปกติ (ต่ำกว่าค่า 100) ผลการวัดจะเชื่อมั่นได้มากกว่าผลที่วัดได้สำหรับบุคคลในระดับที่สูงกว่าปกติ แบบทดสอบมาตรฐานบางแบบแสดงค่าผิดพลาดมาตรฐานในแง่ของอายุของบุคคลที่จะถูกวัด เช่น แบบทดสอบเขาวน Wechsler Intelligence Scale for Children ระบุค่ามาตรฐานไว้ดังนี้

ตารางที่ 12

ค่าผิดพลาดมาตรฐานของ Wechsler Intelligence Scale for Children*

ระดับ	ค่าผิดพลาดมาตรฐานของแต่ละระดับอายุ
$7 \frac{1}{2}$ ปี	4.2 ของค่าเขาวน
$10 \frac{1}{2}$ ”	3.4 ”
$13 \frac{1}{2}$ ”	3.7 ”

*แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 137.

ความยาวของเครื่องมือวัด แบบเครื่องมือวัดที่สั้นหรือที่วัดบุคคลโดยใช้เวลาน้อยย่อมมีค่าความเชื่อมั่นที่แม่นยำน้อยกว่าเครื่องมือวัดที่ยาวกว่า สูตรที่จะใช้คำนวณความแม่นยำของความเชื่อมั่น กำหนดไว้ดังนี้¹⁶

$$r_{nn} = \frac{nr_{tt}}{1 + (n-1)r_{tt}}$$

r_{nn} = ค่าความเชื่อมั่นที่ยาวเป็น n เท่าของเครื่องมือวัดเดิม

n = ความยาว หรือส่วนประกอบที่เพิ่มขึ้น

r_{tt} = สัมประสิทธิ์ของความเชื่อมั่นเดิม

จะเห็นได้ว่า สูตรนี้มีลักษณะคล้ายกับสมการที่ใช้คำนวณค่าความเชื่อมั่นวิธีแบ่งครึ่งเครื่องมือวัดนั่นเอง สมมติว่า แบบทดสอบมาตรฐานหนึ่งใช้วัดเขาวนของข้าราชการ

¹⁶ Thorndike and Hagen, *op. cit.*, p. 137.

ระดับหัวหน้ากอง มีข้อความที่ใช้ทดสอบ 100 รายการ หรือมีคำถาม 100 ข้อ แบบทดสอบนี้มีสัมประสิทธิ์ความเชื่อมั่น 0.50 ต้องการทราบว่า ถ้าขยายแบบทดสอบนี้ให้ มีข้อความที่ใช้ทดสอบ 500 รายการ จะทำให้มีความเชื่อมั่นได้เท่าใด ก็อาจแทนค่าในสมการได้ดังต่อไปนี้

$$\begin{aligned} r_{nn} &= \frac{5(0.50)}{1 + 4(0.50)} \\ &= 0.83 \end{aligned}$$

เครื่องมือวัดยังมีความยาวขึ้นเท่าใด ค่าผิดพลาดมาตรฐานก็ยิ่งน้อยลงไปตามนั้น เพราะผลที่วัดได้จะแสดงความสามารถในการปฏิบัติงานและพฤติกรรมของบุคคลในกลุ่มได้มากยิ่งขึ้น แต่ทั้งนี้มิได้แปลว่า เครื่องมือวัดจะต้องมีความยาวโดยไม่มีขอบเขตจำกัด เพราะโดยหลักธรรมดา การวัดก็ต้องจำกัดไม่ให้เวลานานเกินไป เพื่อมิให้เสียเวลาที่ผู้วัดและผู้ถูกวัด และไม่ทำให้ผู้ถูกวัดเหนื่อย หรือเบื่อหน่าย นอกจากนี้ เนื้อหาที่จะมาสร้างเป็นข้อความที่ใช้ทดสอบก็ยังมีจำกัด

วิธีปฏิบัติการในการคำนวณ ค่าสัมประสิทธิ์ความเชื่อมั่นที่หาได้จะมีความแม่นยำมากหรือน้อย ยังขึ้นอยู่กับวิธีปฏิบัติการในการคำนวณอีกด้วย ในตารางที่ 10 ในเรื่องความแปรปรวนที่มีผลกระทบต่อค่าความเชื่อมั่น จะสังเกตเห็นได้ว่า ถ้าผู้สร้างเครื่องมือวัดใช้วิธีการต่างกันในการคำนวณค่าความเชื่อมั่น ผลค่าความเชื่อมั่นที่ได้ จะสะท้อนความแปรปรวนในลักษณะต่างกัน การคำนวณค่าความเชื่อมั่นด้วยวิธีแบ่งครึ่งเครื่องมือวัด ถือเอาความแปรปรวนของบุคคลและงานที่จะวัดมาพิจารณาประกอบน้อยมาก วิธีเปรียบเทียบเครื่องมือวัดนั้น ๆ กับเครื่องมือวัดอื่น ๆ พิจารณาความแปรปรวนต่าง ๆ มากที่สุด เพราะบุคคลที่ถูกทดลองวัดจะต้องแสดงความคงเส้นคงวาในการปฏิบัติงานหรือแสดงพฤติกรรมตามที่กำหนดไว้ในเครื่องมือวัดทั้งสองอันที่ใช้เทียบกัน ในตารางที่ 9 ในเรื่องวิธีวิเคราะห์รายการ ได้แสดงค่าความเชื่อมั่นที่คำนวณได้โดยวิธีต่าง ๆ แต่ไม่ได้ระบุว่า ถ้าหากคำนวณค่าความเชื่อมั่นของแต่ละแบบ โดยไม่ได้ใช้วิธีตามที่แสดงไว้ นั้น จะได้ค่าต่างกันไปอย่างไร แม้กระนั้นก็จะเห็นได้ว่า วิธีแบ่งครึ่งเครื่องมือวัดแสดงสัมประสิทธิ์ของความเชื่อมั่นไว้สูง

จากสถิติในตารางที่ 13 จะเห็นได้ชัดเจนขึ้นว่า หากแบบทดสอบมาตรฐานหนึ่งๆ เปลี่ยนวิธีการคำนวณค่าความเชื่อมั่นแล้ว จะได้ผลต่างกันประการใด

ตารางที่ 13

เปรียบเทียบค่าความเชื่อมั่นของวิธีคำนวณบางแบบ *

ชื่อแบบทดสอบมาตรฐาน	ผลที่ได้จากวิธีคำนวณค่าความเชื่อมั่นต่างกัน	
	วิธีแบ่งครึ่งเครื่องมือวัด	วิธีเปรียบเทียบเครื่องมือวัด
Otis Quick-Scoring Intelligence Test-Beta	.90	.84
Pintner-Durost Intelligence Test		
Scale 1, Picture Content	.92	.78
Scale 2, Reading	.97	.92

* แหล่งที่มา: Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 138.

จากตารางที่ 13 จะเห็นได้ว่า ค่าความเชื่อมั่นที่หาได้โดยวิธีแบ่งครึ่งเครื่องมือวัดนั้น จะมีผลสูงกว่าที่หาได้โดยวิธีเปรียบเทียบเครื่องมือวัด นักทฤษฎีการวัดจึงเห็นพ้องต้องกันว่า ค่าความเชื่อมั่นที่นั่น ควรจะเป็นค่าที่หาได้โดยวิธีเปรียบเทียบเครื่องมือวัด แต่เนื่องจากการหาค่าความเชื่อมั่นด้วยวิธีนี้ อาจมีความลำบากในบางกรณี ในบางครั้งจึงต้องอนุโลมให้ใช้ค่าความเชื่อมั่นที่หาได้โดยวิธีอื่น

นอกจากวิธีที่ใช้ในการคำนวณสัมประสิทธิ์ความเชื่อมั่นแล้ว ผู้ใช้เครื่องมือวัดยังจะต้องคำนึงอีกด้วยว่า ในการสร้างเครื่องมือวัดนั้น ตัวอย่างประชากรที่ใช้ในการทดลอง มีลักษณะถูกต้องตามหลักการวิจัยหรือไม่ ถ้าหากว่าตัวอย่างที่นำมาทดลองวัดมีน้อยมาก ค่าความเชื่อมั่นที่ระบุไว้ก็ย่อมจะไม่มีค่า แต่ทั้งนี้มิได้แปลว่า ถ้าหากตัวอย่างที่ทดลองมีจำนวนมาก ก็จะทำให้ค่าความเชื่อมั่นเชื่อถือได้และจะต้องพิจารณาต่อไปอีกว่า กลุ่มตัวอย่างที่นำมาทดลองนั้นมีลักษณะเป็นตัวแทนของประชากรจริง ๆ หรือไม่

ถ้าหากว่าจะสร้างเครื่องมือวัดผู้ที่เข้ารับราชการในตำแหน่งประจำแผนก แต่ผู้สร้างเครื่องมือวัดใช้นักศึกษาที่กำลังจะจบปริญญาจากมหาวิทยาลัย หรือผู้ปฏิบัติงานในฐานะหัวหน้าหน่วยงานในบริษัทธุรกิจเป็นตัวอย่างที่ใช้ทดลอง ค่าความเชื่อมั่นที่ได้มาก็ย่อมขาดความแม่นยำและความเที่ยงตรง

ระดับความเชื่อมั่นที่ควรเชื่อถือได้

ถ้าหากเครื่องมือวัดสองอันมีคุณสมบัติทุกอย่างเหมือนกันและเท่ากันหมด ยกเว้นแต่สัมประสิทธิ์ความเชื่อมั่น เครื่องมือวัดที่มีค่าความเชื่อมั่นสูงกว่าย่อมดีกว่า ในกรณีที่เครื่องมือวัดหนึ่งมีค่าความเชื่อมั่นสูงกว่าเครื่องมือวัดอีกอันหนึ่ง แต่ขาดคุณสมบัติบางอย่าง เช่น ต้องเสียค่าโสหุ้ยในการใช้มากกว่า มีความเที่ยงตรง (validity) น้อยกว่า แม้จะกล่าวได้ว่า เครื่องมือวัดอันแรกมีความเชื่อมั่นสูงกว่าอันหลัง แต่ก็มิได้แปลว่า เครื่องมือวัดอันแรกจะมีคุณสมบัติดีกว่า กล่าวสั้นๆ ในเรื่องนี้ก็คือว่า เครื่องมือวัดหนึ่งๆ จำเป็นที่จะต้องมีความเชื่อมั่นอย่างน้อยถึงระดับหนึ่ง แต่ก็อาจไม่จำเป็นต้องสูงไปกว่านั้น อย่างไรก็ตาม ยังไม่มีหลักการตายตัวที่จะกำหนดว่า เครื่องมือวัดหนึ่งๆ ควรจะมีค่าความเชื่อมั่นอย่างน้อยเท่าใด

การที่จะตัดสินว่า เครื่องมือวัดหนึ่งๆ ควรจะมีค่าความเชื่อมั่นสักเท่าใดจึงควรนำมาใช้ ขึ้นอยู่กับค่าความเชื่อมั่นของเครื่องมือวัดประเภทเดียวกันที่มีอยู่ ถ้าหากว่ามีแบบทดสอบมาตรฐานที่มีค่าความเชื่อมั่น 0.85 สำหรับใช้วัดในการคัดเลือกบรรจุข้าราชการในตำแหน่งเสมียนพนักงาน แบบทดสอบอันนี้อาจจะไม่นำไปใช้แม้จะมีค่าความเชื่อมั่นสูง เพราะมีแบบทดสอบมาตรฐานอีกแบบหนึ่งที่ใช้วัดในเรื่องเดียวกันได้ แต่มีค่าความเชื่อมั่นถึง 0.90 ในทำนองเดียวกัน แบบทดสอบมาตรฐานที่มีค่าความเชื่อมั่นเพียง 0.60 อาจจะเป็นแบบที่สมควรใช้ก็ได้ เพราะไม่มีแบบทดสอบใด ๆ ที่มีค่าความเชื่อมั่นมากกว่านี้

ถึงแม้จะไม่มีหลักการกำหนดตายตัวว่า เครื่องมือวัดหนึ่งๆ ควรจะมีค่าความเชื่อมั่นอย่างน้อยเท่าใด แต่ก็มีวิธีคำนวณว่า ค่าความเชื่อมั่นในระดับหนึ่งๆ จะทำให้เชื่อถือได้เพียงใด ระดับความแม่นยำของค่าความเชื่อมั่นนี้พิจารณาได้จากโอกาสความน่าจะเป็น ในตารางที่ 14 แสดงค่าเปรียบเทียบของโอกาสความน่าจะเป็นกับระดับความเชื่อมั่น

ตารางที่ 14

ค่าความแม่นยำของสัมประสิทธิ์ความเชื่อมั่น*

ค่าความเชื่อมั่น	โอกาสความน่าจะเป็นของผลการวัด แต่ละบุคคลที่จะเปลี่ยนลำดับที่
.00	.500
.40	.403
.50	.368
.60	.325
.70	.271
.80	.197
.90	.087
.95	.022
.98	.0005

* แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 140.

จากตารางข้างบนนี้จะเห็นได้ว่า เมื่อสัมประสิทธิ์ของความเชื่อมั่นเท่ากับ 0.00 จะมีโอกาสครึ่งต่อครึ่งที่บุคคลที่ถูกวัดจะเปลี่ยนลำดับที่ใหม่ เมื่อถูกวัดซ้ำด้วยเครื่องมือเดิม เมื่อค่าความเชื่อมั่นสูงขึ้น ผลการวัดซ้ำด้วยเครื่องมือเดิมจะมีโอกาสใกล้เคียงกับผลการวัดครั้งแรกมากขึ้น เช่น เมื่อสัมประสิทธิ์เท่ากับ 0.50 จะมีโอกาสเพียง 1 ใน 3 เท่านั้นที่บุคคลคนหนึ่งจะเปลี่ยนลำดับที่ในการวัดซ้ำ และเมื่อสัมประสิทธิ์ถึง 0.98 จะมีโอกาสเพียง 1 ใน 200 ที่ผลการวัดซ้ำของบุคคลหนึ่ง ๆ จะเปลี่ยนไป

นอกจากจะพิจารณาความแม่นยำของสัมประสิทธิ์ความเชื่อมั่นในแง่ความสัมพันธ์กับความน่าจะเป็นในการที่จะได้ผลที่วัดซ้ำแล้ว ผู้วิเคราะห์ความแม่นยำของความเชื่อมั่นยังอาจพิจารณาได้ในแง่ความสัมพันธ์ของเครื่องมือวัดในเรื่องหนึ่งกับเครื่องมือวัดอีกเรื่องหนึ่ง เช่น ถ้าหากว่า มีเครื่องมือวัดความรู้เรื่องเศรษฐศาสตร์และความรู้ในเรื่องภาษาไทยสำหรับข้าราชการในตำแหน่งเศรษฐกรตรี ก็อาจจะนำค่าสหสัมพันธ์และค่าความเชื่อมั่นของเครื่องมือวัดทั้งสองมากำหนด เพื่อหาความแม่นยำที่แท้จริงในการใช้

เครื่องมือวัดทั้งสอง เพื่อวัดบุคคลที่จะปฏิบัติงานในตำแหน่งเศรษฐกรตรี ดังสมการต่อไป¹⁷

$$r_{1\infty 2\infty} = \frac{r_{12}}{\sqrt{r_{11} r_{22}}}$$

$r_{1\infty 2\infty}$ = ค่าสหสัมพันธ์ที่แสดงความแม่นยำระหว่างผลที่ถูกต้องในการวัด (error free) ของการใช้เครื่องมือวัดทั้งสองร่วมกัน

r_{12} = ค่าสหสัมพันธ์ระหว่างผลการวัดของเครื่องมือทั้งสอง

r_{11} และ r_{22} = สัมประสิทธิ์ความเชื่อมั่นของเครื่องมือวัดแต่ละอัน

ถ้าสมมติว่า เครื่องมือวัดความรู้เรื่องเศรษฐศาสตร์ และความรู้เรื่องภาษาไทยดังตัวอย่าง มีค่าความเชื่อมั่นเท่ากับ 0.71 และ 0.90 ตามลำดับ และค่าสหสัมพันธ์ระหว่างผลการวัดของเครื่องมือทั้งสองเท่ากับ 0.56 ก็อาจหาความแม่นยำของการใช้เครื่องมือวัดทั้งสองร่วมกัน เพื่อให้ทราบว่า แม่นยำเพียงใดในการใช้ประกอบการตัดสินใจบรรจุบุคคลเข้าในตำแหน่งเศรษฐกรตรีได้โดยแทนค่าดังนี้

$$\begin{aligned} r_{1\infty 2\infty} &= \frac{0.56}{\sqrt{(0.71)(0.90)}} \\ &= 0.70 \end{aligned}$$

¹⁷ *Ibid.*, p. 141.

ค่าความเที่ยงตรงของเครื่องมือวัด

หลังจากที่ได้สร้างเครื่องมือวัดแล้ว ขั้นตอนต่อไปก็คือ การวิเคราะห์ว่า เครื่องมือนั้น มีประโยชน์หรือไม่ หรือกล่าวอีกนัยหนึ่งว่า เครื่องมือนั้นวัดสิ่งที่ต้องการวัดได้หรือไม่ ครบถ้วนหรือไม่ และวัดนอกเหนือไปจากสิ่งที่ต้องการวัดหรือไม่ การวิเคราะห์ในขั้นนี้ เรียกว่า การหาความเที่ยงตรงหรือความตรงประเด็น (validity) ศัพท์ “ความเที่ยงตรง” นี้ ใช้กันแพร่หลายพอสมควรในประเทศไทย ฉะนั้น ในหนังสือเล่มนี้ จึงจะใช้ศัพท์นี้ ถึงแม้ว่า คำ “ความตรงประเด็น” จะให้ความหมายที่เข้าใจง่ายกว่า ทั้งนี้เพื่อช่วยมิให้ผู้อ่านเข้าใจสับสน เมื่อไปพบศัพท์นี้ในที่อื่น

เครื่องมือวัดบางชนิดหาความเที่ยงตรงได้ง่าย เช่น เทอร์มิเตอร์หรือเครื่องชั่ง ตวงที่ยกตัวอย่างไว้ในบทที่ 4 เครื่องมือวัดประเภทนี้แทบจะไม่ต้องนำมาวิเคราะห์เพื่อหาความเที่ยงตรง การที่เป็นเช่นนั้นก็เพราะเหตุผลสองประการ ประการแรก เครื่องมือประเภทนี้สามารถวัดในเรื่องที่ตรงกับแนวความคิดของการวัดนั้น ๆ อยู่แล้ว ประการที่สอง เครื่องมือนั้นมีความสัมพันธ์กับตัวแปรอื่น ๆ อยู่แล้ว หรือกล่าวอีกอย่างหนึ่งได้ว่า ไม่ว่าสภาพการวัดจะเป็นเช่นใด และเพื่อประโยชน์แก่กิจกรรมใด เครื่องมือวัดประเภทนี้ก็สามารถวัดให้ได้ผลในเรื่องที่ต้องการเสมอ ถ้าหากว่าเครื่องมือวัดบุคคลมีลักษณะเช่นนี้ ก็แทบจะไม่ต้องมีการวิจัยเพื่อหาค่าความเที่ยงตรงเลย แต่โดยความเป็นจริงแล้ว เครื่องมือวัดบุคคลมีลักษณะพิเศษออกไป กล่าวคือ เครื่องมือวัดบุคคลต้องวัดคุณสมบัติทางจิตวิทยาของบุคคลที่มีลักษณะสลับซับซ้อน เช่น ความสนใจ ความกลัว เป็นต้น การวัดซึ่งพจร ปริมาณของเหงื่อ ความถี่ของลมหายใจ ฯลฯ สามารถทำได้ แต่ลักษณะทางจิตวิทยา เช่น ความสนใจ และความกลัวนั้นวัดได้ยาก การที่วัดได้ยากเช่นนั้น ทำให้ต้องมีการพิสูจน์ความเที่ยงตรง

การหาความเที่ยงตรงของเครื่องมือวัดบุคคล จึงต้องอาศัยการวิเคราะห์เหตุผลของประเด็นที่จะวัดเป็นหลักในการพิจารณา กล่าวโดยทั่วไปแล้ว เครื่องมือวัดต่างก็มีความเที่ยงตรงด้วยกันทั้งนั้น ไม่เหมือนกันตรงที่ว่า มีความเที่ยงตรงมากน้อยต่างกัน จะไม่มีเครื่องมือวัดอันใดที่มีค่าความเที่ยงตรงเป็นศูนย์ หรือมีค่าความเที่ยงตรงเป็น 100 เปอร์เซ็นต์¹ โดยหลักการที่แท้จริง การหาความเที่ยงตรงของเครื่องมือวัด ไม่ใช่เป็นการวิเคราะห์ความเที่ยงตรงของเครื่องมือวัด แต่เป็นการวิเคราะห์ความเที่ยงตรงที่จะได้รับจากการใช้เครื่องมือวัดให้เป็นประโยชน์ ยกตัวอย่างเช่น เครื่องมือวัดบุคคลที่สมัครเข้ารับราชการในตำแหน่งประจำแผนกที่เที่ยงตรง จะมีความเที่ยงตรงต่อเมื่อใช้ในการคัดเลือกบุคคลเพื่อบรรจุบุคคลในตำแหน่งนี้เท่านั้น แต่จะไม่เที่ยงตรงเมื่อใช้วัดบุคคลที่สมัครเข้ารับราชการในตำแหน่งวิทยากรตรีหรือในตำแหน่งอื่น ๆ เครื่องมือวัดบุคคลที่มีความเที่ยงตรงดี จะต้องใช้ประโยชน์ได้ประการใดประการหนึ่งใน 3 ประการ² คือ สามารถสร้างความสัมพันธ์เชิงปฏิบัติการ (functional relationship) กับตัวแปรเฉพาะหนึ่ง ๆ ประการที่สอง มีลักษณะที่เป็นตัวแทนสาระสำคัญของที่อยู่ในโลกของเรื่องนั้น หรือประการที่สาม วัดค่าของคุณสมบัติพฤติกรรมของบุคคล (trait) ได้ การวัดได้ตามประการแรก หมายความว่า ผลการวัดของเครื่องมืออื่น ๆ สามารถใช้คาดคะเนได้ว่าจะมีการแสดงพฤติกรรมอย่างใดอย่างหนึ่ง ถ้าวัดได้ตามประการที่สอง หมายความว่า สาระสำคัญของกรวัด ตรงกับสาระสำคัญของสิ่งที่ตั้งเป้าหมายไว้ในกรวัด ส่วนการวัดได้ตามประการสุดท้ายนั้น หมายความว่า ผลของการวัดแสดงลักษณะอย่างหนึ่งอย่างใดที่เป็นคุณสมบัติทางจิตวิทยาของบุคคลได้ การที่สามารถวัดได้เที่ยงตรงนี้อาจแสดงเป็นสัมประสิทธิ์ได้ โดยการหาค่าสหสัมพันธ์ระหว่างผลการวัดของเครื่องมือวัดหนึ่ง ๆ กับค่าของสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรงของเครื่องมือวัดนั้น ๆ เช่น ถ้าประสงค์ให้มีความเที่ยงตรงในแง่ความสัมพันธ์เชิงปฏิบัติกับตัวแปรเฉพาะหนึ่ง ๆ ผู้สร้างเครื่องมือวัดก็อาจคำนวณค่าสหสัมพันธ์ระหว่างผลการวัดของเครื่องมือวัดของตนกับผลการปฏิบัติงาน หรือการแสดงผลพฤติกรรมจริง ๆ

ความเที่ยงตรงกับความแปรปรวน

เครื่องมือวัดที่มีความเที่ยงตรงจะแสดงผลการวัดที่มีค่าสหสัมพันธ์มากกว่าศูนย์ กับสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรงของเครื่องมือวัดนั้น ๆ กล่าวอีกอย่างหนึ่ง

¹ Nunnally, J.C., *Psychometric Theory*, p. 95.

² Ghiselli, E.E., *Theory of Psychological Measurement*, pp. 298—299.
Nunnally, *op. cit.*, p. 76.

ได้ว่า ผลการวัดของเครื่องมือวัดนั้น มีความแปรปรวนร่วมกันกับค่าของสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรง เนื่องจากค่าของสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรงนั้นเป็นค่าคนละชุดกับค่าที่ได้จากผลการวัด ฉะนั้น ค่าทั้งสองนี้ จะมีลักษณะพิเศษประจำชุด หรือเรียกอีกอย่างหนึ่งว่า ค่าความแปรปรวนเฉพาะ (specific variance) ในบทที่ 4 ได้กล่าวไว้แล้วว่า ค่าความเชื่อมั่นอาจหาได้จากสูตร

$$r_{tt} = \frac{\sigma_{\infty}^2}{\sigma_t^2}$$

ค่านี้ เป็นค่าที่แทนสัมประสิทธิ์สหสัมพันธ์ของผลการวัดของเครื่องมือวัดหนึ่ง ๆ กับตัวของมันเอง ฉะนั้น หากจะเป็นสัมประสิทธิ์สหสัมพันธ์ของผลการวัด ของเครื่องมือวัดหนึ่ง ๆ กับค่าชุดอื่น ๆ ซึ่งมีความแปรปรวนเฉพาะประจำชุด ค่าสหสัมพันธ์ซึ่งจะเป็นค่าความเที่ยงตรง จะต้องเขียนเป็นสมการใหม่ คือ

$$r_{xy} = \frac{\sigma_{\infty}^2 - \sigma_{sp}^2}{\sigma_t^2}$$

$$r_{xy} = \text{ค่าความเที่ยงตรง}$$

$$\sigma_{sp}^2 = \text{ค่าความแปรปรวนเฉพาะ}$$

จากสมการข้างบนนี้ อาจคำนวณต่อไปได้ดังนี้

$$\therefore \sigma_{\infty}^2 = \sigma_{xy}^2 + \sigma_{sp}^2$$

$$\therefore r_{xy} = \frac{(\sigma_{xy}^2 + \sigma_{sp}^2) - (\sigma_{sp}^2)}{\sigma_t^2}$$

$$= \frac{\sigma_{xy}^2}{\sigma_t^2}$$

ได้อธิบายมาแล้วในบทที่ 3 ว่า σ_{xy}^2 คือ ค่าความแปรปรวนสัมพันธ์ ฉะนั้น ค่าความเที่ยงตรงก็คือ อัตราส่วนระหว่างความแปรปรวนสัมพันธ์กับค่าความแปรปรวนของจำนวนรวมทั้งหมด

ค่าความแปรปรวนขององค์ประกอบร่วมกัน ได้อธิบายมาแล้วว่า เป็นค่าที่ข้อมูลตั้งแต่ 2 ชุด เกิดความแปรปรวนร่วมกัน ส่วนค่าความแปรปรวนเฉพาะนั้น มีความหมายตรงข้ามกัน คือเป็นค่าที่ข้อมูลในชุดหนึ่ง ๆ มีความแปรปรวนโดยไม่ซ้ำกับข้อมูลในชุดอื่น ๆ ค่าความเที่ยงตรงเป็นค่าที่พิเศษยิ่งไปกว่าค่าความเชื่อมั่นได้ในแง่ที่ว่า

จะต้องพิจารณาแยกค่าความแปรปรวนเฉพาะนี้ออกเสียก่อน ถ้าหากจะพิจารณาวิธีการคำนวณข้างล่างนี้ จะเห็นความสัมพันธ์ระหว่างค่าความเชื่อมั่น กับค่าความเที่ยงตรงได้ชัดเจนขึ้น

$$\begin{aligned}\therefore \sigma_t^2 &= \sigma_\infty^2 + \sigma_e^2 \\ &= \sigma_{xy}^2 + \sigma_{sp}^2 + \sigma_e^2\end{aligned}$$

เอา σ_t^2 ทหารตลอด

$$\frac{\sigma_t^2}{\sigma_t^2} = \frac{\sigma_{xy}^2}{\sigma_t^2} + \frac{\sigma_{sp}^2}{\sigma_t^2} + \frac{\sigma_e^2}{\sigma_t^2}$$

$$\therefore \frac{\sigma_{xy}^2}{\sigma_t^2} = \frac{\sigma_t^2}{\sigma_t^2} - \frac{\sigma_{sp}^2}{\sigma_t^2} - \frac{\sigma_e^2}{\sigma_t^2}$$

$$\text{หรือ } r_{xy} = \frac{\sigma_t^2}{\sigma_t^2} - \frac{\sigma_e^2}{\sigma_t^2} - \frac{\sigma_{sp}^2}{\sigma_t^2}$$

$$\text{แต่ } \therefore r_{tt} = \frac{(\sigma_t^2 - \sigma_e^2)}{\sigma_t^2} \dots\dots\dots \text{ตามสมการในบทที่ 4}$$

$$= \frac{\sigma_t^2}{\sigma_t^2} - \frac{\sigma_e^2}{\sigma_t^2}$$

$$\therefore r_{xy} = r_{tt} - \frac{\sigma_{sp}^2}{\sigma_t^2}$$

จะเห็นได้ว่า ค่าของความเชื่อมั่นมีค่าเท่ากับ อัตราส่วนของสัมประสิทธิ์ 2 ค่าแรก ของสมการค่าความเที่ยงตรง

ค่าความเที่ยงตรงในการคาดคะเน (predictive validity) มีสัดส่วนสัมพันธ์กับค่าความเชื่อมั่น ดังสูตรข้างล่างนี้³

$$r_{xy} = r_{x\infty} r_{y\infty} \sqrt{r_{xx} r_{yy}}$$

$r_{x\infty} r_{y\infty}$ = ค่าสหสัมพันธ์ระหว่างตัวแปรของการวัดผลที่แท้จริงกับค่าของสิ่งที่เป็นหลักแท้จริงในการพิจารณาความเที่ยงตรง [ค่าสหสัมพันธ์ระหว่าง “คะแนนจริง” ของแบบทดสอบ (predictor) กับ “คะแนนจริง” ของเกณฑ์ (criteria)]

³ Ghiselli, *op. cit.*, p. 352.

r_{xx} = ค่าความเชื่อมั่นของตัวแปรการวัดผล

r_{yy} = ค่าความเชื่อมั่นของสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรง

จากสมการข้างบนนี้ จะเห็นว่า ค่าสหสัมพันธ์ระหว่างตัวแปรของการวัดผลที่แท้จริงกับค่าของสิ่งที่เป็นหลักแท้จริงในการพิจารณาความเที่ยงตรง จะเป็นค่าที่คงที่ แต่ค่าความเชื่อมั่นของตัวแปรการวัดผล และค่าความเชื่อมั่นของสิ่งที่เป็นหลักการในการพิจารณาความเที่ยงตรง จะเปลี่ยนไปตามคุณภาพของเครื่องมือวัด และลักษณะของเรื่องที่จะวัด ค่า r_{xy} จึงเปลี่ยน เมื่อค่า r_{xx} และค่า r_{yy} เปลี่ยน ค่า r_{xy} จะเปลี่ยนไปตามอัตราดังที่แสดงไว้ในตารางที่ 15

ตารางที่ 15

ค่าสูงสุดของ r_{xy} เมื่อเทียบกับค่า r_{xx} และ r_{yy} *

r_{xx}	r_{yy}										
	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.10	0.00	0.10	0.14	0.17	0.20	0.22	0.24	0.26	0.26	0.30	0.32
0.20	0.00	0.14	0.20	0.24	0.28	0.32	0.25	0.37	0.40	0.42	0.45
0.30	0.00	0.17	0.24	0.30	0.35	0.39	0.42	0.46	0.49	0.52	0.55
0.40	0.00	0.20	0.28	0.35	0.40	0.45	0.49	0.53	0.57	0.60	0.63
0.50	0.00	0.22	0.32	0.39	0.45	0.50	0.55	0.59	0.63	0.67	0.71
0.60	0.00	0.24	0.35	0.42	0.49	0.55	0.60	0.65	0.69	0.73	0.77
0.70	0.00	0.26	0.37	0.46	0.53	0.59	0.65	0.70	0.75	0.79	0.84
0.80	0.00	0.28	0.40	0.49	0.59	0.63	0.69	0.75	0.80	0.85	0.89
0.90	0.00	0.30	0.42	0.52	0.60	0.67	0.73	0.79	0.85	0.90	0.90
1.00	0.00	0.32	0.45	0.59	0.63	0.71	0.77	0.84	0.89	0.95	1.00

*แหล่งที่มา : Ghiselli, E.E., *Theory of Psychological Measurement*, p. 271.

จากตารางข้างบนนี้จะเห็นได้ว่า เมื่อค่าสหสัมพันธ์ระหว่างตัวแปรของการวัดผลที่แท้จริงกับค่าของสิ่งที่เป็นหลักแท้จริงในการพิจารณาความเที่ยงตรงคงที่ และค่าความเชื่อมั่นของตัวแปรการวัดผลและสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรงลดน้อยลงไปเท่าใด ค่าความเที่ยงตรงก็ยิ่งลดน้อยตามลงไป ถ้าตัวแปรการวัดผล หรือสิ่ง

เป็นหลักในการพิจารณาความเที่ยงตรงอย่างหนึ่งอย่างใดขาดความเชื่อมั่น กล่าวคือ มีค่าสหสัมพันธ์กับตัวของมันเองเป็น 0.00 ค่าความเที่ยงตรงก็จะมีค่าเท่ากับ 0.00 ด้วย ไม่ว่าค่าของตัวแปรการวัดผลหรือสิ่งที่เป็นหลักในการพิจารณาความเที่ยงตรง จะมีค่าสูงเท่าไรก็ตาม

ประเภทความเที่ยงตรง

ความเที่ยงตรงแบ่งประเภทได้หลายอย่าง สมาคมจิตวิทยาอเมริกันและสมาคมวิจัยการศึกษาอเมริกันได้ร่วมกันพิจารณาเมื่อปี ค.ศ. 1954 และตกลงแบ่งประเภทความเที่ยงตรงออกเป็น 4 ประเภท คือ ความเที่ยงตรงในการคาดคะเน (predictive) ความเที่ยงตรงกับการวัดอื่น ๆ (concurrent) ความเที่ยงตรงในแง่เนื้อหา (content) และความเที่ยงตรงในแง่แนวคิด (construct)⁴ ในตำราบางเล่ม แยกประเภทความเที่ยงตรงออกเป็นอีกแบบหนึ่ง เช่น แยกความเที่ยงตรงว่ามีประเภท ความเที่ยงตรงที่เห็นชัด (face validity) ความเที่ยงตรงในเชิงปฏิบัติ (empirical) ความเที่ยงตรงในแง่เนื้อหาและความเที่ยงตรงในแง่แนวคิด⁵ แต่วิธีการแยกประเภทความเที่ยงตรงที่น่าจะทำให้เข้าใจชัดเจน คือ การแยกออกเป็นสองประเภทใหญ่ ได้แก่ความเที่ยงตรงในเชิงเหตุผล (rational) และความเที่ยงตรงในเชิงสถิติ (statistical)⁶ ความเที่ยงตรงในเชิงเหตุผลแยกย่อยออกไปอีกสองชนิด คือ ความเที่ยงตรงในแง่เนื้อหา และความเที่ยงตรงในแง่แนวคิด ส่วนความเที่ยงตรงในเชิงสถิติก็แยกย่อยออกเป็น 3 ชนิด คือ ความเที่ยงตรงกับเครื่องมือวัดอื่น ๆ ประเด็นเดียวกัน (congruent) ความเที่ยงตรงกับการวัดอื่น ๆ และความเที่ยงตรงในการคาดคะเน ถ้าจะพิจารณาแนวการแยกประเภททั้ง 3 วิธีนี้แล้ว ก็จะเห็นว่ามัลักษณะคล้ายคลึงกันมาก ต่างกันเพียงแต่รายละเอียดในเรื่องแนวความคิดของการแยกประเภท ในที่นี้จะได้อธิบายลักษณะของความเที่ยงตรงแต่ละประเภทตามวิธีการแยกประเภทวิธีสุดท้าย

ความเที่ยงตรงในแง่เนื้อหา การวิเคราะห์ว่า เครื่องมือวัดหนึ่ง ๆ เที่ยงตรงหรือไม่ อาจทำได้โดยการวิเคราะห์หาเหตุผล หรือใช้ดุลพินิจทางวิชาการเพื่อให้ทราบว่า เครื่องมือที่วัดควรครอบคลุมประเด็นและสาระสำคัญอันใดบ้าง ประเด็นและสาระสำคัญนั้น ก็เป็นเนื้อหาของเครื่องมือวัด ถ้าผู้ที่สร้างเครื่องมือวัดบุคคลเข้ารับราชการใน

⁴ Kerlinger, F.N., *Foundations of Behavioral Research*, p. 445.

⁵ Lyman, H.B., *Test Scores and What They Mean*, pp. 26–30.

⁶ Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, pp. 109–110.

ตำแหน่งประจำแผนก ต้องการทราบว่า เครื่องวัดของเรามีความตรงประเด็นในแง่เนื้อหาหรือไม่ ก็จะต้องใช้ดุลพินิจทางวิชาการวิเคราะห์ดูว่า งานของข้าราชการในระดับประจำแผนกมีประเด็นและสาระสำคัญอันใดบ้าง ผู้สร้างเครื่องมือวัดจะต้องได้เล็งดูว่า ผู้ที่ต้องปฏิบัติงานนี้จะต้องเผชิญกับสภาพการณ์ และต้องสามารถปฏิบัติงานหรือแสดงพฤติกรรมอย่างไรแล้วจัดเลือกสิ่งเหล่านี้ออกมาเป็นกลุ่ม เป็นประเภทๆ ถ้าเนื้อหาของเครื่องมือวัดเป็นตัวแทนของรายละเอียดต่างๆ เหล่านี้ได้ ก็กล่าวได้ว่า เครื่องวัดนี้มีความเที่ยงตรง ความเที่ยงตรงในแง่เนื้อหาจึงอาจนิยามได้ว่า เป็นความเป็นตัวแทน หรือความเพียงพอของสุ่มตัวอย่าง (sampling adequacy) ของเนื้อหาของเครื่องมือวัด ผู้ที่จะสร้างเครื่องมือวัดที่มีความเที่ยงตรงในแง่นี้ จะต้องตั้งคำถามตนเองเสมอว่า สาระเนื้อหาของเครื่องมือวัดเป็นตัวแทนของเนื้อหาของสิ่งที่จะวัดเท่าที่มีปรากฏอยู่ทั้งหมดหรือไม่ การที่จะถือได้ว่า สาระเนื้อหาของเครื่องมือวัดเป็นตัวแทนของสิ่งที่จะวัดทั้งหมดได้ มีหลักกว้างๆ ว่า เมื่อสิ่งที่จะวัดทั้งหมดมีประเภทย่อยๆ เป็นจำนวนเท่าใดก็ตาม สาระสำคัญของประเภทย่อยๆ ทุกประเภทเหล่านี้ จะต้องปรากฏอยู่ในเครื่องมือวัด

โดยทั่วไปแล้ว คุณสมบัติของพฤติกรรมในเรื่องหนึ่งๆ อาจแยกประเภท และคัดเลือกตัวแทนที่จะนำมาใส่ไว้ให้ปรากฏในเครื่องมือวัดได้ยาก ฉะนั้นการที่จะสร้างเครื่องมือวัดให้มีความเที่ยงตรงในแง่เนื้อหาตามหลักการดังกล่าวจึงเป็นเรื่องยาก ผู้สร้างเครื่องมือวัดบุคคลในตำแหน่งนักบัญชีตรี อาจจะพบว่า งานที่นักบัญชีตรีต้องทำ และความรู้ที่นักบัญชีตรีจะต้องมี แยกออกเป็นประเภทต่างๆ ได้มากมาย เฉพาะงานและความรู้ที่สำคัญที่สุดที่จำเป็นจะต้องทำหรือต้องมี ก็อาจมีอยู่หลายประการ จนแทบจะนำมาใส่ในเครื่องมือวัดให้ครบถ้วนไม่ได้ ฉะนั้น การหาความเที่ยงตรงประเภทนี้จึงต้องอาศัยดุลพินิจทางวิชาการเป็นหลักสำคัญ แล้วตัดสินใจว่า รายการใดในเครื่องมือวัดมีน้ำหนักและความสำคัญเกี่ยวข้องกับประเด็นใดที่จะวัด

ความเที่ยงตรงในแง่ความคิด แนวคิด (construct) หรือแนวความคิด (concept) เป็นคำที่อาจใช้แทนความหมายเดียวกันได้⁷ เมื่อกล่าวถึงแนวความคิด จะหมายถึงคำที่แสดงนามธรรม ซึ่งอธิบายสิ่งต่างๆ เช่น น้ำหนัก พลังงาน ความสูง เป็นต้น ส่วนแนวคิดนั้น เป็นแนวความคิดที่สร้างขึ้นเพื่อใช้อธิบายความหมายทางวิชาการ เช่น ความฉลาด ความริเริ่ม แรงกระตุ้น การที่จะสร้างเครื่องมือวัดให้เที่ยงตรงในแง่แนวคิด

⁷ Kerlinger, *op. cit.*, pp. 31-32.

ผู้สร้างจะต้องทราบว่า คุณสมบัติทางจิตวิทยาตัวใดของบุคคลที่มีความแปรปรวนสัมพันธ์ร่วมกับเครื่องมือวัดของตน ผู้สร้างเครื่องมือวัดเชาว์ (intelligence) จะต้องถามตนเองว่า มีองค์ประกอบตัวใดที่ทำให้ผลการใช้เครื่องมือวัดเกิดความแปรปรวน ผู้สร้างเครื่องมือวัดอาจจะตั้งคำถามเฉพาะลงไปก็ได้ว่า เครื่องมือวัดของตนวัดองค์ประกอบ ก. หรือองค์ประกอบ ข. หรือเปล่า และวัดองค์ประกอบ ค. ซึ่งเป็นองค์ประกอบที่ไม่พึงประสงค์อีกด้วยหรือเปล่า กล่าวอีกนัยหนึ่งก็คือ ในการสร้างเครื่องมือวัดที่เที่ยงตรงในแง่แนวคิด จะต้องอธิบายได้ว่า การที่บุคคลแต่ละคนมีความแตกต่างกันทางจิตวิทยา ทำให้ค่าคะแนนที่ได้จากเครื่องมือวัดนั้นแปรปรวนไปประการใด

การหาความเที่ยงตรงในแง่แนวคิด จะต้องมีการวิเคราะห์หาเหตุผลในกิจกรรมหรืองาน และขบวนการที่เกี่ยวข้องกับแนวคิดที่วัด ผู้สร้างเครื่องมือวัดจะต้องตอบคำถามได้ว่า แนวคิดในเรื่องที่จะวัดนั้นหมายความประการใด การวิเคราะห์แนวคิดในเรื่องที่จะวัดนั้นไม่ใช่เป็นการวินิจฉัยเนื้อหา เช่น ในเรื่องความเที่ยงตรงในแง่เนื้อหา แต่เป็นการวิเคราะห์ว่า แนวคิดที่เกี่ยวข้องกับเรื่องที่จะวัดนั้นมีความหมายกว้างขวางเท่าใด เช่น ถ้าต้องการสร้างเครื่องมือวัดผู้บังคับบัญชาที่ดี ผู้สร้างก็อาจจะทำการวิจัย และอาจพบว่า ผู้บังคับบัญชาที่มีหัวสมัยใหม่ เป็นผู้บังคับบัญชาที่ดี ทั้งนี้ โดยหาค่าสหสัมพันธ์ระหว่างผู้บังคับบัญชาหัวสมัยใหม่ กับผู้บังคับบัญชาที่ดี แต่การที่จะแปลความหมายออกมาว่า เหตุใดผู้บังคับบัญชาหัวสมัยใหม่ จึงเป็นผู้บังคับบัญชาที่ดี นักวัดผู้หนึ่งจะต้องแปลแนวคิดของคำว่า “หัวสมัยใหม่” กับคำว่า “ที่ดี” เสียก่อน การแปลความหมายของแนวคิดนี้จะทำให้ทราบว่า จะต้องวัดความสามารถและพฤติกรรมอันใดบ้าง เช่น สมมติว่า จะต้องการสร้างเครื่องมือวัดบุคคลเข้าทำงานในตำแหน่งเจ้าหน้าที่สารบรรณ และในการสร้างเครื่องมือวัดนั้น ปรากฏว่า เจ้าหน้าที่สารบรรณต้องมีความสามารถในการใช้ถ้อยคำภาษาที่จะก่อให้เกิดผลดี นักสร้างเครื่องมือวัดก็จะต้องวิเคราะห์แนวคิดของความสามารถในการใช้ภาษาที่จะก่อให้เกิดผลดี เพื่อให้ทราบว่า จะต้องวัดพฤติกรรมอันใดบ้าง ซึ่งอาจจะแยกได้ดังต่อไปนี้⁸

1. ความสามารถในการเลือกความคิดที่จะพูดหรือเขียน ได้แก่ ความคิดที่น่าสนใจหรือมีความสำคัญ และมีความประสานกัน
2. การประมวลความคิด ได้แก่การจัดลำดับความคิดให้เป็นเหตุเป็นผลให้ประทับใจผู้ฟัง หรือผู้อ่าน ตลอดจนการเรียบเรียงรายละเอียดของประเด็นใหญ่ ๆ

⁸ Thorndike and Hagen, *op. cit.*, p. 112.

3. การจัดส่วนตอนของความคิด ถ้าเป็นการเขียนก็เป็นการจัดย่อหน้า เพื่อประมวลข้อความที่อธิบายความคิดเป็นตอน ๆ ถ้าเป็นการพูดก็เป็นการจัดกลุ่มแนวความคิดที่จะพูดให้เป็นพวก ๆ

4. การใช้ประโยคที่ประทับใจ เป็นการเรียบเรียงประโยคที่สามารถอธิบายความคิดหนึ่ง ๆ ได้ (single complete idea) ตลอดจนการเรียบเรียงประโยคให้สั้นและยาวอย่างสละสลวย

5. การใช้คำที่เหมาะสม หมายถึงการที่สามารถเลือกคำที่มีความหมายตรงกับความคิด และสามารถใช้คำต่าง ๆ ที่น่าสนใจ อธิบายความคิดได้

6. ถ้อยคำ คือ ความสามารถในการกล่าวถึงความคิดทั้งหมดตามแนวที่เหมาะสมกับเรื่องที่จะเล่า และวิธีที่จะเล่า ตลอดจนผู้ที่จะฟังหรือผู้ที่จะอ่าน

รายการทั้งหมดข้างต้นนี้ เป็นการวิเคราะห์แนวคิดของความสามารถในการใช้ภาษาที่จะก่อให้เกิดผลดีในตำแหน่งเจ้าหน้าที่สารบรรณ ที่วิเคราะห์มาทั้งหมดนั้นเป็นเพียงตัวอย่าง มิได้แปลว่า ความสามารถในเรื่องนี้จะมีรายละเอียดเพียงเท่าที่ยกมา แต่ที่แสดงรายการนี้ก็เพื่อจะให้เห็นว่า เมื่อวิเคราะห์แนวคิดของความสามารถแล้ว ผู้สร้างเครื่องมือวัดก็จะมีประเด็นย่อย ๆ ที่จะวัดได้

ความเที่ยงตรงกับเครื่องมือวัดประเด็นเดียวกัน ผู้สร้างเครื่องมือวัด อาจหาความเที่ยงตรงได้โดยวิธีสถิติ ทั้งนี้ เพื่อจะได้ค่าที่แสดงปริมาณความเที่ยงตรงที่มองเห็นได้ วิธีแรกก็คือ การนำเครื่องมือวัดประเด็นเดียวกันมาทดลองเทียบค่าสหสัมพันธ์กับผลการวัดของเครื่องมือที่กำลังสร้างขึ้น เช่น ถ้าต้องการทราบเครื่องมือวัดชาวน์ของข้าราชการ ก็อาจนำผลการทดลองของเครื่องมือที่สร้างขึ้นนี้ ไปคำนวณหาค่าสหสัมพันธ์กับผลการทดลองวัดชาวน์โดยใช้แบบทดสอบของ Binet ที่ทดลองกับกลุ่มตัวอย่างกลุ่มเดียวกันและเวลาเดียวกัน ผลการทดลองของเครื่องมือวัดชาวน์ข้าราชการที่สร้างขึ้น กับเครื่องมือวัดชาวน์ที่มีอยู่แล้วนี้จะแสดงระดับความสามารถของเครื่องมือวัดที่สร้างขึ้นใหม่ในการวัดสิ่ง ๆ เดียวกันกับเครื่องมือวัดชาวน์ที่มีอยู่แล้ว ถ้าหากว่ามีการยอมรับว่า เครื่องมือวัดชาวน์ของ Binet มีความเที่ยงตรงในการวัดชาวน์ข้าราชการไทย สัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้ก็จะแสดงค่าความเที่ยงตรงของเครื่องมือวัดที่สร้างขึ้นใหม่ที่เชื่อถือได้

การหาความเที่ยงตรงวิธีนี้ หากไม่มีเครื่องมือวัดประเด็นเดียวกันที่เที่ยงตรงอยู่แล้ว อาจดำเนินการได้ลำบาก เพราะค่าสหสัมพันธ์ที่เทียบระหว่างเครื่องมือวัดใหม่กับเครื่องมือวัดเก่า จะเชื่อถือได้ว่าเป็นค่าความเที่ยงตรงก็ต่อเมื่อเครื่องมือวัดเก่ามีความเที่ยงตรง

อยู่แล้ว ฉะนั้น ในกรณีที่ไม่มีเครื่องมือวัด นักทฤษฎีการวัดจึงเสนอให้เทียบกับผลงานวิจัย ยกตัวอย่างเช่น ในการสร้างเครื่องมือวัดความกังวล⁹ ผู้วิจัยอาจทำการวิจัยความกังวลโดยตั้งสมมติฐานว่าบุคคลที่จะรับการผ่าตัด จะมีความกังวลมากกว่าปกติ ในการนี้ได้สร้างเครื่องมือวัดที่เรียกว่างานหมุน ซึ่งประกอบด้วยส่วนสีดำ และสีขาว ซึ่งเมื่อหมุนเร็ว ๆ แล้วจะเห็นเป็นสีเทา ในการทดลองนี้ ได้นำบุคคลกลุ่มหนึ่งที่กำลังจะได้รับการผ่าตัดมาฟังดูงานหมุนนี้ ซึ่งหมุนในอัตราความเร็วต่าง ๆ กัน ปรากฏว่าบุคคลกลุ่มนี้ มองเห็นงานหมุนเป็นสีเทา ทั้ง ๆ ที่งานหมุนช้ากว่าอัตราความเร็วที่บุคคลกลุ่มนี้มองเห็นเป็นสีเทาลงหลังจากได้รับการผ่าตัดแล้ว การที่อัตราความเร็วของงานหมุนนี้ได้ว่า บุคคลมองเห็นงานหมุนเป็นสีเทาเมื่อมีความกังวล (กำลังจะรับการผ่าตัด) แสดงว่างานหมุนมีความเที่ยงตรง เมื่อเทียบกับโอกาสที่จะผ่าตัดหรือไม่ผ่าตัด จึงกล่าวได้ว่า งานหมุนมีความเที่ยงตรงกับเครื่องมือวัดประเด็นเดียวกัน

ในตำราการวัดผลบุคคลหลายเล่ม ไม่ได้กล่าวถึงความเที่ยงตรงประเภทนี้เลย ทั้งนี้เพราะได้นำแนวคิดในเรื่องนี้ไปรวมกับวิธีคำนวณหาความเที่ยงตรงกับการวัดอื่น ๆ

ความเที่ยงตรงกับการวัดอื่น ๆ ถ้าแยกความเที่ยงตรงให้มีความเที่ยงตรงประเภทความเที่ยงตรงกับเครื่องมือวัดประเด็นเดียวกัน ความเที่ยงตรงกับการวัดอื่น ๆ ก็จะมีหมายถึงแต่เพียง ความสัมพันธ์ระหว่างเครื่องมือวัดที่สร้างขึ้นกับข้อเท็จจริงเกี่ยวกับบุคคลที่จะถูกวัด ข้อเท็จจริงเกี่ยวกับบุคคลที่จะนำมาเทียบหาความสัมพันธ์กับเครื่องมือวัดที่สร้างขึ้น อาจเป็นผลการวิเคราะห์งานของบุคคลที่กำลังปฏิบัติอยู่ ถ้าหากว่าเครื่องมือวัดมีสัมประสิทธิ์สหสัมพันธ์ร่วมกัน ผลการวิเคราะห์นั้นก็กล่าวได้ว่า เครื่องมือวัดที่สร้างขึ้นมีความเที่ยงตรงกับการวัดอื่น ในการปฏิบัติทั่วไป ผู้สร้างเครื่องมือวัดย่อมไม่มีโอกาสที่จะวิเคราะห์รายละเอียดทั้งหมดของการปฏิบัติงานที่บุคคลทุกคนกำลังทำอยู่ ฉะนั้น ผู้สร้างเครื่องมือวัดก็อาจนำวิธีการสุ่มตัวอย่างของการวิจัยมาใช้เพื่อคัดเลือกเฉพาะบุคคลที่เป็นตัวแทนในการปฏิบัติงาน ซึ่งบุคคลที่จะถูกวัดโดยเครื่องมือวัดนี้ จะต้องปฏิบัติในระยะต่อไป ตัวอย่างเช่น ถ้าต้องการสร้างเครื่องมือวัดความสามารถเริ่มของข้าราชการระดับหัวหน้าแผนก โดยให้มีความเที่ยงตรงกับการวัดอื่น ๆ ผู้สร้างก็ต้องนำเครื่องมือที่สร้างขึ้นนั้น มาทดลองกับกลุ่มตัวอย่างของข้าราชการระดับหัวหน้าแผนก เมื่อได้ผลประการใดแล้วก็อาจนำมาคำนวณเทียบกับผลการให้คะแนนความสามารถในเรื่องนี้ของผู้ที่เป็นผู้บังคับบัญชาของหัวหน้าแผนกเหล่านี้ ค่าสหสัมพันธ์ที่ได้จะแสดงความเที่ยงตรงของเครื่องมือวัด เมื่อใช้วัดบุคคลที่จะทำงานในหน้าที่หัวหน้าแผนก

⁹ *Ibid.*, pp. 114—115.

ลักษณะที่สำคัญของความเที่ยงตรงประเภทนี้ อยู่ตรงที่ว่า ความเที่ยงตรงที่ได้มิใช่เป็นการทำนายว่า บุคคลจะมีความสำเร็จในการปฏิบัติงานนั้น ๆ ในระดับใดในตัวอย่าง เรื่องหัวหน้าแผนกที่ยกมา ถึงแม้ค่าความเที่ยงตรงที่ได้จะมีปริมาณสูง เช่น 0.80 หรือ 0.85 ก็แปลได้อย่างมากแต่เพียงว่า หัวหน้าแผนกที่ดีเป็นผู้ที่มีความริเริ่ม แต่จะแปลไปไม่ได้ว่า ผู้ที่มีความสามารถริเริ่ม จะต้องเป็นผู้ที่ได้รับความก้าวหน้าในการทำงานในฐานะหัวหน้าแผนก หรือผู้ที่ขาดคุณสมบัตินี้ จะได้รับผลตรงข้าม ค่าความเที่ยงตรงประเภทนี้ เป็นค่าที่แสดงลักษณะหนึ่ง ๆ ที่เกี่ยวกับพฤติกรรมในการปฏิบัติงานเท่านั้น ไม่มีความหมายเป็นการทำนายความสามารถในการทำงานได้ หรือไม่ได้

ความเที่ยงตรงในการคาดคะเน ความเที่ยงตรงประเภทนี้ เป็นอุดมการณ์ของนักวัดผลบุคคล เครื่องมื่อวัดที่มีความเที่ยงตรงในการคาดคะเนจะต้องแสดงผลได้ว่า เกี่ยวข้องกับความสามารถทำงานได้หรือทำไม่ได้ในระยะต่อไป ระดับที่เครื่องมือวัดสามารถคาดคะเนได้ในเรื่องนี้ เป็นการแสดงคุณประโยชน์ของเครื่องมือวัดหนึ่ง ๆ

ในเรื่องความเที่ยงตรงกับการวัดอื่น ๆ ได้กล่าวไว้แล้วว่า เป็นความเที่ยงตรงที่ไม่สามารถทำนายความสามารถในการทำงานได้หรือไม่ได้ ในที่นี้ควรจะพิจารณาเหตุผลที่แสดงว่า ความเที่ยงตรงประเภทนี้ ทำนายความสามารถในการทำงานไม่ได้เสียก่อนก่อนที่จะทำความเข้าใจกับความเที่ยงตรงในการคาดคะเน การที่ความเที่ยงตรงกับการวัดอื่น ๆ ไม่มีความสามารถในการทำนายก็เพราะเหตุผลสองประการ ประการแรก ค่าของการวัดอื่น ๆ ที่นำมาเทียบกับผลการทดลองเครื่องมือวัด อาจเป็นค่าที่ผิดจากความจริง เพราะผลการทำงานของหัวหน้าแผนกที่ผู้บังคับบัญชาให้คะแนนนั้น เป็นผลที่ยังไม่ได้เพิ่มหรือลดแรงกระตุ้นของคนที่เข้าไปทำงานใหม่จริง ๆ การวิจัยในปี ค.ศ. 1953 ได้พิสูจน์ว่า บุคคลที่ดำรงตำแหน่งอยู่แล้ว ไม่ได้มีแรงกระตุ้นที่จะทำงานเท่ากับบุคคลที่จะมาทำงาน¹⁰ ประการที่สอง ผลการทำงานของหัวหน้าแผนกไม่ได้รวมผลของประสบการณ์เข้าไปด้วย เมื่อบุคคลได้ทำงานหนึ่ง ๆ ไปชั่วระยะเวลาหนึ่งแล้ว ย่อมได้เรียนรู้สิ่งต่าง ๆ ที่เกี่ยวกับงานนั้น ๆ ทำให้การแสดงผลพฤติกรรมที่เกี่ยวกับงานของตนต่างไปจากบุคคลที่ยังไม่เคยทำ ประสบการณ์นี้อาจทำให้บุคคลแสดงผลพฤติกรรมได้ดียิ่งกว่าเดิม และเร็วกว่าบุคคลที่ไม่เคยทำงาน ฉะนั้น จึงสรุปได้ว่าความเที่ยงตรงกับการวัดอื่น ๆ ไม่เกี่ยวข้องกับความเที่ยงตรงในแง่การคาดคะเนเลย

¹⁰ Guion, R.M., *Personnel Testing*, p. 127.

ถ้ากล่าวในแง่ทฤษฎี ความเที่ยงตรงในการคาดคะเนเป็นระดับความสัมพันธ์ระหว่างความแปรปรวนของเครื่องมือวัดกับความแปรปรวนของมาตรวัดความสำเร็จของงาน (criterion measure) ได้อธิบายไว้ในเรื่องตัวแปรการวัดผลแล้วว่า เครื่องมือวัดมีความแปรปรวนประการใด ส่วนมาตรวัดความสำเร็จของงานก็คือ ค่าของตัวแปรตามในเรื่องความสามารถในการทำงาน หรือพฤติกรรมในการทำงานในระยะต่อจากการวัดผล¹¹ มาตรวัดความสำเร็จของงานนี้ จะได้กล่าวโดยละเอียดในอีกบทหนึ่งต่างหาก

การหาความเที่ยงตรงในการคาดคะเน มุ่งที่จะแก้ปัญหาในเชิงปฏิบัติ เพื่อให้เครื่องมือวัดสามารถใช้ในการคัดเลือก หรือจัดบุคคลให้เหมาะแก่ตำแหน่งงาน วิธีการคำนวณหามาตรวัดความสำเร็จของงานที่จะนำมาเทียบมีอยู่หลายวิธี ความสามารถในการคาดคะเนของเครื่องมือวัดจึงมีลักษณะต่าง ๆ กันออกไปตามมาตรฐานที่นำมาใช้เทียบ วิธีง่าย ๆ ที่ใช้กันทั่วไปก็คือ วิธีการทดลองเครื่องมือวัดกับบุคคลที่มาสมัครเข้าทำงาน และเมื่อผ่านไประยะหนึ่ง หลังจากที่บุคคลบางส่วนเหล่านั้นได้เข้าปฏิบัติงานแล้ว ผู้สร้างเครื่องมือวัดก็จะรวบรวมข้อมูลที่เกี่ยวข้องกับความสามารถที่จะปฏิบัติงานได้หรือไม่ได้นำมาเทียบกับผลการทดลองวัดที่ได้รวบรวมไว้ในครั้งแรก การเปรียบเทียบค่าทั้งสองอาจคำนวณเป็นค่าสหสัมพันธ์ หรืออาจจะแสดงเป็น ไดอะแกรมตัวแปรสองตัวก็ได้ (bivariate scatter diagram)

เครื่องมือวัดที่มีความเที่ยงตรงในการคาดคะเน จะแสดงค่าที่เชื่อถือได้พอสมควรในเรื่องความสามารถในการปฏิบัติ หรือพฤติกรรมที่บุคคลหนึ่ง ๆ มีหรือปรากฏในระยะที่ผ่านมา หรือในระยะปัจจุบัน หรือในระยะต่อไป การที่จะคาดคะเนในแบบใดแบบหนึ่งนี้ ต้องอาศัยมาตรวัดความสำเร็จของงานในลักษณะต่าง ๆ ในการวัดว่า บุคคลหนึ่ง ๆ มีคุณสมบัติของพฤติกรรมอย่างไร ผู้วัดจะต้องหาค่าของมาตรวัดความสำเร็จของงานที่จะอธิบายลักษณะนิสัยที่บุคคลนั้น ๆ ได้มีในระยะก่อน การที่ใช้เครื่องมือวัดวัดคุณสมบัติทางจิตวิทยาที่มีในอดีตนี้ ดูคล้ายกับว่าจะไม่ได้ผลดีเท่ากับการพิจารณาคุณสมบัติทางจิตวิทยาของบุคคลนั้นโดยตรง แต่แท้ที่จริงแล้ว การที่จะศึกษาวิเคราะห์คุณสมบัติทางจิตวิทยาของบุคคลหนึ่ง ๆ โดยตรงมักจะทำได้ยาก หรือต้องเสียเวลามาก หรือขาดหลักฐานข้อมูล เช่น สมมติตัวอย่างง่าย ๆ ว่า เจ้าหน้าที่ในหน่วยงานแห่งหนึ่ง ต้องการทราบ ว่า บุคคลที่มาสมัครงานเคยถูกตำรวจจับเพราะขับขี่ยวดยานผิดกฎจราจร

¹¹ Good, C.V., ed., *Dictionary of Education*, p. 487.
Nunnally, *op. cit.*, p. 76.

ถ้าหากว่าไม่ใช้วิธีการวัดผล เจ้าหน้าที่ผู้นั้นก็จะต้องไปสืบข้อเท็จจริงจากสถานีตำรวจทุกแห่งในท้องที่ที่บุคคลผู้นั้นเคยขับขี่รถยนต์ไป แต่ถ้าใช้วิธีการวัดผลก็อาจสะดวกรวดเร็วยิ่งขึ้น โดยเพียงแต่ถามว่า “ท่านเคยถูกตำรวจจับเพราะฝ่าฝืนกฎจราจรหรือไม่”¹² ย่อมเป็นสิ่งแน่นอนว่า คำตอบของคำถามนี้ อาจมีทั้งคำตอบที่ผิดความจริงเพราะตั้งใจหรือเพราะความจำเลอะเลือน แต่ถ้ามีวิธีการพิสูจน์ได้ว่า คำตอบในเรื่องนี้ไม่ขัดกับประวัติของบุคคลนั้น ๆ และมีความเที่ยงตรงกับพฤติกรรมในอดีต การวัดเช่นนั้นก็ย่อมเป็นสิ่งที่น่าสนใจมากกว่าการยอมเสียเวลาไปค้นคว้าข้อเท็จจริงจากสถานีตำรวจ แต่ในการวัดผลเพื่อประโยชน์ในแง่การบริหารงานบุคคลนั้น มักมุ่งจะวัดคุณสมบัติทางจิตวิทยาของบุคคลในระยะปัจจุบันหรือในระยะต่อไปมากกว่า เพราะหวังที่จะให้บุคคลนั้น ๆ สามารถปฏิบัติงานได้ผลเมื่อบรรจุ หรือแต่งตั้งแล้ว ผลการแสดงผลพฤติกรรมตามเครื่องมือวัดที่มีความเที่ยงตรงในการคาดคะเน ย่อมจะบอกได้ว่า บุคคลหนึ่ง ๆ สามารถแก้เครื่องยนต์หรือปฏิบัติงานใดงานหนึ่งในปัจจุบัน (ในระยะที่วัดผล) ได้หรือไม่ การวัดเช่นนั้นย่อมเป็นการประหยัดเวลากว่าการที่จะให้บุคคลนั้นไปทดลองแก้เครื่องยนต์ให้ครบทุกชนิด หรือปฏิบัติงานให้ครบทุกประเภท เป็นระยะเวลาานพอสมควรที่จะแน่ใจได้ว่า บุคคลนั้นแก้เครื่องยนต์หรือปฏิบัติงานนั้นเป็นจริง โดยเฉพาะเครื่องมือวัดที่สามารถคาดคะเนพฤติกรรมในระยะต่อไปได้เที่ยงตรง ย่อมมีประโยชน์ที่สุด การคาดคะเนได้เช่นนี้ หมายความว่า จากลักษณะนิสัยหรือคุณสมบัติทางจิตวิทยาที่บุคคลนั้นมีอยู่ในปัจจุบัน บุคคลนั้นจะมีพฤติกรรมอย่างไรในสภาพการณ์หนึ่ง ๆ ในระยะต่อไป หรือกล่าวอีกอย่างหนึ่งได้ว่า การที่บุคคลนั้นมีลักษณะประจำตัวอย่างหนึ่ง ๆ นั้น ทำให้บุคคลนั้นมีความสามารถในการทำงานต่อไปอย่างไร เช่น ถ้ามีบุคคลมาสมัครเข้ารับราชการ หากจะทราบได้ว่า ต่อไปบุคคลนั้นจะสามารถปฏิบัติงานประเภทใดได้ดีที่สุด ก็ย่อมเป็นการมีประโยชน์ต่อทางราชการเป็นอย่างยิ่ง การวัดเช่นนั้น เป็นการวัดความถนัดทางธรรมชาติ ซึ่งจะกล่าวโดยละเอียดในบทที่ 10

ขั้นตอนดำเนินการในการสร้างเครื่องมือวัดงานบุคคลที่เที่ยงตรง

เครื่องมือวัดบุคคลที่ใช้ประโยชน์ในทางจิตวิทยาหรือการศึกษา มีวิธีดำเนินการหลายอย่างที่จะสร้างความเที่ยงตรงให้แก่เครื่องมือวัด ทั้งนี้แล้วแต่ประเภทของความเที่ยงตรงตามที่กล่าวมา แต่เฉพาะเครื่องมือวัดที่จะใช้ประโยชน์ในการบริหารงานบุคคลนั้น

¹² Ghiselli, *op. cit.*, p. 341.

มุ่งความเที่ยงตรงในการคาดคะเนเป็นประการสำคัญ การดำเนินการเพื่อสร้างเครื่องมือวัดที่เที่ยงตรงจึงมีแนวพิเศษออกไป อาจแบ่งอย่างหยาบ ๆ ได้เป็น 4 ชั้น คือ¹³

1. วิเคราะห์งาน
2. คัดเลือกเครื่องมือวัดเพื่อทดลอง
3. ทดลองเครื่องมือวัด
4. วัดมาตรฐานความสำเร็จของงาน

การวิเคราะห์งาน คือการที่พยายามหาข้อมูลที่ตรงประเด็น (pertinent) เกี่ยวกับงานนั้น ๆ โดยวิธีการที่ละเอียดลออ วิธีการวิเคราะห์งานมีมากมายหลายวิธี แต่โดยสรุปแล้วก็คือ การที่ได้ข้อมูลเกี่ยวกับงานหนึ่ง ๆ มาโดยการสังเกตงานนั้น ๆ โดยตรง หรือการสัมภาษณ์บุคคลที่รู้เรื่องงานนั้น ๆ ดี หรือศึกษาเอกสารข้อเท็จจริงที่อธิบายงานนั้น ๆ รายละเอียดการวิเคราะห์งานมีกล่าวไว้แล้วในบทที่ 2 แต่ในที่นี้ควรชี้ว่า การวิเคราะห์งานเพื่อประโยชน์ในการวัดผลนั้น จะต้องชี้ให้เห็นประเด็นปลีกย่อยทุกอย่างที่เกี่ยวกับงานที่บุคคลจะต้องทำให้ทราบ ว่า เครื่องมือวัดจะต้องมีสาระสำคัญประการใดบ้าง เครื่องมือวัดที่สร้างขึ้นนี้อาจมีหลายชุด ผู้สร้างเครื่องมือวัดก็จำเป็นต้องคัดเลือกทดลองแต่ละชุดให้เหมาะสม ทั้งนี้ อาจอาศัยผลการค้นคว้าที่ได้ดำเนินไปแล้วมาพิจารณาเปรียบเทียบ ในสหรัฐอเมริกา มีหนังสือที่รายงานผลการค้นคว้าในเรื่องการทดลองเครื่องมือวัดที่อาจเป็นประโยชน์ เช่น *Mental Measurements Yearbook* ที่พิมพ์เป็นประจำปี และ *Handbook of Employee Selection*¹⁴ ในประเทศไทยยังไม่มีเอกสารทางวิชาการประเภทนี้โดยตรง แต่มีเอกสารของสำนักงานคณะกรรมการข้าราชการพลเรือนที่เกี่ยวข้อง และอาจเป็นประโยชน์ในเรื่องนี้ได้บ้างคือ คู่มือการสอบข้อเขียนของกองวิชาการและสอบไล่¹⁵ จำนวนชุดของเครื่องมือวัดที่จะทดลองนี้ ขึ้นอยู่กับความสะดวกในการหาบุคคลมาเป็นกลุ่มทดลอง และความสำคัญของเนื้อหาที่จะวัด ถ้าต้องการที่จะวัดเพื่อคัดเลือกบุคคลที่ดีจริง ๆ เพื่อปฏิบัติงาน ก็จำเป็นต้องสร้างเครื่องมือวัดให้มากที่สุด และจะต้องยอมเสียเวลาในการทดลองให้มากขึ้น

ในการทดลองเครื่องมือวัดนี้ อาจทำได้เป็นสองอย่าง คือ แบบติดตามผลหรือแบบทดลองกับผู้ปฏิบัติงานในปัจจุบัน แบบติดตามผลนั้น เป็นการทดลองเครื่องมือวัดกับผู้ที่มาสมัครงาน แล้วหลังจากที่คัดเลือกบรรจุบุคคลเหล่านั้น (โดยไม่ได้ใช้ผลจาก

¹³ Lawshe, C.H., and Balma, M.J., *Principles of Personnel Testing*, p. 22.

¹⁴ *Ibid.*, p. 24.

¹⁵ สำนักงาน ก.พ., คู่มือการสอบข้อเขียน, หน้า 24.

เครื่องมือที่กำลังทดลอง) จึงวิเคราะห์ค่ามาตรฐานความสำเร็จของบุคคลเหล่านี้อีกครั้ง หนึ่ง วิธีนี้เป็นวิธีที่เปลืองเวลา เพราะต้องมีการรอรระยะที่จะวัดมาตรฐานความสำเร็จของงาน วิธีหลังทดลองกับบุคคลที่ทำงานในปัจจุบัน แล้วเทียบกับค่าของมาตรฐานความสำเร็จของงานซึ่งวิเคราะห์ได้ในทันที เพราะบุคคลเหล่านี้ปฏิบัติงานอยู่แล้ว วิธีนี้เป็นวิธีที่ใช้แพร่หลายมากกว่าวิธีแรก อย่างไรก็ตาม วิธีนี้มีจุดอ่อนในแง่ที่ว่า คุณสมบัติทางจิตวิทยาซึ่งบุคคลในกลุ่มทดลองแสดงตามเครื่องมือวัดอาจจะไม่ใช่คุณสมบัตินั้นที่บุคคลที่ยังไม่ได้ปฏิบัติงานจะมีจริง ๆ ก็ได้ วิธีแก้ในเรื่องนี้คือ พยายามคัดเลือกคนที่เป็นตัวแทนของงานจริง ๆ ไม่ใช่คัดเลือกเฉพาะบุคคลที่ดีหรือไม่ดี และไม่มีการบังคับให้มาทดลอง ทุกคนในกลุ่มทดลองจะต้องได้รับคำสั่งแจ้งว่า ตนไม่ได้กำลังถูกวัด แต่ผู้สร้างเครื่องมือวัดต้องการที่จะวัดเครื่องมือวัด ฉะนั้น ผลการทดลองจะไม่กระทบกระเทือนสถานภาพปัจจุบันของบุคคลเหล่านี้ และผู้ใดที่ไม่ประสงค์จะร่วมมือ ก็อาจจะหยุดการทดลองเมื่อใดก็ได้

หลังจากได้ผลของการทดลองเครื่องมือวัดออกมาเป็นค่าคะแนนแล้ว ผู้สร้างเครื่องมือวัดก็ต้องหาข้อมูลที่แสดงว่า การปฏิบัติงานในตำแหน่งหน้าที่นั้น ๆ จะต้องปฏิบัติในระดับใด จึงจะเรียกว่าปฏิบัติได้หรือไม่ได้ ดีหรือไม่ดี การดำเนินงานในขั้นนี้เป็นขั้นสำคัญและยากที่สุด ในขั้นนี้ กล่าวโดยย่อ ๆ ได้ว่า ผู้สร้างเครื่องมือวัดจะต้องประเมินว่าบุคคลแต่ละคนที่อยู่ในกลุ่มทดลองนั้นปฏิบัติงานได้ดีเพียงไร การปฏิบัติงานดีหรือไม่ดีนี้อาจพิจารณาได้ในแง่ปริมาณและคุณภาพของงาน หรืออาจแสดงเป็นค่าประเมินที่ผู้บังคับบัญชาของบุคคลเหล่านี้กำหนด ผู้สร้างเครื่องมือวัดจะต้องแยกประเภทให้ชัดเจนลงไปว่า บุคคลใดสามารถปฏิบัติงานได้ผลดีที่สุดกว่าบุคคลอื่น หรือบุคคลใดอยู่ในประเภทที่น่าพอใจ หรือใช้การไม่ได้

การแปลความหมายของค่าความเที่ยงตรง

ในตารางที่ 2 บทที่ 3 ได้แสดงศักยภาพในการคาดคะเนของเครื่องมือวัด ศักยภาพในการคาดคะเนก็คือ ค่าความเที่ยงตรงในการคาดคะเนนั่นเอง ค่าที่แสดงไว้ในตารางเป็นผลการค้นคว้าเปรียบเทียบเครื่องมือวัดที่มีในท้องตลาดก่อนปี ค.ศ. 1955 โดยแยกประเภทเครื่องมือวัดตามลักษณะที่จะวัด จากตารางนี้ ผู้สร้าง หรือผู้ใช้เครื่องมือวัดย่อมเห็นได้ว่า หากจะวิเคราะห์ว่าเครื่องมือวัดใดมีคุณสมบัติที่ดีในการคาดคะเน ผู้ผู้นยอมไม่มองแต่เพียงค่าสหสัมพันธ์ที่แสดงความเที่ยงตรงว่าสูงหรือต่ำ แต่ยังจะต้อง

พิจารณาความสัมพันธ์ระหว่างคุณสมบัติทางจิตวิทยาซึ่งเครื่องมือวัดนั้นวัดได้ กับงานที่บุคคลจะต้องปฏิบัติอีกด้วย

อย่างไรก็ดี โดยทั่วไป หากทุกสิ่งทุกอย่างดีพร้อม เครื่องมือวัดที่แสดงสัมประสิทธิ์ความเที่ยงตรงสูง ย่อมดีกว่าเครื่องมือวัดที่มีสัมประสิทธิ์ความเที่ยงตรงต่ำ แต่ข้อเท็จจริงยังมีอีกว่า มาตรฐานความสำเร็จของงานที่นำมาใช้มีลักษณะต่าง ๆ กัน เครื่องมือวัดที่ใช้วัดมาตรฐานความสำเร็จของงานที่มีคุณภาพต่ำอาจให้ค่าความเที่ยงตรงสูง ฉะนั้น ค่าความเที่ยงตรงสูงในกรณีเช่นนี้ก็อาจมีประโยชน์น้อยหรือแทบไม่มีเลย ในตารางที่ 16 แสดงสัมประสิทธิ์ของความเที่ยงตรงของเครื่องมือวัดที่มีในท้องตลาดบางแบบ เมื่อเทียบกับมาตรฐานความสำเร็จของงาน

ตารางที่ 16

ค่าความเที่ยงตรงเทียบกับมาตรฐานความสำเร็จของงาน

ชื่อแบบทดสอบ	ลักษณะมาตรฐานความสำเร็จของงาน	ค่าความเที่ยงตรง
Seashore Tonal Memory Test	ผลการเล่นเครื่องดนตรีที่มีสาย	0.28
Short Employment Test		
Word Knowledge Score	ผลการปฏิบัติงานของผู้ใช้เครื่องจักรทำบัญชี 80 คน	0.10
Word Knowledge Score	ผลงานของพนักงานขวเลขพิมพ์ดีด 106 คน	0.53
Arithmetic Skill Score	ผลการปฏิบัติงานของผู้ใช้เครื่องทำบัญชี 80 คน	0.26
Arithmetic Skill Score	ผลงานของพนักงานขวเลขพิมพ์ดีด 106 คน	0.60

ในตารางที่ 16 จะเห็นว่า เครื่องมือวัดงานประเภทเดียวกัน เช่น ความรู้เกี่ยวกับถ้อยคำ เมื่อเปลี่ยนมาตรฐานความสำเร็จของงานที่จะนำมาเทียบ ค่าความเที่ยงตรงเปลี่ยนไป

สัมประสิทธิ์ความเที่ยงตรงของเครื่องมือวัดหนึ่ง ๆ ตามปกติ จะมีค่าใกล้เคียง 1.00 น้อยมาก ฉะนั้นคำถามที่สำคัญก็คือว่า เครื่องมือวัด ควรจะมีค่าความเที่ยงตรงอย่างน้อยเท่าใดจึงจะเป็นที่น่าพอใจ คำถามเช่นนี้ ถ้าพิจารณาในแง่ของหน่วยงานก็เท่ากับถามว่า ถ้าหากจะใช้เครื่องมือวัดหนึ่ง ๆ เพื่อการบรรจุแต่งตั้งบุคคล จะมีโอกาสผิดพลาดในการบรรจุแต่งตั้งบุคคลไม่เหมาะสมเป็นปริมาณเท่าใด การที่จะตอบคำถามนี้ จะต้องพิจารณาทั้งค่าความเที่ยงตรงของเครื่องมือวัดนั้น ๆ ประกอบกับอัตราร้อยละของบุคคลที่จะคัดเลือกจากทั้งหมดด้วย การที่จะคัดเลือกคนเพียง 10 เปอร์เซ็นต์ ย่อมมีลักษณะต่างกับการที่จะคัดเลือกคน 90 เปอร์เซ็นต์จากจำนวนทั้งหมด ถ้าสมมติว่า จะคัดเลือกบุคคลครั้งหนึ่ง โอกาสที่จะผิดพลาดในการใช้เครื่องมือวัดที่มีค่าความเที่ยงตรงระดับต่าง ๆ จะเป็นไปตามที่แสดงไว้ในตารางที่ 17

ตารางที่ 17

โอกาสผิดพลาดของเครื่องมือวัดที่มีค่าความเที่ยงตรงระดับต่าง ๆ ในการคัดเลือกคนไว้ 50 เปอร์เซ็นต์*

ค่าความเที่ยงตรง	อัตราร้อยละที่ผิดพลาด
0.00	50.00
0.20	43.60
0.40	36.90
0.50	33.30
0.60	29.50
0.70	25.30
0.80	20.50
0.90	14.40

* แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 121.

จากตารางที่ 17 จะเห็นได้ว่า ถ้าเครื่องมือวัดมีค่าความเที่ยงตรงเป็นศูนย์ โอกาสที่จะบรรจุแต่งตั้งคนที่ไม่เหมาะสมมีถึง 50 เปอร์เซ็นต์ ถ้าจะคัดเลือกคนไว้เพียงครั้งหนึ่งของทั้งหมด ทั้งนี้เพราะเครื่องมือวัดนั้นไม่มีความเที่ยงตรงเลย การบรรจุแต่งตั้งบุคคลโดยอาศัยผลการวัดของเครื่องมือที่มีลักษณะเช่นนี้ ก็เป็นเสมือนกับการตัดสินใจโดยวิธี

โยนหัวโยนก้อย แต่ถ้าหากเครื่องมือวัดมีค่าความเที่ยงตรงถึง 0.50 โอกาสที่จะบรรจุแต่งตั้งบุคคลโดยเหมาะสมจะมีถึง 66.70 เปอร์เซ็นต์ และยิ่งเครื่องมือวัดมีค่าความเที่ยงตรงสูงขึ้นเท่าใด โอกาสที่จะอาศัยเครื่องมือวัดนี้เป็นเครื่องพิจารณาในการบรรจุแต่งตั้งบุคคลก็ยิ่งมีอัตราร้อยละที่จะถูกมากขึ้นเท่านั้น ทั้งนี้โดยมีข้อแม้โดยตลอดว่าจะคัดเลือกบุคคลไว้เพียงครั้งหนึ่งของทั้งหมด

การที่จะพิจารณาให้เห็นเด่นชัดขึ้นว่า ค่าความเที่ยงตรงของเครื่องมือวัดที่สูงขึ้นจะทำให้การตัดสินใจเกี่ยวกับงานบุคคลโดยอาศัยผลของเครื่องมือวัดนั้นๆ จึงอาจทำได้โดยเทียบค่าสถิติของโอกาสความน่าจะเป็น ระหว่างค่าความเที่ยงตรงระดับต่างๆ เมื่อมีการคัดเลือกบุคคลไว้จากจำนวนทั้งหมดในอัตราร้อยละต่างๆ กัน Lawshe ได้เสนอวิธีทำแผนภูมิความคาดหวัง (expectancy chart) ซึ่งมีทั้งประเภทบุคคล และประเภทหน่วยงาน ซึ่งผู้ที่มีความสนใจในทางสถิติอาจจะศึกษารายละเอียดได้¹⁶ ในที่นี้จะยกค่าสถิติที่จะแสดงภาพรวมว่า ความสามารถในการคาดคะเนของเครื่องมือวัดสัมพันธ์กับค่ามาตรฐานความสำเร็จประการใด ดังในตารางที่ 18

ตารางที่ 18

ความเที่ยงตรงเทียบกับค่ามาตรฐานความสำเร็จของงาน*
(แต่ละบรรทัดและสดมภ์แสดงจำนวน 1,000 ราย)

ค่าความเที่ยงตรง = 0.00					ค่าความเที่ยงตรง = 0.60				
$\frac{1}{4}$ ของมาตรฐานความสำเร็จ					$\frac{1}{4}$ ของมาตรฐานความสำเร็จ				
ส่วนที่					ส่วนที่				
$\frac{1}{4}$ ของค่าคะแนน	4	3	2	1	$\frac{1}{4}$ ของค่าคะแนน	4	3	2	1
ส่วนที่ 1	250	250	250	250	ส่วนที่ 1	45	141	277	537
2	250	250	250	250	2	141	264	318	277
3	250	250	250	250	3	277	318	264	141
4	250	250	250	250	4	537	277	141	45

¹⁶ Lawshe, *op. cit.*, pp. 300-317.

ค่าความเที่ยงตรง = 0.50					ค่าความเที่ยงตรง = 0.80				
$\frac{1}{4}$ ของมาตรวัดความสำเร็จ					$\frac{1}{4}$ ของมาตรวัดความสำเร็จ				
ส่วนที่					ส่วนที่				
$\frac{1}{4}$ ของค่าคะแนน	4	3	2	1	$\frac{1}{4}$ ของค่าคะแนน	4	3	2	1
ส่วนที่ 1	73	168	279	480	ส่วนที่ 1	6	66	253	675
2	168	268	295	279	2	66	271	410	253
3	279	295	268	168	3	253	410	271	66
4	480	279	168	73	4	675	253	66	6

* แหล่งที่มา : Thorndike, R.L., and Hagen, E., *Measurement and Evaluation in Psychology and Education*, p. 123.

ในตารางที่ 18 จะเห็นว่า แต่ละบรรทัดแสดงจำนวนของบุคคลที่แบ่งออกได้เป็น 4 ส่วน เมื่อเทียบกับมาตรวัดความสำเร็จของงาน ในช่องสดมภ์แสดงจำนวนบุคคลที่ตรงกับมาตรวัดความสำเร็จของงานในส่วนหนึ่ง ๆ ของ 4 ส่วน เช่น เมื่อค่าความเที่ยงตรงเท่ากับ 0.50 จะเห็นว่า บุคคล 480 คน จาก 1,000 คน ซึ่งเป็นบุคคลประเภทที่ได้ค่าคะแนนจากเครื่องมือวัดต่ำสุดจะสัมพันธ์กับมาตรวัดความสำเร็จของงานในส่วนที่ต่ำสุดของ 4 ส่วน เมื่อคิดเป็นอัตราร้อยละจะได้ 48 เปอร์เซ็นต์ หมายความว่า บุคคลที่ได้คะแนนต่ำสุดจำนวน 48 คน จาก 100 คน จะเป็นผู้ที่สามารถทำงานได้ดีเป็นลำดับที่ 4 ถ้าแบ่งความสามารถในการทำงานเป็น 4 ระดับ ในทำนองเดียวกัน บุคคลจำนวน 27.90 เปอร์เซ็นต์ 16.80 เปอร์เซ็นต์ และ 7.30 เปอร์เซ็นต์ ของบุคคลที่ได้คะแนนต่ำสุด จะมีความสามารถในการปฏิบัติงานเป็นลำดับที่ 3, ที่ 2 และ ที่ 1 ตามลำดับ ถ้าพิจารณาเฉพาะบุคคลที่ได้ค่าคะแนนดีที่สุด บุคคล 7.30 เปอร์เซ็นต์ 16.80 เปอร์เซ็นต์ 27.90 เปอร์เซ็นต์ และ 48.00 เปอร์เซ็นต์ที่ได้ค่าคะแนนเป็นที่ 1 ถึงที่ 25 จะมีโอกาสปฏิบัติงานได้ดีเป็นที่ 4, ที่ 3, ที่ 2 และที่ 1 ตามลำดับ ตารางนี้จะมีประโยชน์ ช่วยในการพิจารณาว่าการคัดเลือกบุคคลที่ได้ค่าคะแนนสูงหรือต่ำจากเครื่องมือวัดที่มีค่าความเที่ยงตรงระดับต่าง ๆ จะมีโอกาสได้บุคคลที่เหมาะสมตามเกณฑ์ของมาตรวัดความสำเร็จเป็นอัตราร้อยละเท่าใด

นอกจากจะพิจารณาในแง่ต่าง ๆ ดังกล่าวแล้ว ผู้ที่รับผิดชอบในการใช้เครื่องมือวัด ยังจะต้องระลึกถึงประเด็นที่กล่าวไว้ในตอนต้นบทนี้อีกด้วย คือ ค่าความเที่ยงตรงนั้นมีต่อเมื่อใช้กับงานหนึ่ง ๆ โดยเฉพาะ ฉะนั้นในการที่จะพิจารณาว่า เครื่องมือวัดหนึ่ง ๆ

จะเที่ยงตรงหรือไม่ จะต้องหวนกลับไปวิเคราะห์ความมุ่งหมายของการวัดอีกด้วย เครื่องมือวัดความสามารถในการแก้เครื่องยนต์ที่ใช้ในโรงเรียนช่างกลไม่จำเป็นต้องมีความเที่ยงตรง เมื่อนำมาใช้วัดบุคคลที่จะสมัครเข้าทำงานในตำแหน่งพนักงานแก้เครื่องยนต์

ปัญหาบางประการเกี่ยวกับความเที่ยงตรง

นอกจากปัญหาในเรื่องการจำกัดมาตรฐานวัดความสำเร็จของงานที่มีความเชื่อมั่น และความเที่ยงตรง ซึ่งจะกล่าวโดยละเอียดในบทที่ 6 โดยเฉพาะแล้วค่าความเที่ยงตรงของเครื่องมือวัด ยังมีปัญหาพิเศษที่ควรจะนำมากล่าวในที่นี้สองประการ คือ ปัญหาแรกได้แก่ ความผิดแผกกันภายในกลุ่ม (heterogeneity) ของกลุ่มตัวอย่างที่นำมาทดลองกับเครื่องมือวัด ปัญหาที่สองคือ เรื่องความยาวของเครื่องมือวัด ค่าความเที่ยงตรงมีความสัมพันธ์กับค่าความเชื่อมั่นอย่างใกล้ชิดดังอธิบายมาแล้วในตอนต้นของบทนี้ และในบทที่ 4 ก็อาจเห็นแล้วว่า พิสัย (range) ของความสามารถและพฤติกรรมของบุคคลในกลุ่มที่ใช้ทดลองกับเครื่องมือวัด มีผลกระทบต่อค่าความเชื่อมั่น ในกรณีของความเที่ยงตรง ความผิดแผกกันภายในกลุ่ม จึงมีอิทธิพลทำให้ค่าที่คำนวณได้แตกต่างออกไปด้วย เมื่อก้าวโดยทั่วไปแล้ว ถ้าพิสัยของความสามารถและพฤติกรรมของบุคคลในกลุ่มที่ใช้ทดลองต่ำลง โอกาสที่จะได้สัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัด กับมาตรฐานวัดความสำเร็จของงานก็จะต่ำลงตามลงไปด้วย ยิ่งถ้าหากบุคคลในกลุ่มทดลองมีความสามารถและพฤติกรรมเหมือนกันทุกประการ ค่าสหสัมพันธ์ที่ได้ก็ต้องเป็นศูนย์ ฉะนั้น ในการทดลองเครื่องมือวัดงานบุคคล ผู้สร้างเครื่องมือวัดจะต้องคำนึงถึงพิสัยของความสามารถและพฤติกรรมของกลุ่มบุคคลที่จะเลือกมาทดลองนี้เสมอ

ในเรื่องความยาวของเครื่องมือวัดนั้น ค่าความเที่ยงตรงก็จะมีโอกาสแปรปรวนเช่นเดียวกับค่าความเชื่อมั่น ต่างกันแต่ว่า ค่าความเที่ยงตรงเปลี่ยนไปตามความยาวของเครื่องมือวัดมากกว่าค่าความเชื่อมั่น ในเรื่องความเชื่อมั่นนั้นได้อธิบายไว้แล้วว่า เครื่องมือวัดที่ยาวเท่าใด ก็ยิ่งเชื่อมั่นได้มากขึ้น ในกรณีของความเที่ยงตรงก็เช่นเดียวกัน แต่เนื่องจากว่าเครื่องมือวัดที่เชื่อมั่นไม่ได้จะเที่ยงตรงไม่ได้ ฉะนั้น ค่าความเที่ยงตรงจะมีมากขึ้นเมื่อเครื่องมือวัดยาวมากขึ้น เครื่องมือวัดที่ยาวขึ้นโดยส่วนที่ยาวขึ้นมีความคล้ายคลึงกับกลุ่ม (homogeneous) จะทำให้ค่าความแปรปรวนของผลที่ถูกต้องในการวัด (true variance) สูงขึ้น และค่าความแปรปรวนของผลที่ผิดพลาดในการวัด (error variance) ต่ำลง ปรากฏการณ์เช่นนี้จะมีผลให้ค่าความแปรปรวนขององค์ประกอบร่วมกัน (common factor variance) เพิ่มขึ้น ซึ่งเป็นการเพิ่มค่าความเที่ยงตรงนั่นเอง

สูตรที่ใช้คำนวณค่าความเที่ยงตรงที่เพิ่มหรือลดตามความยาวของเครื่องมือวัดนั้นดังต่อไปนี้¹⁷

$$r_{nxy} = \frac{r_{xy}}{\sqrt{\frac{1-r_{xx}}{n} + r_{xx}}}$$

r_{nxy} = ค่าความเที่ยงตรง (ค่าสหสัมพันธ์ระหว่างเครื่องมือวัด x กับมาตรฐานความสำเร็จ y) เมื่อเครื่องมือวัดยาวเป็น n เท่า

r_{xy} = ค่าความเที่ยงตรงเดิม (เมื่อความยาวยังไม่เปลี่ยนแปลง)

r_{xx} = ค่าความเชื่อมั่นของเครื่องมือวัด x

n = อัตราที่ความยาวของเครื่องมือวัด x เพิ่ม

ถ้าหากนำสมการข้างบนมาคำนวณ จะทำให้ทราบได้ว่า เครื่องมือวัดที่สร้างขึ้นนั้น ควรจะมีความยาวเท่าใด จึงจะได้ค่าความเที่ยงตรงตามต้องการ

$$\therefore r_{nxy} = \frac{r_{xy}}{\sqrt{\frac{1-r_{xx}}{n} + r_{xx}}}$$

$$\therefore \frac{1-r_{xx}}{n} = \frac{r_{xy}^2}{r_{nxy}^2} - r_{xx}$$

$$n = \frac{1-r_{xx}}{\frac{r_{xy}^2}{r_{nxy}^2} - r_{xx}}$$

จากผลที่ได้ใหม่นี้จะเห็นว่า การพิจารณาปัญหาว่า เครื่องมือวัดควรจะยาวเท่าใด จึงจะมีค่าความเที่ยงตรงในระดับที่เชื่อมั่นได้นั้น ต้องอาศัยว่าเครื่องมือวัดที่ทดลองนั้น เดิมมีความเที่ยงตรง เมื่อเทียบกับมาตรฐานความสำเร็จของงานอยู่แล้วในระดับใด ประกอบกับข้อเท็จจริงที่ว่าเครื่องมือวัดนั้นเชื่อมั่นได้เพียงใด ถ้าหากว่าระดับของค่าทั้งสอง สูงอยู่แล้ว อัตราความยาวที่ควรจะเป็นที่จะต้องเพิ่มขึ้น ก็อาจจะไม่จำเป็นต้องสูง

¹⁷ Ghiselli, *op. cit.*, p. 407.

บทที่ 6

มาตรวัดความสำเร็จของงาน

การวัดในการบริหารงานบุคคลมีลักษณะพิเศษกว่าการวัดทั่วไปทางจิตวิทยาในแง่ที่ว่า การวัดทางจิตวิทยาเป็นเพียงการวัดลักษณะที่ต้องกรวัดเท่านั้น แต่การวัดในการบริหารงานบุคคลเป็นการวัดเพื่อให้ได้ประโยชน์ในการคาดคะเนความสามารถในการปฏิบัติงานของบุคคลต่าง ๆ อีกด้วย เช่น ในการวัดสติปัญญา นักวิชาการในการวัดเพื่อการบริหารงานบุคคลจะไม่ประสงค์แต่เพียงที่จะวัดเขาวัดเขาโดยทั่ว ๆ ไป แต่จะมุ่งที่จะทราบว่า การปฏิบัติงานหนึ่ง ๆ จำเป็นต้องมีสติปัญญาในด้านใดบ้างและอย่างไร การที่จะทำได้เช่นนั้น นักวิชาการในการวัดจะต้องวิเคราะห์พฤติกรรมในการปฏิบัติงาน (job behavior) เพื่อให้ทราบว่า บุคคลที่จะปฏิบัติงานในตำแหน่งหนึ่ง ๆ อย่างได้ผลสำเร็จจะต้องแสดงพฤติกรรมที่เกี่ยวข้องโดยตรงกับงานนั้น ๆ อย่างไรบ้าง พฤติกรรมนี้จะเป็นตัวแปรที่นำมาเทียบค่ากับผลการทดลองใช้เครื่องมือวัด เพื่อให้ทราบว่า เครื่องมือวัดหนึ่ง ๆ มีคุณค่าในการคาดคะเนว่าบุคคลจะปฏิบัติงานในตำแหน่งต่าง ๆ ได้ผลสำเร็จประการใด การที่นำพฤติกรรมในการปฏิบัติงานมาวิเคราะห์ค่าเทียบคุณค่าของเครื่องมือวัดเช่นนั้น จึงได้มีการเรียกพฤติกรรมเฉพาะในเรื่องนี้ว่า มาตรวัดความสำเร็จของงาน (criterion measure)

เครื่องมือวัดบุคคลจะมีค่าในการบริหารงานบุคคล ก็ต่อเมื่อเครื่องมือวัดนั้น ๆ ช่วยให้ผู้บริหารงานคาดคะเนได้ว่า ผู้ถูกวัดจะแสดงพฤติกรรมเช่นใดเมื่อได้รับการแต่งตั้งให้ทำงานหนึ่ง ๆ การที่จะทำได้เช่นนั้น ก็ต่อเมื่อนักวิชาการในการวัดได้กำหนดมิติของมาตรวัดความสำเร็จของงานไว้เรียบร้อยแล้ว และเครื่องมือวัดที่ใช้กันนั้น ๆ สามารถจัดลำดับหรือแบ่งกลุ่มผู้ถูกวัดในแง่พฤติกรรมการทำงานให้สอดคล้องกับมิติที่ได้กำหนดไว้

ในบทก่อนได้อธิบายลักษณะต่าง ๆ ของค่าความเที่ยงตรงของแบบทดสอบมาตรฐานตลอดจนวิธีการที่ใช้ในการวิเคราะห์ค่านี้ รายละเอียดเรื่องมาตรวัดความสำเร็จของงานที่กล่าวไว้ในบทนี้ก็คือ ค่าอธิบายตัวแปรภายนอก (external variable) ที่จะนำมาใช้เทียบกับค่าของผลการทดสอบที่ได้จากเครื่องมือวัดนั้นเอง

แนวคิดในการสร้างมาตรวัดความสำเร็จของงาน การที่กล่าวว่า ให้ใช้พฤติกรรมในการปฏิบัติงานเป็นมาตรวัดความสำเร็จของงานเพื่อเทียบคุณค่าของเครื่องมือวัดนั้นพิจารณาดูเผิน ๆ ก็ไม่น่าจะมีปัญหาอะไร แต่โดยความเป็นจริงแล้วเป็นปัญหาที่ยุงยากมาก เพราะพฤติกรรมการปฏิบัติงานที่จะนำมาใช้เป็นตัวแทนของความสำเร็จผลในการปฏิบัติงานจริง ๆ มีลักษณะยากต่อการจำกัดความในการที่จะให้ได้ค่าที่จะคำนวณเทียบได้กับเครื่องมือวัด ยิ่งไปกว่านั้นยังมีปัญหาอื่น ๆ ตามมาอีกหลายปัญหา การวิเคราะห์ค่าพฤติกรรมของงานหนึ่ง ๆ ควรวิเคราะห์เพื่อให้มีมาตร ๆ เพียงอันเดียว ซึ่งจะมีลักษณะแสดงการปฏิบัติงานทุก ๆ ด้านหรือจะต้องให้มีมาตร ๆ หลาย ๆ อันที่เป็นอิสระแก่กัน แต่เป็นมาตร ๆ ที่วัดพฤติกรรมต่าง ๆ ของงานตำแหน่งเดียวกัน ปัญหาต่อไปก็คือว่า มีตัวแปรพื้นฐานอะไรอยู่บ้างที่เกี่ยวพันกับมาตรฐานการประเมินค่าความสำเร็จในการปฏิบัติงาน ปัญหาที่เกี่ยวข้องกับเรื่องที่ว่า ควรพิจารณาตัวแปรในด้านเศรษฐกิจที่สัมพันธ์กับผลการปฏิบัติงานหรือควรถือตัวแปรในเรื่องความประพฤติในการทำงานเป็นหลักในการกำหนดมาตร ๆ และอีกเรื่องหนึ่งที่ว่า ตำแหน่งงานที่วัดนั้นมีลักษณะสอดคล้องกับนโยบายของหน่วยงานอย่างไร เมื่อทราบวิธีแก้ปัญหเหล่านี้ มาตร ๆ ที่วิเคราะห์มาได้จึงจะมีคุณค่าเป็นมาตร ๆ ที่ดี

มาตรวัดความสำเร็จของงานชั้นสุดท้าย ในทางปฏิบัติผู้สร้างมาตรวัดความสำเร็จของงานย่อมไม่สามารถสร้างมาตร ๆ ที่มีคุณสมบัติครบถ้วนดีทุกประการโดยแก้ไขปัญหาดังต่าง ๆ ที่กล่าวมาแล้วได้ทั้งหมด อย่างไรก็ตาม การสร้างมาตร ๆ ตามหลักวิชาที่ต้องตั้งความมุ่งหมายให้ได้มาตร ๆ ที่มีความเที่ยงตรงและสมบูรณ์ที่สุด ฉะนั้น Thorndike¹ ซึ่งเป็นนักวิชาการในการวัดผลงานบุคคลชั้นนำของสหรัฐ ฯ จึงได้เสนอให้แบ่งมาตร ๆ ออกเป็น 3 ชั้น คือ ชั้นสุดท้าย ชั้นกลาง และชั้นต้น มาตร ๆ ชั้นสุดท้าย คือ มาตร ๆ ในอุดมคติในการคัดเลือกบรรจุแต่งตั้ง หรือในการคัดเลือกผู้ผ่านการฝึกอบรม เพื่อให้ได้บุคคลที่ดีที่สุดได้ปฏิบัติงาน มาตร ๆ ชั้นกลาง คือ มาตร ๆ ที่หาขนาดแทน (substitute criterion) ซึ่งได้มาโดยการวิเคราะห์ให้ค่าการปฏิบัติงานโดยมีหลักการ ส่วนมาตร ๆ

¹ Thorndike, Robert L., *Personnel Selection*, pp. 121-124.

ขั้นต้น หมายถึงข้อเท็จจริงเกี่ยวกับการปฏิบัติงานเป็นผลสำเร็จหรือไม่สำเร็จที่อาจหา
พิจารณาได้ในทันที

มาตราฯ^๒ ขั้นสุดท้าย มีลักษณะเป็นนามธรรมที่ประมวลทุกสิ่งทุกอย่างซึ่งอธิบายว่า
การปฏิบัติงานได้ผลสำเร็จอย่างใดนั้นมิชอบเขต และความลึกซึ้งแค่ไหน² Thorndike
ยกตัวอย่างว่า มาตราฯ^๓ ขั้นสุดท้ายของการขายประกันชีวิตก็คือ ปริมาณกรมธรรม์จำนวน
มากที่สุดจำนวนหนึ่งที่ผู้ขายประกันจะต้องขายได้ ในช่วงระยะเวลาที่กำหนดไว้ช่วง
หนึ่ง การกำหนดมาตราฯ^๔ นี้จะต้องคำนึงถึงปริมาณกรมธรรม์ที่ขายได้ อัตราการต่ออายุ
สัญญากรมธรรม์ ความคงเส้นคงวาของการปฏิบัติงาน ผลกระทบที่มีต่อการขายกรม
ธรรม์ของผู้ขายประกันคนอื่น ๆ ผลการปฏิบัติงานเกี่ยวกับเอกสารและหลักฐานสัญญา
ต่าง ๆ ความร่วมมือในการช่วยฝึกอบรมพนักงานใหม่ และลักษณะการประชาสัมพันธ์
งาน การที่จะรวบรวมข้อมูลทั้งหมดที่ได้ ก็จะต้องดำเนินการตลอดชั่วอายุการปฏิบัติ
งานของผู้ขายประกัน ซึ่งเป็นการเสียเวลามาก และยิ่งถ้าเป็นงานที่วัดปริมาณได้โดย
ยากกว่าการขายประกันชีวิต แม้จะรอนชว่อายุการปฏิบัติงานของบุคคลหนึ่ง ๆ ก็ยัง
อาจไม่ได้ข้อมูลครบถ้วน ถึงแม้จะสามารถรวบรวมข้อมูลได้ในชั่วอายุการปฏิบัติงาน
ช่วงระยะเวลาที่ผ่านมาไปก็ทำให้ข้อมูลหย่อนคุณภาพลงในแง่ความเป็นจริง

ถึงแม้มาตราฯ^๕ ขั้นสุดท้ายจะเป็นสิ่งที่วิเคราะห์ออกมาได้โดยยากก็ตาม ผู้สร้างมาตราฯ^๖
ก็มีหน้าที่สร้างมาตราฯ^๗ ที่มีลักษณะเข้าข่ายเป็นมาตราฯ^๘ ขั้นสุดท้ายให้มากที่สุดที่จะมากได้
ในการนี้ อาจใช้หลักสถิติมาประกอบในการพิจารณาความสำเร็จของงาน จึงทำให้เกิด
มาตราฯ^๙ ที่มีลักษณะทดแทนซึ่งเรียกว่า มาตราฯ^{๑๐} ชั้นกลาง เช่น ถ้าหากผู้รู้จักงานที่จะวัด
หนึ่ง ๆ ได้ตกลงเห็นพ้องกันว่า การที่จะปฏิบัติงานให้สำเร็จด้วยดี ผู้ปฏิบัติงานจะ
ต้องมีพฤติกรรมอย่างหนึ่งอย่างใดที่เห็นได้ชัดในช่วงเวลาหนึ่ง ข้อตกลงเห็นพ้องกันนี้
ก็จะเป็นข้อมูลที่ใกล้เคียงกับมาตราฯ^{๑๑} ขั้นสุดท้ายมากที่สุด หลังจากที่ได้มีการเทียบค่า
พฤติกรรมที่กำหนดนี้ออกมาเป็นตัวเลขที่นับได้ และพิสูจน์ค่าสหสัมพันธ์ระหว่างค่า
พฤติกรรมดังกล่าวกับการประเมินการปฏิบัติงานจริง ๆ ว่ามีความเกี่ยวข้องกันเชื่อถือได้
ผู้สร้างมาตราฯ^{๑๒} ก็อาจนำเอาผลการประเมินการปฏิบัติงานจริง ๆ มาถือเป็นมาตราฯ^{๑๓} ชั้นกลาง
เพื่อใช้เทียบกับเครื่องมือวัดของตนได้ อย่างไรก็ตาม การพิสูจน์ค่าสหสัมพันธ์ระหว่าง
ผลการประเมินการปฏิบัติงานกับมาตราฯ^{๑๔} ขั้นสุดท้าย เป็นงานที่ยุ่งยากและเปลืองเวลา
มาก เพราะจะต้องมีการติดตามผลการประเมินเป็นระยะ ๆ เป็นเวลานาน ๆ และยังมี
ปัญหาในเรื่องการกำหนดมาตราฯ^{๑๕} ขั้นสุดท้าย ข้อมูลที่นำมาใช้สร้างมาตราฯ^{๑๖} ชั้นกลางจึงมัก

² Ibid.

จะต้องอาศัยหลักการวิเคราะห์หาเหตุผล (rational basis) เป็นแนวในการพิจารณาว่าจะเชื่อถือได้หรือไม่ การวิเคราะห์หาเหตุผลนี้ทำได้โดยพิจารณาข้อมูลทั้งหมดอย่างรอบคอบว่าเป็นเหตุเป็นผลกันตามหลักการวิจัยหรือไม่ และมีความสอดคล้องกับเป้าหมายในการคัดเลือกบุคคลเพื่อปฏิบัติงานในหน่วยงานหนึ่ง ๆ เพียงใด นอกจากนี้ยังอาจใช้วิธีการทางสถิติหาค่าสหสัมพันธ์ระหว่างข้อมูลต่าง ๆ ที่จะนำมาพิจารณาเป็นมาตรฐานชั้นกลาง แม้ว่าข้อมูลเหล่านี้มิได้เป็นมาตรฐานชั้นสุดท้าย และค่าสหสัมพันธ์ที่ได้ก็ไม่ใช่ค่าสหสัมพันธ์ของข้อมูลที่มีอยู่กับค่ามาตรฐานอันใดก็ตาม ค่าสหสัมพันธ์ระหว่างกันของข้อมูลก็อาจช่วยเสริมให้เห็นจริงว่า เหตุและผลที่ได้ใช้เป็นหลักในการพิจารณาข้อมูลในขั้นตอนก่อนหน้านั้น มีหลักฐานสนับสนุนพอที่จะเชื่อได้ว่าข้อมูลต่าง ๆ ที่มีความสอดคล้องกันทำให้ใช้ประกอบการพิจารณาสร้างมาตรฐานได้

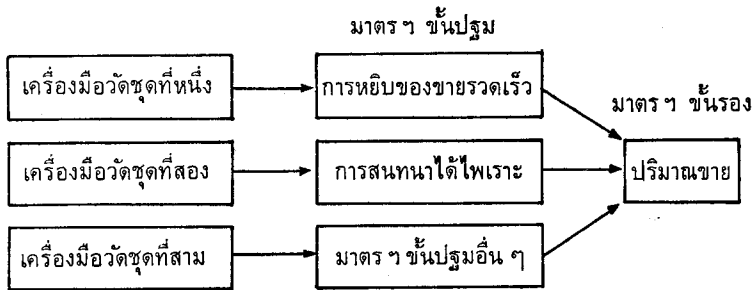
ข้อมูลเกี่ยวกับผู้ปฏิบัติงาน ที่สามารถรวบรวมได้ในทันทีถือว่าเป็นมาตรฐานชั้นต้น เช่น ความเห็นของผู้บังคับบัญชาที่ได้จากการประเมินในทันทีที่ผู้ปฏิบัติงานเริ่มรับตำแหน่งหน้าที่ หรือผลการทดลองปฏิบัติงานในระยะเวลาสั้น ๆ ข้อมูลเหล่านี้เป็นเพียงส่วนหนึ่งของข้อเท็จจริงที่ว่าผู้ปฏิบัติงานได้ทำงานในหน้าที่ของตนได้ดีเพียงไร ต่อเมื่อได้มีการวิเคราะห์เปรียบเทียบข้อมูลเหล่านี้กับผลการปฏิบัติงานในช่วงระยะเวลา จึงจะสามารถสร้างมาตรฐานที่มีคุณภาพดีขึ้น เพื่อทำให้เป็นมาตรฐานชั้นกลางต่อไปดังกล่าวมาแล้ว

มาตรวัดความสำเร็จของงานชั้นปฐมและชั้นรอง (primary and secondary criteria) Guion³ ได้อ้างผลการค้นคว้าวิจัยของนักวิชาการกว่า 10 คน ที่ยืนยันว่า มาตรวัดความสำเร็จของงาน พิจารณาได้เป็นหลายมิติ (dimensions) ฉะนั้น การที่ Thorndike กำหนดให้มีมาตรฐานชั้นสุดท้าย จึงเป็นการขัดต่อความเป็นจริงในเรื่องมิติ อย่างไรก็ตาม ความคิดในเรื่องมาตรฐานชั้นสุดท้ายมิได้จำกัดว่า ผู้สร้างมาตรฐาน จะต้องมองมาตรฐานชั้นสุดท้ายนี้ให้เป็นมิติเดียว ฉะนั้นจึงกล่าวในทันทีได้ว่า ความเห็นเรื่องมาตรฐานชั้นสุดท้ายนี้ยังลบล้างไม่ได้ ส่วนความเห็นในเรื่องมิติ เป็นเรื่องที่น่าสนใจและควรนำมาพิจารณาประกอบกันไปในการสร้างมาตรฐานด้วย ในแง่มิติ มาตรฐานประกอบด้วยหลายมิติ ทั้งในแง่เวลาที่ปฏิบัติงาน และในแง่ตัวบุคคลแต่ละคนที่ครองตำแหน่งหน้าที่ ในแง่เวลา อาจแบ่งเป็นสองระยะ เช่น ระยะที่เริ่มปฏิบัติงานและระยะเมื่อปฏิบัติงานไปแล้วช่วงเวลาหนึ่ง ผลผลิตของงานในสองระยะนี้ต้องต่างกัน ในแง่ตัวบุคคลแต่ละคนบำเพ็ญตนต่าง ๆ กัน เช่น พนักงานขายของในตำแหน่งเหมือนกัน คนหนึ่งหิวของให้ลูกค้าได้รวดเร็ว

³ Guion, Robert M., *Personnel Testing*, pp. 114-115.

แต่อีกคนหนึ่งทำให้ลูกค้าสนใจและสนทนากับลูกค้าได้เพราะ มิติเหล่านี้ทำให้ต้องมีเครื่องมือวัดหลาย ๆ อย่าง เพื่อให้สามารถคาดคะเนผลการปฏิบัติงานในแง่ต่าง ๆ ได้

การที่ความสำเร็จของงานมีมิติหลายมิติเช่นนี้ กล่าวได้อีกอย่างหนึ่งว่าผู้ปฏิบัติงานแต่ละคนอาจมีพฤติกรรมไม่เหมือนกัน แต่ก็สามารถปฏิบัติงานให้ได้ผลสำเร็จเหมือนกัน ในตัวอย่างเรื่องพนักงานขายในย่อหน้าก่อน ถ้าพนักงานทั้งสองคนจะทำให้บริษัทได้รับประโยชน์ในทางธุรกิจคำนวณเป็นปริมาณเงินได้เท่า ๆ กัน พฤติกรรมของพนักงานขายที่ต่างกันเช่นนี้ ก็ถือได้ว่าเป็นพฤติกรรมที่ทำให้ปฏิบัติงานสำเร็จได้ผลดี จึงแยกได้ว่า พฤติกรรมที่แตกต่างกันแต่ละบุคคลในระหว่างผู้ปฏิบัติงานเป็นมาตราฯ ชั้นปฐมและประโยชน์ในทางธุรกิจของบริษัทเป็นมาตราฯ ชั้นรอง มาตราฯ ชั้นปฐมอาจมีได้ตั้งแต่สองขึ้นไป ทำให้ต้องมีเครื่องมือวัดตามจำนวนของมาตราฯ ที่มีอยู่ ลักษณะของความสัมพันธ์ระหว่างมาตราฯ ทั้งสองประเภทกับเครื่องมือวัดมีแสดงไว้ในแผนภาพที่ 6



แผนภาพที่ 6 มาตราฯ ชั้นปฐมและชั้นรอง

ไม่ว่าจะใช้แนวคิดของ Thorndike หรือของ Guion ประเด็นสำคัญก็อยู่ตรงที่ว่า จะประเมินอะไรจึงจะได้ค่าเป็นตัวแทนความสำเร็จของงานที่แท้จริง สิ่งที่จะต้องปฏิบัติก็คือต้องมีการวิเคราะห์งาน (job analysis) ซึ่งได้มีกล่าวรายละเอียดไว้ในบทที่ 2 แล้ว เมื่อได้วิเคราะห์งานแล้วจึงจะเก็บข้อเท็จจริงในการปฏิบัติงานมาเทียบค่าประกอบการสร้างมาตราฯ ต่อไป ตามที่ได้ถือปฏิบัติกันมา การเก็บข้อเท็จจริงนี้อาจทำได้ในวิธีต่าง ๆ ก็คือ เก็บข้อมูลโดยวิธีปรนัย ทั้งนี้อาจทดสอบความรู้และความเข้าใจงาน หรือหาหลักฐานการปฏิบัติงานที่เป็นจริง หรือสังเกตตัวอย่างงาน มิฉะนั้น ก็อาจถือเอาข้อวินิจฉัยส่วนบุคคลเป็นข้อมูลในการสร้างมาตราฯ

มาตรวัดความสำเร็จของงานที่ได้โดยวิธีปรนัย

มาตร ๑ ที่นำมาใช้ควรจะต้องเป็นมาตร ๑ ที่ได้มาโดยวิธีปรนัย เพื่อให้ค่าของมาตร ๑ นั้น ไม่บิดเบือนหรือผันผวนตามอารมณ์หรือความรู้สึกของผู้ประเมินค่า พฤติกรรมการปฏิบัติงาน หรือกล่าวอีกนัยหนึ่งว่า หากได้มาตร ๑ มาโดยวิธีปรนัยย่อมมีโอกาสที่จะได้มาตร ๑ ที่มีค่าความเชื่อมั่นสูงกว่ามาตร ๑ ที่ได้โดยวิธีอื่น

แต่เป็นที่น่าเสียดายว่ามาตร ๑ ที่จะได้มาโดยวิธีปรนัยนั้น มีลักษณะจำกัดที่ไม่สามารถเป็นตัวแทนลักษณะทุกอย่างของพฤติกรรมการปฏิบัติงาน ข้อเสียในเรื่องความจำกัดนี้จะพิจารณาเห็นได้จากลักษณะของข้อมูลที่จะได้มาโดยวิธีนี้ ข้อมูลที่จะนำมาพิจารณาประกอบเป็นมาตร ๑ ประเภทนี้ แยกได้เป็น ๓ ประเภท คือ คะแนนจากการทดสอบความรู้และความเข้าใจงาน (examination of knowledge and understanding) หลักฐานการปฏิบัติงานที่เป็นจริง (objective performance record) และผลจากการสังเกตตัวอย่างงาน (observer-scored job performance)

คะแนนการทดสอบความรู้และความเข้าใจงาน ตำราส่วนใหญ่ไม่นับข้อมูลประเภทนี้ไว้ในจำพวกที่จะใช้พิจารณาหามาตรวัดความสำเร็จของงาน ทั้งนี้อาจเป็นเพราะความรู้และความเข้าใจงานนั้นมีลักษณะห่างจากการปฏิบัติงานให้สำเร็จ จึงไม่อาจถือว่ามีความสัมพันธ์กับมาตร ๑ ขั้นสุดท้ายได้อย่างแท้จริง แต่ Thorndike⁴ ก็ยังอ้างว่า ข้อมูลประเภทนี้อาจนำมาใช้สร้างมาตร ๑ ได้ และมาตร ๑ นี้จะมีความเชื่อถือได้ว่าไม่บิดเบือนเพราะอคติของผู้ประเมิน และโดยเฉพาะงานในตำแหน่งหน้าที่ที่เป็นนามธรรม (abstract) ควรจะพิจารณาเทียบหามาตร ๑ ด้วยข้อมูลประเภทนี้อย่างยิ่ง เพราะเป็นข้อมูลที่ให้ค่าความสำเร็จของงานอย่างใกล้ชิดที่สุด การทดสอบความรู้ และความเข้าใจงานอาศัยข้อสอบปรนัยเป็นหลัก ลักษณะและวิธีการสร้างข้อสอบอาจนำเอาหลักการของการทดสอบความสัมฤทธิ์มาใช้ได้ (achievement test) เรื่องนี้จะมีกล่าวโดยละเอียดต่อไปในบทที่ 10 ในที่นี้อาจกล่าวโดยย่อได้ว่า ผู้สร้างข้อสอบวัดความรู้และความเข้าใจงาน จะต้องแน่ใจว่าเนื้อหาของข้อสอบครอบคลุมความรู้ทั้งหมดที่ผู้ปฏิบัติงานที่รู้งานนี้จะต้องมีตามสัดส่วนที่แท้จริง คำถามแต่ละข้อจะต้องถามประเด็นที่สำคัญ และการถือว่าประเด็นไหนสำคัญนั้นจะต้องเป็นที่ตกลงกันระหว่างผู้ออกข้อสอบและหัวหน้างาน หรือผู้เชี่ยวชาญในงานตำแหน่งนั้น ๆ การที่จะจัดทำข้อสอบนี้ให้สมบูรณ์ได้จึงต้องอาศัยความร่วมมือของบุคคลสามฝ่าย คือ นักจิตวิทยาในการวัดงานบุคคล นักวิเคราะห์งานหรือ

⁴ Thorndike, *op. cit.*, pp. 145—149.

ผู้เชี่ยวชาญในด้านตำแหน่งงาน และหัวหน้างานหรือผู้บังคับบัญชาโดยตรง นักจิตวิทยาในการวัดงานบุคคลมีหน้าที่นำความเห็นและข้อวิพากษ์วิจารณ์ของนักวิเคราะห์งาน และหัวหน้างาน มาดำเนินการตั้งหัวข้อคำถามตามแบบของการทดสอบที่ดี

หลักฐานการปฏิบัติงานที่เป็นจริง ข้อมูลประเภทนี้ได้แก่บันทึกข้อเท็จจริงเกี่ยวกับผลการปฏิบัติงานที่เก็บไว้เป็นหลักฐานประจำตัวผู้ปฏิบัติงาน ข้อเท็จจริงนี้ต้องเป็นข้อเท็จจริงที่นับได้หรือสังเกตได้ (countable or observable)⁵ ฉะนั้นพฤติกรรมในการปฏิบัติงานที่รวบรวมมาเป็นข้อมูลประเภทนี้ จึงเป็นเพียงพฤติกรรมบางด้านเท่านั้น ไม่ใช่พฤติกรรมทั้งหมด และเป็นพฤติกรรมของงานบางตำแหน่ง ไม่ใช่พฤติกรรมของงานทุก ๆ ตำแหน่ง พฤติกรรมที่ได้มาเป็นข้อมูลประเภทนี้ อาจแยกออกได้เป็น 2 ประเภท คือ ข้อเท็จจริงเกี่ยวกับผลิตผลงาน และข้อเท็จจริงในด้านงานบุคคล

ข้อเท็จจริงในเรื่องผลิตผลงานหมายถึง ปริมาณและคุณภาพของงานที่ทำได้รวมทั้งความสามารถในการรับการฝึกอบรม (trainability) ข้อเท็จจริงในเรื่องนี้ต้องเป็นข้อเท็จจริงที่นับได้และแสดงออกมาเป็นตัวเลขได้ ด้านปริมาณของผลงาน ผู้สร้างมาตรฯ อาจนับจำนวนชิ้นของงานที่ผู้ปฏิบัติงานทำสำเร็จในหนึ่งชั่วโมง หรือ วัน หรือสัปดาห์ เช่น ปริมาณปอที่คัดเลือกในหนึ่งชั่วโมง ปริมาณกระดาษที่พิมพ์ดีดในหนึ่งวันหรือปริมาณการขายของพนักงานขายในหนึ่งปี หรืออาจพิจารณาปริมาณของผลงานในแง่เวลาก็ได้ คือ ปริมาณเวลาที่ต้องใช้ในการปฏิบัติงานชิ้นหนึ่ง ๆ เช่น ปริมาณเวลาที่ใช้ในการดำเนินการโครงการวิจัยหนึ่ง ๆ ในด้านคุณภาพ หมายถึงว่าผู้ปฏิบัติงานคนหนึ่ง ๆ ปฏิบัติงานได้ดีเพียงไร แทนที่จะพิจารณาในแง่ที่ว่าแต่ละคนปฏิบัติงานได้มากและได้เร็วเพียงใด แต่การวัดว่าบุคคลปฏิบัติงานได้ดีเพียงไรเป็นเรื่องทำได้ยาก ฉะนั้นในการลงบันทึกข้อเท็จจริงเกี่ยวกับคุณภาพของงานจึงเล็งพิจารณาไปในแง่ที่ว่า การปฏิบัติงานหนึ่ง ๆ สอดคล้องกับอุปสงค์ของงานเพียงใด (job demands)⁶ กล่าวคือ วัดว่า มีชิ้นงานส่วนใดบ้างที่ใช้ไม่ได้ หรือต้องปฏิบัติใหม่ ค่าใช้จ่ายที่ต้องเสียไปในการทำงานที่ใช้ไม่ได้เสียใหม่เป็นปริมาณเท่าใด เช่น คุณภาพงานของพนักงานพิมพ์ดีด ก็หมายถึงจำนวนคำที่พิมพ์ผิด คุณภาพของสมุหบัญชี หมายถึง ปริมาณข้อผิดพลาดที่ตรวจพบในระยะเวลาหนึ่ง ๆ ส่วนความสามารถในการรับการฝึกอบรมนั้นเป็นผลการบันทึกความเป็นจริงว่า ผู้ปฏิบัติงานมีความเร็วในการเรียนรู้งานเพียงไร ความสามารถนี้คือ ปริมาณเวลาที่ผู้ปฏิบัติงาน

⁵ Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, pp. 35—37. Guion, *cp. cit.*, p. 91.

⁶ *Ibid.*, p. 92.

ต้องใช้ในการที่จะสร้างความรู้ความสามารถ (proficiency) ให้แก่ตนเองในการปฏิบัติงานหนึ่ง หรือกล่าวอีกนัยหนึ่งก็คือ ระดับความรู้ความสามารถของผู้ปฏิบัติงานในระยะเวลาที่กำหนดให้ช่วงหนึ่ง ๆ ในงานที่อุปสงค์และหน้าที่ความรับผิดชอบค่อนข้างตายตัว (static) ระยะเวลาที่ผู้ปฏิบัติงานต้องใช้ในการสร้างความรู้ความสามารถมีปริมาณค่อนข้างน้อย งานประเภทนี้ได้แก่งานง่าย ๆ ที่ต้องทำซ้ำ ๆ ซาก ๆ เช่น งานขับรถรับจ้าง ใช้เวลาเพียงประมาณ 18 สัปดาห์ ก็สามารถมีความรู้ ความสามารถในการปฏิบัติงานเป็นอย่างดี (หลังจากขับรถเป็นแล้ว) แต่งานที่มีความสลับซับซ้อน รวมทั้งงานที่ใช้สมอง จะต้องใช้เวลานานกว่าจะมีความรู้ความสามารถเต็มที่ เช่น ผู้ปฏิบัติหน้าที่พนักงานขายเพื่อการลงทุนธุรกิจ (investment salesman) สามารถเพิ่มความรู้ความสามารถในการปฏิบัติงานได้เรื่อย ๆ ไปจนถึงสิบปี⁷ ข้อมูลในด้านนี้จะช่วยให้ทราบว่าผู้ปฏิบัติงานที่สำเร็จด้วยดี ใช้เวลาเท่าใดจึงจะมีความรู้ความสามารถในการรับการฝึกอบรม มิใช่เป็นข้อมูลประเภทเดียวกับความรู้ความสามารถในการผลิตผลงาน

ข้อเท็จจริงในด้านงานบุคคล หมายถึงพฤติกรรมในการปฏิบัติงานในส่วนที่ไม่เกี่ยวกับเนื้อหาของงานโดยตรง แต่พฤติกรรมประเภทนี้ก็สามารถแสดงได้ว่า ผู้ปฏิบัติงานที่ถือได้ว่าดีมีคุณลักษณะเช่นใด พฤติกรรมนี้ได้แก่การทดสอบการลา ระยะเวลาดำรงตำแหน่ง อัตราความเร็วในการเลื่อนชั้น เลื่อนตำแหน่ง และอุบัติเหตุในระหว่างการปฏิบัติงาน ในเรื่องการขาดการลา อาจลงบันทึกหลักฐานในแง่จำนวนชั่วโมงและวันที่ขาดลา หรือจำนวนครั้งที่ขาดลาก็ได้ หลักฐานการขาดลานี้มีประโยชน์ในแง่ที่ว่า *ceteris paribus* ผู้ที่มาปฏิบัติงานประจำโดยสม่ำเสมอ ย่อมมีประโยชน์ต่อหน่วยงานมากกว่าผู้ที่ขาด ๆ ลา ๆ ข้อเท็จจริงเรื่องระยะเวลาดำรงตำแหน่งทำให้ทราบว่า ผู้ใดบ้างปฏิบัติงานกับหน่วยงานหนึ่ง ๆ ได้ยั่งยืนหรือไม่ยั่งยืนเพียงใด ข้อมูลนี้จะช่วยให้ผู้สร้างมาตรฐานมีบรรทัดฐานที่จะกล่าวได้ว่า ผู้ใดค่าคะแนนตามเครื่องมือวัดจะปฏิบัติงานกับหน่วยงานอย่างน้อยก็ช่วงเวลาที่กำหนดไว้หรือไม่ อัตราความเร็วในเรื่องการเลื่อนชั้นเลื่อนตำแหน่งช่วยให้ทราบว่า บุคคลต่าง ๆ ที่ปฏิบัติงานหนึ่ง ๆ มีความมีโอกาสที่จะได้เลื่อนชั้นเลื่อนเงินเดือนและตำแหน่งอย่างไร (promotability) ความมีโอกาสนี้เป็นข้อเท็จจริงในประเด็นที่ว่า แต่ละคนต้องเสียเวลาเท่าใดจึงจะได้รับการเลื่อน และสามารถนำไปเทียบกับเครื่องมือวัด ทำให้แยกแยะได้ว่า ผู้ถูกวัดจะมีความมีโอกาสได้รับการเลื่อนในระยะเวลาที่กำหนดไว้หรือไม่ ส่วนอุบัติเหตุในระหว่างการปฏิบัติงานนั้น เป็นหลักฐานที่

⁷ *Ibid.*, p. 93.

อาจเทียบกับเครื่องมือวัดที่สร้างขึ้น เพื่อให้แยกแยะได้ว่าผู้ใดมีโอกาสประสบอุบัติเหตุ
มากน้อยเพียงใด (accident proneness)

โดยหลักการ หลักฐานการปฏิบัติงานที่เป็นจริงย่อมใช้เป็นข้อมูลในการสร้างมาตรา
ที่ดี เพราะมาตราขั้นสุดท้ายก็คือ ผลการปฏิบัติงานที่ตนเอง ข้อเท็จจริงที่บันทึกไว้
ย่อมมีความเป็นปรนัย (objectivity) มากกว่าผลการปฏิบัติที่ประเมินได้ด้วยวิธีอื่น ๆ
เพราะได้เก็บประมวลไว้เป็นระยะยาวและรวบรวมจากความเป็นจริง แต่ข้อมูลก็มีข้อ
จำกัดอยู่ 2 ประการ คือ ในเรื่องผลิตผลงานโดยตรงนั้น ข้อเท็จจริงที่บันทึกเป็นข้อ
มูลเฉพาะที่นับได้เท่านั้น แต่ในงานต่าง ๆ นั้น ก็มีงานที่มีผลงานที่นับไม่ได้อยู่หลาย
อย่าง อีกประการหนึ่ง ข้อเท็จจริงนี้มีตัวแปรหลายอย่าง (variables) ที่ทำให้นามาแปล
ความหมายเปรียบเทียบกัน เพื่อถือเป็นบรรทัดฐานได้ยาก ตัวแปรเหล่านี้ได้แก่สภาพ
การต่าง ๆ ที่เกี่ยวข้องกับการปฏิบัติงานหนึ่ง ๆ กล่าวคือ ผลิตผลการปฏิบัติงานของ
บุคคลหนึ่ง ๆ นั้นไม่ได้ขึ้นอยู่กับความสามารถและความตั้งใจพยายามของผู้ปฏิบัติงาน
แต่ฝ่ายเดียว เช่น ผู้ปฏิบัติงานที่ต้องใช้เครื่องจักร คุณภาพและความเร็วของเครื่องจักร
มีส่วนในการกำหนดปริมาณงานที่บุคคลหนึ่ง ๆ จะทำสำเร็จ ในกรณีที่ปฏิบัติงานกันเป็น
คณะ หรืองานที่ปฏิบัติต้องมีความเกี่ยวข้องกันหลายฝ่าย การปฏิบัติตนของผู้อื่นก็มีผล
กระทบต่อปริมาณ และคุณภาพของผลิตผลงานของผู้ปฏิบัติงานแต่ละคน

ผลจากการสังเกตตัวอย่างงาน ในโอกาสที่ผู้สร้างมาตราฯ ไม่สามารถค้นหาหลักฐาน
การปฏิบัติงานที่เป็นจริงได้ และเห็นว่าข้อตัดสินและความเห็นของผู้บังคับบัญชาขาด
ความเป็นปรนัยเชื่อถือไม่ได้ ดังจะกล่าวรายละเอียดถึงในหัวข้อต่อไป ผู้สร้างมาตราฯ
ก็อาจสุ่มตัวอย่างของงานที่ปฏิบัติในตำแหน่งหนึ่ง ๆ และไปทำการสังเกตด้วยตนเอง
ในระหว่างสังเกตผู้สร้างมาตราฯ ก็จะทำให้ค่าคะแนนของการปฏิบัติงานไปด้วย หลักการที่
สำคัญในเรื่องนี้ คือ การที่ผู้สร้างมาตราฯ ได้ใช้ความระมัดระวังในการเลือกตัวอย่างที่จะ
ไปสังเกต ให้เป็นส่วนหนึ่งของงานที่บุคคลทุกคนที่ปฏิบัติงานในหน้าที่นั้นต้องทำเสมอ
เหมือนกัน และเมื่อได้ทำเช่นนั้นแล้ว ในระหว่างการสังเกตก็ต้องพยายามกำหนด
มาตรฐานในการให้ค่าคะแนนการปฏิบัติงาน เพื่อให้ได้ผลที่มีความเป็นปรนัยมากที่สุด⁸
ในการปฏิบัติจริง ๆ นั้น ผู้สร้างมาตราฯ ย่อมไม่สามารถหลีกเลี่ยงอุปสรรคทั้งสอง
ประการนี้ได้อย่างสิ้นเชิง ในกรณีที่เป็งานง่าย ๆ ปฏิบัติไปตามคัลลอง (routine) ก็อาจ

⁸ Thorndike, *op. cit.*, pp. 141—145.
Lawshe and Balma, *op. cit.*, pp. 51—52.

เลือกตัวอย่างงานที่เป็นตัวแทนจริง ๆ ของการปฏิบัติงานได้โดยสะดวก และยังสามารถกำหนดคุณค่าความสำคัญของการปฏิบัติงานส่วนนั้นในขั้นตอนต่าง ๆ ได้โดยมีอคติของผู้สร้างมาตรฯ ไปปะปนเป็นส่วนน้อย แต่ในกรณีที่เป็นงานสลับซับซ้อนนอกจากจะมีความลำบากในการกำหนดมาตรฐานการปฏิบัติงานแล้ว ผู้สังเกตยังไม่สามารถทำตัวให้เป็นเสมือนเครื่องจักรที่บันทึกผลการปฏิบัติงานส่วนนั้น ๆ ในขั้นตอนต่าง ๆ โดยไม่มีความลำเอียง เพราะในการให้ค่าของการปฏิบัติ ผู้สร้างมาตรฯ จำต้องแปลความหมายของการกระทำต่าง ๆ ในการปฏิบัติงาน โดยอาศัยดุลพินิจของตนทำให้ค่าที่วัดออกเป็นผลงานมีน้ำหนักต่าง ๆ กันไปตามแต่ผู้สร้างมาตรฯ แต่ละคน เช่น ถ้าจะสังเกตตัวอย่างงานของคนขับรถโดยสารสาธารณะก็อาจทำได้โดยสะดวก ผู้สร้างมาตรฯ อาจกำหนดตัวอย่างงานไว้ว่า มีการขับถอยหลัง การเลี้ยว การติดเครื่อง และการเปลี่ยนเกียร์บนทางชัน ความนุ่มนวลและการปฏิบัติในเรื่องเหล่านี้โดยไม่บกพร่องในระดับต่าง ๆ ก็จะทำให้ผู้สร้างมาตรฯ สามารถกำหนดค่าเป็นตัวเลขได้ทันที แต่การปฏิบัติงานในหน้าที่ผู้ตรวจการหรืองานอื่นที่มีความซับซ้อนมากกว่า ย่อมทำให้ข้อมูลประเภทนี้มีความเที่ยงตรงและความเชื่อมั่นลดลงไปตามส่วน

มาตรวัดความสำเร็จของงานที่ได้จากการวินิจฉัยส่วนบุคคล

เพื่อที่จะให้ได้ทราบพฤติกรรมทุกด้านที่ปรากฏในการปฏิบัติงาน ผู้สร้างเครื่องมือวัดจำต้องอาศัย ภาพประทับใจ (impression) ที่บุคคลที่สามมีต่อผู้ปฏิบัติงาน บุคคลที่สามนี้อาจจะเป็นผู้บังคับบัญชาของผู้ปฏิบัติงานโดยตรง หรืออาจเป็นผู้ที่มีหน้าที่ประเมินความประพฤติและการปฏิบัติของผู้ดำรงตำแหน่งหน้าที่หนึ่ง ๆ ก็ได้ ภาพประทับใจจำเป็นต้องเป็นสิ่งที่ดีค่าออกมาเป็นปริมาณได้ ในการนี้จึงได้มีวิธีการให้ค่า (rating scale) ในรูปต่าง ๆ กัน วิธีการให้ค่านี้กำหนดลักษณะนิสัยหนึ่ง ๆ ของผู้ปฏิบัติงานเป็นตัวหลักที่ผู้ปฏิบัติงานคนหนึ่ง ๆ จะได้รับคะแนนสูงหรือต่ำตามแต่ลักษณะการประพฤติและปฏิบัติตนของผู้ปฏิบัติงานนั้น ๆ วิธีการให้ค่านี้ทำได้ในหลายลักษณะ ลักษณะวิธีการให้ค่าที่ทำได้สะดวกและแพร่หลายมาก คือ วิธีการให้ค่าแบบกราฟ (graphic rating scale)⁹ ดังตัวอย่างในแผนภาพที่ 7

⁹ Thorndike, Robert L., and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*, pp. 350—351.

ลักษณะพฤติกรรม	ค่าพฤติกรรม			
	ดีกว่าคนอื่น เสมอ	ดีกว่าคนอื่น บางครั้ง	เฉลี่ยแล้ว ปานกลาง	เฉลี่ยแล้ว ใช้ไม่ได้
ความรับผิดชอบใน การทำงานให้เสร็จ
ความแม่นยำ
ความขยันหมั่นเพียร
ความประณีต

แผนภาพที่ 7 ตัวอย่างวิธีการให้ค่าแบบกราฟ

ข้อบกพร่องของวิธีการให้ค่า เนื่องจากวิธีการให้ค่าในลักษณะต่าง ๆ มีความสัมพันธ์
เกี่ยวข้องกับประเมินบุคลิกภาพ รายละเอียดของวิธีการนี้ จึงจะกล่าวไว้ในบท
ที่ 10 ในที่นี้จะพิจารณาแต่ในแง่ข้อเสียของการใช้วิธีนี้เพื่อสร้างมาตรวัดความสำเร็จ
ของงาน

เมื่อทราบข้อบกพร่องของวิธีนี้แล้ว จึงจะได้เปรียบเทียบผลการค้นคว้าวิจัยในการ
ที่จะแก้ไขการใช้วิธีนี้ให้ได้ผลในการที่จะได้พฤติกรรมที่แท้จริงในการปฏิบัติงาน

ข้อเสียของวิธีการให้ค่าโดยตรงที่ความบกพร่องของผู้ให้ค่า (rater) ความบกพร่องนี้
แยกได้เป็น 2 นัย คือ ผู้ให้ค่าอาจขาดความเต็มใจที่จะให้ค่าอย่างตรงไปตรงมา หรือผู้
ให้ค่าอาจขาดความสามารถที่จะให้ค่าได้อย่างเที่ยงตรง¹⁰

ในเรื่องการขาดความเต็มใจที่จะให้ค่าอย่างตรงไปตรงมา เป็นปัญหาที่พิจารณาแล้ว
ดูจะขัดความเป็นจริง เพราะย่อมเป็นที่เข้าใจกันว่า ผู้ให้ค่าแต่ละคนย่อมพยายามอย่าง
ที่สุดที่จะปฏิบัติตามระเบียบข้อบังคับในการให้คะแนน และพยายามที่จะให้ความเป็น
ธรรมแก่ผู้ปฏิบัติงานในตำแหน่งที่ถูกประเมินทุกคน ถ้าหากจะมีการผิดพลาดในการ
ประเมินงานก็จะต้องเป็นเพราะความบกพร่องส่วนบุคคล ไม่ใช่ด้วยความตั้งใจที่จะไม่
ให้ความยุติธรรม แต่ตามความเป็นจริงทางวิชาการนั้น ผู้ให้ค่าก็อาจจะให้คะแนนงานที่
ประเมินอย่างบิดเบือนด้วยความตั้งใจได้ การที่มีการบิดเบือนในการให้ค่าคะแนนนี้ อาจ
เป็นเพราะเหตุหนึ่งเหตุใดจากสาเหตุสองประการ คือ ประการแรก ผู้ให้ค่าไม่ยอมอุทิศ

¹⁰ Remmers, H.H., et al., *A Practical Introduction to Measurement and Evaluation*, pp. 359—360.
Thorndike and Hagen, *op. cit.*, pp. 338—344.

คนที่จะลำบากในการดำเนินการอย่างละเอียดถี่ถ้วนในการให้ค่าคะแนน หรือ ประการที่สองผู้ให้ค่าคะแนนมีความสัมพันธ์กับผู้ปฏิบัติงานที่ถูกประเมิน ในลักษณะที่ทำให้ผู้ให้ค่าไม่ประสงค์ให้ผู้ถูกประเมินต้องเสียประโยชน์ เป็นความจริงที่ว่า การให้คะแนนในเรื่องใด ๆ ก็ตาม หากจะให้เป็นที่เที่ยงตรงจะต้องเสียเวลามาก เพราะจะต้องมีการวิเคราะห์พฤติกรรมในเรื่องนั้นออกมาเป็นประเด็นที่สำคัญ แล้วกำหนดประเด็นเหล่านั้นเป็นเกณฑ์ในการที่จะนำผู้ถูกประเมินมาเทียบกัน จะเห็นได้จากแบบกรอกรายการที่ใช้ในการประเมินข้าราชการนั้นมีค่าเป็นเพียงแบบฟอร์มเท่านั้น ผู้บังคับบัญชาที่มีหน้าที่ตอบแบบกรอกรายการนี้มิได้สะดวกเวลาจริง ๆ ในการพิจารณาเทียบค่าของพฤติกรรมในการปฏิบัติงาน ผลการประเมินจึงเป็นเพียงภาพประทับใจที่เป็นอัตนัยของผู้บังคับบัญชามิได้วัดค่าการปฏิบัติงานจริง ๆ ตลอดปีของผู้อยู่ใต้บังคับบัญชา เมื่อถึงขั้นตอนที่พิจารณาให้ค่าความดีความชอบ จึงอาศัยการอภิปรายในที่ประชุมของผู้บังคับบัญชาเป็นหลัก มิได้ถือข้อเท็จจริงของแบบกรอกรายการนี้เป็นหลัก การที่เป็นเช่นนั้น จึงทำให้มีผู้กล่าวว่า เป็นเพราะการพิจารณาความดีความชอบของข้าราชการอาศัยการประชุมดังกล่าว จึงไม่จำเป็นต้องเสียเวลากรอกแบบรายการประเมินงานให้ละเอียดถี่ถ้วนเที่ยงตรง อย่างไรก็ตามการประเมินพฤติกรรมของงานเป็นเรื่องที่จะต้องยอมอุทิศเวลาหากประสงค์จะให้ค่าประเมินเที่ยงตรง โดยผู้บังคับบัญชาจะต้องหมั่นประเมินผู้อยู่ใต้บังคับบัญชาเป็นระยะ ๆ อยู่เสมอ และเมื่อถึงวาระที่จะต้องลงบันทึกให้ค่าประเมิน ก็สามารถทำได้โดยมิได้อาศัยความรู้สึกชั่วขณะหรือความลำเอียง แต่อาศัยหลักฐานการประเมินครั้งต่าง ๆ เปรียบเทียบกัน ส่วนใหญ่ผู้มีหน้าที่ให้ค่าพฤติกรรมการปฏิบัติงานก็ได้แก่ผู้บังคับบัญชาชั้นต้น ทั้งนี้เพราะเจ้าหน้าที่ฝ่ายงานบุคคลโดยตรง เช่น สำนักงาน ก.พ. หรือ หัวหน้าฝ่ายการเจ้าหน้าที่ของบริษัทธุรกิจย่อมเป็นผู้ที่อยู่ห่างไกลผู้ปฏิบัติงานมากกว่าหัวหน้าแผนกหรือหัวหน้ากองของผู้ปฏิบัติงานนั้น ๆ ในกรณีเช่นนั้น จึงเป็นเหตุให้เกิดผลเสีย อีกประการหนึ่ง คือ ในบางครั้งผู้บังคับบัญชาจะพยายามให้ค่าคะแนนแก่บุคคลในบังคับบัญชาสูง เพื่อไม่ให้ผู้น้อยหน้าบุคคลที่อยู่ใต้บังคับบัญชาของผู้อื่น แทนที่จะพยายามให้ค่าคะแนนตรงกับพฤติกรรมของงาน และโดยเฉพาะในการที่การประเมินค่าพฤติกรรมของงานเป็นเรื่องเกี่ยวข้องกับการเลื่อนขั้นอันดับเงินเดือนโดยตรง ก็ยังทำให้ผู้บังคับบัญชาต้องพยายามรักษาผลประโยชน์ของบุคคลในกลุ่มของตนในหน่วยงานทั่ว ๆ ไป ค่าคะแนนของพฤติกรรมในการปฏิบัติงานตำแหน่งต่าง ๆ จึงมักมารวมกลุ่มกันอยู่ในระดับใดระดับหนึ่ง ทำให้ไม่สามารถแยกแยะออกไปได้ว่า พฤติกรรมในการปฏิบัติงานตำแหน่งใดควรจะได้รับค่าคะแนนสูงกว่าตำแหน่งอื่น ๆ เท่าใด

ในด้านความสามารถของผู้ให้คำที่กล่าวว่ายังมีทางบกพร่องนั้น เป็นเพราะเหตุหนึ่งเหตุใดในสาเหตุ 5 ประการดังต่อไปนี้ คือ ความกำกวม (ambiguity) ของลักษณะพฤติกรรมที่จะต้องประเมิน ลักษณะนิสัยที่จะให้คำนั้นเห็นไม่ได้เด่นชัดจากการสังเกตภายนอก ผู้ให้คำไม่มีโอกาสหรือไม่ได้ใช้โอกาสเต็มที่ในการสังเกตพฤติกรรม เกณฑ์จะให้คำไม่ตรงกันระหว่างผู้ให้คำและลักษณะพิเศษประจำตัว (idiosyncrasies) ของผู้ให้คำแต่ละคน¹¹

การที่กล่าวว่าลักษณะพฤติกรรมที่ประเมินมีความกำกวมก็เพราะว่า ในการประเมินโดยทั่วไป ผู้ให้คำจะต้องให้คะแนนรายการพฤติกรรมที่มีความหมายกว้าง ๆ เช่นที่แสดงไว้ในแผนภาพที่ 7 คำที่ว่า ความขยันหมั่นเพียร หรือความรับผิดชอบในการทำงานให้เสร็จ มีความหมายต่อผู้ให้คำแต่ละคนต่าง ๆ กันไป ที่ว่าขยันนั้นหมายถึงว่า อยู่ปฏิบัติหน้าที่ในสำนักงานตลอดเวลาราชการ และเมื่อผู้บังคับบัญชาเรียกตัวเมื่อเวลาบ่ายสี่โมงครึ่งก็ยังพร้อมที่ไปรายงานตัวได้ หรือหมายความว่าปฏิบัติงานที่รับมอบหมายเสร็จเรียบร้อยก่อนเวลาทุกครั้ง ในกรณีที่ผู้ให้คำถือความหมายแรกเป็นเกณฑ์บุคคลบางคนก็จะได้ค่าคะแนนสูง แต่เมื่อผู้ให้คำถือความหมายประการหลัง ผู้ที่ได้ค่าคะแนนสูงจะกลายเป็นบุคคลต่างประเภทออกไป สภาพการณ์เช่นนี้ย่อมทำให้การให้คำพฤติกรรมขาดความคงเส้นคงวา ปัญหาที่ว่าลักษณะนิสัยที่จะให้คำ เห็นไม่ได้เด่นชัดจากการสังเกตภายนอก หมายความว่าถึงลักษณะนิสัยหลายอย่างที่ใช้เป็นเครื่องอธิบายพฤติกรรม เป็นสิ่งที่ซ่อนอยู่ในตัวผู้ปฏิบัติงาน อากัปกริยาท่าทางที่แสดงออกภายนอกเป็นเพียงเครื่องสะท้อนอย่างผิวเผิน จะถือเอาเป็นลักษณะที่แท้จริงไม่ได้ เช่น ผู้ให้คำถือความเชื่อมั่นในตัวเองของผู้ปฏิบัติงาน จะถือเอาเพียงน้ำเสียงในการพูด และความสามารถในการสนทนาของบุคคลหนึ่งเป็นหลักในการตัดสินใจไม่สมควร ผู้ให้คำจะประเมินลักษณะนี้ได้ก็ต่อเมื่อได้คลุกคลีสังเกตผู้ปฏิบัติงานเป็นเวลานาน ๆ และอาศัยความรู้ทางจิตวิทยาแยกแยะความประพฤติของผู้ทำงาน จึงจะสรุปได้ว่าผู้ปฏิบัติงานนั้น ๆ มีลักษณะนิสัยอย่างไร ปัญหานี้ก็เกี่ยวข้องไปถึงข้อเท็จจริงอีกประการหนึ่งที่ว่า ผู้ให้คำไม่มีโอกาสคลุกคลีสังเกตผู้ปฏิบัติงานอย่างแท้จริง ผู้บังคับบัญชามีหน้าที่ให้คำพฤติกรรมในงานต่างก็มีผู้อยู่ใต้บังคับบัญชาหลายคน ถึงแม้ว่าจะมีเวลาปฏิบัติงานร่วมกันเป็นเวลาหลายชั่วโมงในวันหนึ่ง ๆ ผู้บังคับบัญชาก็มีภารกิจอื่น ๆ ที่สำคัญกว่าที่จะต้องทำให้สำเร็จ ปัญหาจึงเกิดขึ้นว่า ผู้ให้คำหรือผู้บังคับบัญชามีโอกาสคลุกคลีสังเกตผู้ปฏิบัติงานในสภาพงานต่าง ๆ กันจนครบทุกแง่ทุกมุม เชื่อได้ว่าการให้คำ

¹¹ Thorndike and Hagen, *op. cit.*, pp. 540-544.

ลักษณะนิสัยต่าง ๆ มีคุณค่าเชื่อถือได้อย่างแท้จริงหรือไม่ ถึงหากจะแก้ปัญหาทั้งสามประการนี้ได้ ก็ยังมีปัญหาใหญ่อื่นๆ อีก คือ ผู้ให้ค่าจะต้องค่าพฤติกรรมออกมาเป็นปริมาณซึ่งไม่สามารถพิสูจน์ว่าวัดได้ตรงกันระหว่างผู้ให้ค่าเอง เช่น การให้ค่าลักษณะอันหนึ่งอันใดของพฤติกรรม จะระบุว่าผู้ปฏิบัติงานมีคุณสมบัติตามลักษณะนั้น ๆ อยู่ในระดับ ดีมาก หรือ ปานกลาง หรือใช้ไม่ได้ การที่ว่าดีมากนั้นหมายความว่าอย่างไร หมายความว่า เป็นที่ดีที่สุดในระหว่างผู้ปฏิบัติงานทั้งหมด หรือเป็นคนหนึ่งในสองสามคนที่ดีกว่าบุคคลที่เหลือ และที่ว่าดีมากนี้เทียบกับผู้ปฏิบัติงานทั้งหมดของแผนกหรือกอง หรือทั่ว ๆ ไป และได้เทียบโดยพิจารณาข้อได้เปรียบหรือเสียเปรียบของการที่บุคคลผู้หนึ่งได้ปฏิบัติงานโดยระยะเวลาที่แตกต่างกันและในสภาพการณ์ที่ไม่เหมือนกันกับบุคคลอื่น ๆ หรือเปล่า ปัญหาสุดท้ายเป็นปัญหาประจำตัวผู้ให้แต่ละคน กล่าวคือ ผู้ให้แต่ละคนชอบหรือรังเกียจพฤติกรรมและปรากฏการณ์ต่าง ๆ ในลักษณะที่ไม่เหมือนกัน ประสบการณ์ทำให้ผู้ให้ค่าคนหนึ่งถือว่าพฤติกรรมของบุคคลคนหนึ่งแสดงความอ่อนแอหรือความใจดำ ในขณะที่ผู้ให้ค่าอีกคนหนึ่งมิได้ถือเช่นนั้น กล่าวอีกอย่างหนึ่งได้ว่า ผู้ให้ค่าแต่ละคนมีแนวโน้มที่จะมีอคติ (bias) ซึ่งมีผลกระทบต่อการประเมินรายการละเอียดของลักษณะนิสัยต่าง ๆ ที่ระบุไว้ในพฤติกรรมการปฏิบัติงาน

ข้อบกพร่องต่าง ๆ เหล่านี้ทำให้ผลการประเมินพฤติกรรมในการปฏิบัติด้วยการให้ค่าตามวิธีที่ใช้ทั่ว ๆ ไปเกิดผลเสีย สามประการ คือ ความบิดเบือนอันเกิดจากอคติ (halo) ความเบ้ให้การให้คะแนน (leniency) และความโน้มสู่ศูนย์กลาง (central tendency) ความบิดเบือนจากอคติคือ การที่ผู้ให้ค่าถือความรู้สึกส่วนตัวอันเกิดจากอคติที่มีต่อผู้ปฏิบัติงานเป็นแนวในการให้ค่าแก่ลักษณะนิสัยทุกอย่างของพฤติกรรม ความเบ้ในการให้คะแนน เป็นลักษณะการที่ผู้ให้ค่าอาจเป็น “คนกระตือรือร้น” หรือ “คนให้คะแนนง่าย” ทำให้ค่าทำให้เกิดความเบ้ (skewed distributions of ratings) ส่วนความโน้มสู่ศูนย์กลางนั้นเป็นการที่ค่าคะแนนมักจะมารวมกระจุกกันอยู่ไม่แยกกระจาย รายละเอียดในเรื่องนี้มีกล่าวไว้ในตำราและเอกสารทางวิชาการหลายแห่งแล้วจึงจะไม่กล่าวซ้ำอีก

การประเมินด้วยการบังคับเปรียบเทียบบุคคล (employee comparisons) การที่วิธีการให้ค่าพฤติกรรมในการปฏิบัติงานตามวิธีทั่ว ๆ ไปไม่มีความบกพร่องและผลเสียดังกล่าว นักวัดบุคคลหลายท่าน เช่น S.N. Stevens, E.F. Wonderlic, R.K. Rosensteel W.K. Kirchner, และ M.D. Dunette เป็นต้น ได้พยายามค้นคว้าวิธีใหม่เพื่อจะให้ได้มาตรฐาน ๆ ที่ไม่มีข้อบกพร่องต่าง ๆ โดยกำหนดหลักการว่า เมื่อจะประเมินพฤติกรรมงานให้ประเมินทุก ๆ คนพร้อมกันที่ละลักษณะนิสัย เพื่อที่จะให้ผู้ประเมินต้องมีการเปรียบเทียบบุคคล

ทั้งหมดจริง ๆ วิธีการใหม่นี้ก็เท่ากับเอาวิธีให้ค่าเดิมมาปรับปรุงใหม่นั้นเอง ตามหลักการใหม่นี้จะมีวิธีประเมินใหญ่ ๆ 2 วิธี คือ วิธีจัดลำดับที่ (ranking) และวิธีจัดบัญชีตรวจสอบ (checklist)¹²

วิธีจัดลำดับที่อาจทำได้ 3 อย่าง ประการแรก จัดลำดับที่สลับ (alternation ranking) อย่างที่สอง จัดลำดับที่โดยเทียบคู่ (paired comparisons) และอย่างที่สาม จัดลำดับที่ตามกำหนดอัตราส่วน (forced distribution) การจัดลำดับที่สลับทำได้โดยผู้ประเมินทำบัญชีชื่อผู้ที่ถูกให้ค่าเป็นแบบสุ่ม (random order) แล้วนำชื่อเหล่านั้นมาพิจารณาว่าผู้ใดดีที่สุด และผู้ใดเลวที่สุดในกลุ่ม พิจารณาไล่ไปที่ละคู่เช่นนี้ ก็จะได้บุคคลที่สมควรได้รับการประเมินให้อยู่ในอันดับแรกไล่ลงไป และในขณะเดียวกันก็ได้บุคคลที่สมควรได้รับการพิจารณาให้อยู่ในอันดับสุดท้ายไล่ขึ้นมา ในการจัดลำดับที่โดยเทียบคู่ ทำได้โดยนำบุคคลแต่ละคนไปพิจารณาเทียบคู่กับบุคคลอื่น ๆ ทุกคนในกลุ่มที่ละคน เช่น ถ้ามีบุคคลในกลุ่ม 10 คน ก็จะมีการจับคู่ 45 คู่ หรือถ้ามีบุคคลในกลุ่ม 150 คน ก็จะมี 11,175 คู่ การที่มีคู่มากเช่นนี้ ทำให้วิธีการจัดลำดับตามแบบนี้ต้องอาศัยเครื่องจักรกล (computer) ส่วนวิธีจัดลำดับที่แบบกำหนดอัตราส่วนนั้นอาศัยหลักการแจกแจงปกติ (normal distribution) กล่าวคือ ตามหลักสถิตินั้นทราบอยู่แล้วว่า ปรากฏการณ์ของสิ่งต่าง ๆ จะเกิดขึ้นโดยมีอัตราส่วนสัมพันธ์กับค่าเบี่ยงเบนมาตรฐาน (standard deviation) จากเกณฑ์เฉลี่ย (mean) ออกไปทั้งข้างบวกและลบเป็นปริมาณ .5 ของค่าเบี่ยงเบนมาตรฐานจะมีกรณีของปรากฏการณ์เกิดขึ้นเป็นอัตราประมาณร้อยละ 38.30 และเมื่อตัดค่าเบี่ยงเบนนี้ออกไปไม่ว่าจะเป็นทางด้านบวกหรือลบอีก .7 ของค่าเบี่ยงเบนมาตรฐานก็มีกรณีปรากฏการณ์เกิดขึ้นอีกเป็นอัตราประมาณร้อยละ 19.34 จากนั้นไปจะมีค่าเบี่ยงเบนมาตรฐานเหลืออีก 2.3 ก็จะมีกรณีของปรากฏการณ์เกิดขึ้นอีกประมาณร้อยละ 11.50 การที่เป็นเช่นนี้จึงสามารถนำบุคคลที่จะถูกประเมินมาแบ่งเป็น 5 ประเภท ประเภทสูงสุดมีร้อยละ 10 (ทอนจาก 11.50 %) ประเภทที่สองร้อยละ 20 (ปรับจาก 19.34 %) ประเภทที่สามร้อยละ 40.00 (ปรับจาก 38.30 %) ประเภทที่สี่ร้อยละ 20 (เท่ากับประเภทที่สองโดยที่ถือค่าเบี่ยงเบนมาตรฐานเท่ากัน) และประเภทที่ห้า ร้อยละ 10 (เท่ากับประเภทที่หนึ่ง) ถ้าหากมีผู้ปฏิบัติงาน 50 คน ผู้ประเมินก็จะต้องจัดให้มีบุคคลในประเภทต่าง ๆ จากสูงสุดมาต่ำสุด เป็นจำนวน 5,10,20,10,5 ตามลำดับ

¹² Guion, *op. cit.*, pp. 100—109.

Thorndike and Hagen, *op. cit.*, pp. 354—355.

วิธีจัดลำดับที่เป็นวิธีประเมินนั้นต้องใช้เวลามาก แต่ก็ได้ผลดีในการที่บังคับให้ผู้ประเมินต้องมีการวิเคราะห์ที่ผู้ปฏิบัติงานที่ถูกประเมินอย่างแท้จริง ผู้ประเมินไม่สามารถจัดให้ผู้ถูกประเมินได้รับค่าคะแนนเท่า ๆ กัน เป็นส่วนใหญ่ได้ ยิ่งไปกว่านั้นยังหลีกเลี่ยงความบิดเบือนจากอคติและความเบ้ในการให้คะแนนอีกด้วย กล่าวคือ ถึงแม้ผู้ประเมินจะเป็น “คนใจดี” ในการให้คะแนน อย่างไรก็ตามก็จำเป็นต้องจัดให้ผู้ถูกประเมินคนใดคนหนึ่งเป็นที่สุดท้าย และถึงแม้ว่าจะเป็น “คนกระดุก” ก็ต้องจัดให้มีคนได้อันดับที่หนึ่ง อย่างไรก็ตาม ปัญหาที่ยังเกิดขึ้นในแง่ที่ว่า ลำดับที่แต่ละคนได้รับในแต่ละกลุ่มนั้นมีค่าเท่ากันหรือเหมือนกันระหว่างกลุ่มหรือไม่ ผู้ที่ได้ที่สามในแผนกที่มีผู้ปฏิบัติงานเพียง 3 คน ย่อมมีค่าต่างจากผู้ที่ได้ที่สามในแผนกที่มีผู้ปฏิบัติงาน 20 คน นอกจากนี้ช่วงห่างระหว่างลำดับที่แต่ละทีมได้มีความห่างเป็นปริมาณเท่า ๆ กัน ฉะนั้นในการประเมินด้วยวิธีจัดลำดับที่นี้ จึงได้มีการเปรียบเทียบระหว่างกลุ่มที่ถูกประเมินด้วยการเปรียบเทียบนี้ทำได้โดยนำบุคคลจำนวนหนึ่งในกลุ่มมาจัดให้เป็นบุคคลหลัก (key men) บุคคลหลักนั้นจะต้องเป็นผู้ที่ผู้ประเมินทุกคนของกลุ่มต่าง ๆ รู้จักเป็นอย่างดีด้วย ลำดับที่ของบุคคลต่าง ๆ ในกลุ่มจะต้องนำมาคำนวณเป็นค่ามาตรฐานปกติ (normalized standard scores) ค่าของบุคคลหลักนั้นก็จะได้เป็นค่าในการลากเส้นสหสัมพันธ์ถดถอย (regression line) เพื่อที่จะเทียบค่าของบุคคลทุก ๆ คนในกลุ่มกับคะแนนบรรทัดฐาน (master scale) ซึ่งกำหนดค่าคะแนนที่เทียบได้เหมือนกันหมดระหว่างบุคคลทุกคนในกลุ่มต่าง ๆ

วิธีจัดบัญชีตรวจสอบเป็นวิธีที่กำหนดให้มีรายการของลักษณะต่าง ๆ ที่ควรจะมีในพฤติกรรมการทำงานทั้งหมด การสร้างรายการที่ถือว่าดีที่สุด คือ การสร้างโดยใช้เทคนิคเหตุการณ์วิกฤติ (critical incident technique) เทคนิคนี้มีหลักว่าผู้สร้างรายการจะต้องวิเคราะห์เหตุการณ์ที่มีความสำคัญที่สุดต่อการปฏิบัติงานให้ได้สำเร็จ หรือทำให้การปฏิบัติงานต้องล้มเหลว เช่น บัญชีตรวจสอบพฤติกรรมการทำงานของพนักงานขายของบริษัท Minnesota Mining and Manufacturing Company กำหนดรายการไว้ดังนี้¹³

- ก. เทียวซุบซิบความลับของลูกค้า
- ข. ไม่มีความรู้เกี่ยวกับผลผลิตที่จะขาย
- ค. สัญญากับลูกค้ามากเกินไป
- ง. ให้ลูกค้าติดหนื้มมากเกินไปจนลูกค้าเองเกิดสงสัย
- จ. เมื่อลูกค้ามีเรื่องไม่พอใจ รีบติดตามผลให้เป็นที่พอใจ

¹³ Guion, *op. cit.*, p. 105.

การประเมินโดยวิธีตรวจสอบอาจทำได้หลายวิธี เช่น วิธีจัดบัญชีสรุป (summed ratings) วิธีจัดช่วงเสมอกัน (equal-appearing ratings) และวิธีบังคับตอบ (forced-choice rating) เนื่องจากวิธีบังคับตอบเป็นวิธีที่มีชื่อเสียงกว่าวิธีอื่น จึงจะอธิบายเฉพาะวิธีนี้เท่านั้นในที่นี่ วิธีนี้ถือหลักการเช่นเดียวกับวิธีจัดลำดับที่ในการพิจารณาเปรียบเทียบลักษณะนิสัยของพฤติกรรมของทุกๆ คนพร้อมๆ กัน ต่างกันแต่ว่า แทนที่จะพิจารณาลักษณะนิสัยทีละอย่างให้พิจารณาลักษณะนิสัยทีละชุดแทน (a set of attributes) โดยทั่วๆ ไป ผู้ประเมินจะทำรายการที่สะท้อนให้เห็นภาพผู้ปฏิบัติงานที่ดีในตำแหน่งหนึ่งๆ เป็นชุดๆ ละ 4 รายการ (tetrad) โดยให้รายการในชุดหนึ่งๆ มีค่าเท่าเทียมกันในสายตาของผู้อธิบายคุณลักษณะของพฤติกรรมการปฏิบัติงานทั่วๆ ไป แต่ในแต่ละชุดนี้จะมีอยู่ 2 รายการที่มีความสำคัญในการชี้ว่า ผู้ปฏิบัติงานที่ดีนั้นต้องมีลักษณะนิสัยอย่างไร เช่น ในบัญชีตรวจสอบซึ่งใช้ประเมินงานในตำแหน่งครูโรงเรียนการช่างของกองทัพอากาศสหรัฐฯ ได้กำหนด

- ก. อุดทนที่จะสอนผู้เรียนซ้ำ
- ข. บรรยายด้วยความมั่นใจ
- ค. ทำให้นักเรียนสนใจและกระตือรือร้น
- ง. อธิบายให้นักเรียนเข้าใจความมุ่งหมายก่อนสอนทุกอย่าง

แต่ละรายการในชุดจะเป็นคำอธิบายครูที่ดีในโรงเรียนทั้งสิ้น แต่รายการหนึ่งๆ จะมีค่าในการวัดว่าเป็นครูที่ดีในปริมาณต่างกัน โดยวิธีวิเคราะห์หองค์ประกอบ ซึ่งจะกล่าวรายละเอียดในบทต่อไป ทำให้ทราบว่า แต่ละรายการในชุดนี้มีค่าพรรณนัแยกคุณลักษณะ (discrimination index) เรียงลำดับดังต่อไปนี้ คือ +1.15, +.55, +1.39 และ +.79 ค่าพรรณนัทำให้สามารถให้ค่าแก่รายการแต่ละอันที่ผู้ประเมินเลือกกว่าคำอธิบายการปฏิบัติงานของผู้ถูกประเมิน เช่น อาจจะให้ค่า 2 แก่รายการข้อ ก. ในฐานะเป็นรายการที่มีค่าพรรณนัแยกลักษณะสูงสุด และให้ค่าศูนย์แก่รายการข้อ ง. เนื่องจากมีค่าพรรณนัต่ำที่สุด ส่วนข้อ ก. และ ข. ได้ 1 เท่าๆ กัน คะแนนรวมของรายการทุกชุดก็จะเป็นค่าประเมินของพฤติกรรมงานของผู้ถูกประเมินคนหนึ่งๆ การที่จัดรายการให้เป็นเช่นนี้มีเหตุผลที่ว่า ผู้ประเมินจะไม่เอนเอียงเพราะรายการทุก ๆ ข้อก็ไม่มีอะไรที่เลวร้ายเป็นพิเศษกว่ากัน ผู้ประเมินเองก็ไม่ทราบว่า รายการข้อใดจะเป็นประโยชน์ หรือเป็นผลเสียต่อผู้ถูกประเมินอย่างไร อีกประการหนึ่ง ค่าคะแนนที่ได้จะมีลักษณะกระจายใกล้เคียงกับเกณฑ์เฉลี่ยตามปกติ (normal distribution) อย่างไรก็ดี การจัดบัญชีตรวจสอบโดยวิธีบังคับตอบ ก็ยังมีปัญหาอยู่ 2 ประการ คือ ผู้ประเมินที่จำต้องใช้วิธีนี้ใน

การประเมินมักจะอ้างว่า มีความลำบากในการเลือกรายการ อีกประการหนึ่ง การเปรียบเทียบลักษณะนิสัยซึ่งต่าง ๆ กัน ยังไม่มีวิธีการมาตรฐานทำให้ค่าที่ประเมินได้ โดยวิธีนี้ขาดความหมายที่แท้จริงที่นำมาใช้อธิบายภาพผู้ปฏิบัติงานที่ถูกประเมิน

การประเมินมาตรฐานวัดความสำเร็จของงาน

ปัญหาสำคัญของเครื่องมือวัดในการบริหารงานบุคคลข้อหนึ่ง คือ การที่ต้องวินิจฉัยว่าจะวัดอะไรจึงจะได้ข้อเท็จจริงที่แสดงพฤติกรรมในการปฏิบัติงานสำเร็จ ถ้าได้ข้อเท็จจริงนั้นมา ก็ทำให้มีมาตรฐานวัดความสำเร็จของงานที่ดี เป็นทางให้มาเทียบค่ากับเครื่องมือวัด เพื่อสร้างเครื่องมือวัดที่ดีต่อไปได้ จากแนวคิดอันนี้ก็ย่อมหมายความว่า มาตรฐาน ๆ ควรเป็นปรนัยให้มาก และควรจะต้องสัมพันธ์สอดคล้องกับตัวแปรร่วมของข้อเท็จจริงเพื่อให้มีความเที่ยงตรง (validity) และความเชื่อมั่น (reliability) นอกจากนี้ มาตรฐานที่ดีก็ยังต้องมีคุณภาพในการแยกแยะว่า ผู้ที่ปฏิบัติงานสำเร็จด้วยดีนั้นดีเป็นระดับอย่างไร และมีความสะดวกและความประหยัดในการรวบรวมข้อมูล¹⁴ ความจริงหลักการประเมินว่า มาตรฐานแต่ละอันดีหรือไม่ก็มีลักษณะคล้าย ๆ กับหลักการประเมินคุณค่าเครื่องมือวัดนั่นเอง รายละเอียดในเรื่องความเชื่อมั่น หรือความเที่ยงตรงก็ได้มีอธิบายไว้โดยเฉพาะแล้วในบทที่ 4 และที่ 5 ตามลำดับ อย่างไรก็ตามในเรื่องมาตรฐานมีรายละเอียดบางอย่างซึ่งเป็นประเด็นเฉพาะ ดังจะกล่าวแยกไว้ในหัวข้อข้างล่าง

ความเที่ยงตรง อาศัยหลักแนวคิดในการสร้างมาตรฐาน ดังได้กล่าวไว้ในตอนต้นบท ผู้อ่านก็อาจเห็นได้ว่ามาตรฐาน จะมีความเที่ยงตรงได้ ก็ต่อเมื่อมีค่าสหสัมพันธ์กับมาตรฐานสูงสุดท้ายในอัตราสูง แต่ในทางปฏิบัติก็ได้กล่าวไว้แล้วว่า เป็นการยากที่จะค้นหามาตรฐาน ขึ้นสุดท้าย ฉะนั้นในการพิจารณาความเที่ยงตรงของมาตรฐาน ก็อาจทำได้โดยวิเคราะห์ว่ารายละเอียดที่จะนำมาสร้างมาตรฐาน เป็นประเด็นที่สำคัญต่อการปฏิบัติงานในตำแหน่งหน้าที่หนึ่ง ๆ หรือไม่ การวิเคราะห์นี้เป็นแบบความเที่ยงตรงในแง่แนวคิด ซึ่งหมายความว่า ผู้สร้างมาตรฐาน หลีกเลี่ยงการใช้ความรู้สึกส่วนตัวได้ลำบาก¹⁵ อย่างไรก็ดี หากผู้สร้างใช้ดุลพินิจวิเคราะห์เหตุผลเป็นอย่างดี ก็อาจหลีกเลี่ยงความมั่วคลำได้ โดยที่ข้อมูลสำหรับบางอย่างเห็นได้ชัดว่า หากผู้สร้างมาตรฐาน ใช้ความระมัดระวังก็อาจคัดเลือกเฉพาะที่ตรงประเด็นได้ เช่น หากมีการวิเคราะห์ความสำเร็จของงานในตำแหน่งเจ้าหน้าที่ประจำปี ความแม่นยำและความรวดเร็วในการยิงไปถูกเป้าหมายก็คือข้อมูล

¹⁴ Thorndike., *op. cit.*, pp. 124—125.

¹⁵ Cronbach, Lee J., and Meehl, Paul E., Construct Validity in Psychological Tests, in Jackson, Douglas N., and Messick, Samuel, eds, *Problems in Human Assessment*, pp. 60—62.

ที่มีความตรงประเด็นที่จะนำมาใช้สร้างมาตรฯ ที่เที่ยงตรงได้ และในกรณีที่ไม่สามารถ
ชั่งไปได้ชัดเจนว่า ข้อมูลนั้นตรงประเด็นกับความสำเร็จของงานหรือไม่ ก็อาจใช้หลัก
สถิติช่วยในการวิเคราะห์ได้¹⁶

ความเชื่อมั่น สัมประสิทธิ์ความเชื่อมั่นของมาตรฯ ควรมีค่าสูงกว่าศูนย์ซึ่งนี้หมายถึง
ความว่า ผู้สร้างมาตรฯ ควรได้มีการทดลองคำนวณค่าทางสถิติของข้อมูลต่าง ๆ ที่นำมา
ประกอบการสร้างมาตรฯ¹⁷ Thorndike ยืนยันว่าไม่มีความจำเป็นถึงขั้นสูงสุดที่จะต้อง
ให้มาตรฯ มีค่าสัมประสิทธิ์ความเชื่อมั่นสูง ต่อเมื่อกำหนดค่าต่ำมากก็ควรจะได้แก้ไข การแก้ไข
พิจารณาเป็น 2 นัย คือ คุณค่าความเชื่อมั่นในแง่เนื้อแท้ (intrinsic reliability) และคุณค่า
ความเชื่อมั่นในแง่อาการภายนอก (extrinsic reliability) หากค่าสัมประสิทธิ์ความเชื่อมั่น
ต่ำตามนัยแรกก็เป็นเพราะ เกิดจากความไม่คงเส้นคงวาส่วนบุคคลที่ปฏิบัติงานนั้น ๆ
แก้ไขได้โดยเพิ่มจำนวนตัวอย่างที่ประเมินให้มากขึ้น แต่ถ้าหากต่ำตามนัยที่สองก็เป็น
เพราะเกิดจากสภาพแวดล้อมที่เกี่ยวข้องกับการปฏิบัติงาน เช่น ดินฟ้าอากาศ เครื่องจักร
ลักษณะการปฏิบัติงานของผู้ร่วมงานของผู้ถูกประเมิน หรือการไม่จำกัดความหมายของ
งานที่จะประเมินให้แน่นอนตายตัว ในกรณีเช่นนี้ต้องมีการควบคุมสภาพแวดล้อมเหล่านี้
โดยอาศัยหลักวิธีการวิจัยอย่างแท้จริง รายละเอียดมีกล่าวไว้ในตำราของผู้วิจัยคนเดียว
กันแล้ว¹⁸

ความสามารถในการแยกแยะ (discrimination) หมายความว่ามาตรฯ ที่สร้างนั้นจะต้อง
มีคุณภาพสามารถแยกได้ว่า พฤติกรรมในลักษณะใดจึงทำให้ปฏิบัติงานได้ผลสำเร็จ
อย่างดีที่สุด และเป็นอย่างไรๆ ลงมาประการใด หากมาตรฯ ที่สร้างขึ้นนั้นไม่สามารถ
แยกได้ว่า บุคคลที่ปฏิบัติอย่างไรจะมีผลสำเร็จของงานต่างกันเป็นระดับอย่างไร ก็ย่อม
ทำให้ไม่สามารถสร้างเครื่องมือวัดที่จะคาดคะเนได้ว่า ควรจะคัดเลือกเฉพาะบุคคล
ประเภทใด จึงจะได้บุคคลที่มีคุณสมบัติตามที่ต้องการ

ความสะดวกและความประหยัด (practicality) การรวบรวมข้อมูลเพื่อการสร้างมาตรฯ
ก็มีลักษณะเช่นเดียวกับการเก็บข้อมูลเพื่อการวิจัยทั่วไป คือ เนื่องจากงานนี้เป็นงาน
ที่ต้องเสียค่าใช้จ่ายและเวลามาก ผู้สร้างมาตรฯ จึงต้องใช้ดุลพินิจให้รอบคอบในการ

¹⁶ Ghiselli, Edwin E., *Theory of Psychological Measurement*, pp. 354—357.

¹⁷ Thorndike, *op. cit.*, pp. 127—128.

¹⁸ สุคนธรังษี, สวัสดิ์, และจุลละนันทน์, สมศรี, *ระเบียบวิธีเบื้องต้นในการวิจัย*, หน้า 112—115.

กำหนดข้อจำกัดเพื่อการวิจัย (limitation) เพื่อให้สอดคล้องกับทุนและเวลาที่มียู่¹⁹ ยิ่งไปกว่านั้น ข้อมูลในการสร้างมาตรฯ ยังมีลักษณะพิเศษกว่าข้อมูลธรรมดาทั่วไปในการวิจัยในแง่ที่ว่า ข้อมูลส่วนใหญ่ที่จะใช้ในการสร้างมาตรฯ เป็นข้อมูลที่ต้องรบกวนเวลาและความสะดวกของผู้ปฏิบัติงานฝ่ายต่าง ๆ เพราะเป็นข้อมูลที่เป็นข้อเท็จจริงในการปฏิบัติงานจริง ๆ ฉะนั้น นักวิจัยงานบุคคลในด้านนี้ย่อมหลีกเลี่ยงไม่ได้ที่จะสร้างความไม่พอใจให้แก่บุคคลฝ่ายต่าง ๆ ถ้าไม่ได้ทำให้บุคคลที่จะต้องเสียสละเวลาและความสะดวกในการให้ข้อเท็จจริงเข้าใจเสียก่อนล่วงหน้าถึงประโยชน์สำคัญของมาตรฯ และถึงแม้ผู้ปฏิบัติงานฝ่ายต่าง ๆ จะเห็นความสำคัญในเรื่องนี้ แต่การที่ผู้วิจัยต้องใช้ระยะเวลา นานกว่าจะสร้างมาตรฯ สำเร็จ ก็อาจทำให้ผู้ให้ความร่วมมือขาดความอดทนเสียก่อนก็ได้ นักวิจัยที่มีหน้าที่สร้างมาตรฯ จึงควรเทียบปัญหาเรื่องนี้กับระยะเวลาที่ใช้ปฏิบัติงาน และเมื่อเห็นว่าถึงขั้นที่ควรจะให้ผู้ปฏิบัติงานที่เกี่ยวข้องได้ทราบหรือตระหนักประโยชน์ของการรวบรวมข้อมูลในตอนใด ก็ควรดำเนินการให้สอดคล้องกับการขอความร่วมมือของทุก ๆ ฝ่าย

¹⁹ Thorndike, *op. cit.*, pp. 131—132.
Guion, *op. cit.*, pp. 120—121.

การวิเคราะห์องค์ประกอบ

เครื่องมือวัดที่สมบูรณ์แบบไม่เพียงแต่มีความเชื่อมั่นและค่าความเที่ยงตรง แต่ยังจะต้องกำหนดให้แน่ชัดลงไปว่า ลักษณะที่วัดนั้นมี “ธาตุ” ตามที่ต้องการ และลักษณะนั้น ๆ ทำให้คาดคะเนได้ว่า ผู้ถูกวัดมีความสามารถ และพฤติกรรมที่จะทำให้ปฏิบัติงานสำเร็จได้ผลตามเป้าหมายของหน่วยงาน ถึงแม้ว่าทดลองเครื่องมือวัดที่สร้างขึ้นกับมาตรวัดความสำเร็จของงานก็ยังไม่สามารถทำให้ได้เครื่องมือวัดที่สมบูรณ์แบบตามที่ต้องการ เพราะนอกจากจะประสบปัญหาในการวิเคราะห์มาตรฐานวัดความสำเร็จที่เที่ยงตรงแม่นยำแล้ว ผู้วัดยังจะต้องใช้เครื่องมือวัดหลายร้อยชุดเพื่อให้ทราบลักษณะที่แท้จริงทั้งหมดของบุคคลหนึ่ง การใช้เครื่องมือวัดเป็นร้อย ๆ ชุดในการวัดบุคคลหนึ่ง เป็นสิ่งที่ต้องเสียเวลามากมายจนไม่สะดวกในทางปฏิบัติ ดังนั้นจึงได้มีการนำหลักคณิตศาสตร์มาประยุกต์ในการสร้างเครื่องมือวัด หลักนี้เรียกว่า การวิเคราะห์องค์ประกอบ (factor analysis) การวิเคราะห์เครื่องมือวัดตามหลักนี้จะทำให้ทราบได้ว่า เครื่องมือวัดแต่ละชุดมี “ธาตุ” อะไรบ้าง เป็นผลดีเกิดขึ้น 2 ประการ คือ ประการแรก อาจใช้เครื่องมือวัดชุดใดชุดหนึ่งทดสอบบุคคล เพื่อให้ทราบระดับความสามารถหรือลักษณะนิสัยที่จำเป็นต่อการปฏิบัติงาน ในการนี้ไม่จำเป็นต้องใช้เครื่องมือวัดเป็นจำนวนมาก ๆ และอีกประการหนึ่งสิ่งที่วิเคราะห์ได้จากเครื่องมือวัด ซึ่งเรียกว่าองค์ประกอบ ทำให้แนวคิด (concept) ใหม่เกิดขึ้นในการอธิบายลักษณะและคุณสมบัติต่าง ๆ ของมนุษย์

ในการวัดบุคคล การวิเคราะห์องค์ประกอบใช้ได้ทั้งในขั้นตอนการวิเคราะห์เครื่องมือวัด และในขั้นการวิเคราะห์มาตรฐานวัดความสำเร็จของงาน เมื่อใช้กับการวิเคราะห์มาตรฐานวัดความสำเร็จของงาน ก็จะทำให้ความสามารถและลักษณะนิสัยที่จำเป็นต่อผู้มีในตัวผู้ปฏิบัติงานมีองค์ประกอบอะไรบ้าง และเมื่อเทียบกับงานระดับต่าง ๆ หรือ

ต่างชนิดกันแล้ว มุ่งเป้าประกอบที่เหมือนกันหรือต่างกันอย่างไร ส่วนในการใช้กับการวิเคราะห์เครื่องมือวัดหรือตัวแปรการวัดผล ที่ทำให้ทราบว่าเครื่องมือวัดหนึ่งมีองค์ประกอบที่ตรงกับมาตรฐานความสำเร็จของงานเพียงใด เมื่อต้องการจะทราบว่านาย ก. ควรปฏิบัติงานในตำแหน่งหนึ่ง แต่ไม่ควรปฏิบัติงานอีกตำแหน่งหนึ่ง ผู้วัดบุคคลจะต้องทราบว่า งานทั้งสองตำแหน่งนั้นต้องมีบุคคลที่มีคุณสมบัติความสามารถและลักษณะนิสัยซึ่งแยกเป็นส่วนต่าง ๆ ไม่ซ้ำกันอย่างไรบ้าง และในขณะเดียวกัน ตัวนาย ก. มีคุณสมบัติความสามารถและลักษณะนิสัยซึ่งแยกออกได้เป็นส่วนอย่างไร หากไม่มีการวิเคราะห์องค์ประกอบคุณสมบัติ ความสามารถ และลักษณะนิสัยที่จะต้องมีในตำแหน่งงานทั้งสอง และที่มีในตัวนาย ก. ก็ย่อมไม่มีทางทราบได้ว่า เหตุใด นาย ก. จึงควรปฏิบัติงานในตำแหน่งแรก แต่ไม่ควรปฏิบัติงานในตำแหน่งที่สอง

หลักการวิเคราะห์องค์ประกอบ อาจนำไปใช้ในการวิจัยทางพฤติกรรมศาสตร์ทั่วไป (behavioral science) เช่น ความเปลี่ยนแปลงเรื่องความรู้สึกของกลุ่มบุคคล รูปแบบของทัศนคติของบุคคล อันเป็นผลจากการเปลี่ยนแปลงของวงจรธุรกิจ เป็นต้น ในที่นี้จะได้อธิบายหลักขั้นมูลฐานเพื่อสร้างความเข้าใจทั่วไปในการใช้วิธีการนี้ ในการวัดผลบุคคลเท่านั้น

แนววิเคราะห์ทางคณิตศาสตร์

ข้อมูลเชิงปริมาณที่รวบรวมได้จากการวิจัยในปัญหาประเภทและเรื่องเดียวกันจะมีลักษณะพิเศษต่าง ๆ กันไป เพราะล้วนแต่เป็นเพียงตัวแทนส่วนหนึ่งของชุดค่าความเป็นจริงในจักรวาล (defined universal set) ลักษณะพิเศษอันหนึ่งของข้อมูลแต่ละชุดนี้คือ ความแปรปรวน (variance) ซึ่งได้กล่าวถึงไว้บ้างแล้วในบทที่ 3

ความแปรปรวนแต่ละค่าสามารถแยกออกได้เป็นสองส่วน คือ ความแปรปรวนที่แท้จริง (true or systematic variance) และความแปรปรวนอันเนื่องมาจากความคลาดเคลื่อน (error variance) ความแปรปรวนที่แท้จริงเป็นผลเกิดจากผลกระทบของตัวแปร ส่วนความแปรปรวนอันเนื่องมาจากความคลาดเคลื่อนเกิดจากความเพ้อผู่

ถึงแม้ข้อมูลในชุดหนึ่ง ๆ จะมีความแปรปรวน ข้อมูลทั้งหลายในแต่ละชุดก็จะมีลักษณะที่คล้ายกันอยู่จนอาจทำให้รวมกันเป็นกลุ่มย่อย ๆ ภายในชุดหนึ่ง ๆ ได้ ความแปรปรวนอันเกิดจากความไม่เหมือนกันระหว่างกลุ่มย่อย ๆ ภายในชุดนี้ ทำให้เกิดความแปรปรวนที่แท้จริง ความแตกต่างระหว่างกลุ่มย่อยอาจเกิดจากการควบคุมทดลองตัวแปร หรือเกิดจากตัวแปรอื่น ๆ ที่ไม่ได้ควบคุมทดลองก็ได้ทั้งสองอย่าง

ในแต่ละกลุ่มย่อย ซึ่งได้กล่าวแล้วว่า มีลักษณะคล้ายกัน จึงได้นำมาเทียบค่าความแปรปรวนที่แท้จริง ข้อมูลภายในกลุ่มก็ไม่จำเป็นที่จะต้องมีความเสมอเหมือนกัน ความแปรปรวนภายในกลุ่มแต่ละกลุ่มเอง เมื่อนำมาเฉลี่ยก็จะทำให้ทราบค่าความแปรปรวนอันบังเอิญ

ความเป็นจริงเรื่ององค์ประกอบของความแปรปรวนนี้ อาจพิสูจน์ได้จากตัวอย่างง่าย ๆ ทางคณิตศาสตร์ดังต่อไปนี้

สมมติว่าในหน่วยงานแห่งหนึ่ง ทดลองใช้วิธีคัดเลือกสองวิธีในการบรรจุบุคคลเข้าปฏิบัติงานในหน้าที่หนึ่ง ซึ่งต้องมีผู้ปฏิบัติงานที่เหมือนกันเป็นจำนวน 10 คน วิธีแรกคัดเลือกโดยพิจารณาระดับคะแนนจากประกาศนียบัตรวิชาชีพที่ได้รับ คัดเลือกไว้เป็นจำนวน 5 คน วิธีที่สอง คัดเลือกโดยให้ทดสอบข้อเขียนและสัมภาษณ์ คัดเลือกไว้เป็นจำนวนเท่า ๆ กัน เมื่อปฏิบัติงานไประยะเวลาหนึ่งแล้ว หัวหน้างานได้ประเมินค่าความสามารถในการปฏิบัติงานของบุคคลทั้ง 10 คน ออกมาดังนี้

กลุ่มที่คัดเลือกโดยวิธีแรก : 5, 4, 3, 2, 1

กลุ่มที่คัดเลือกโดยวิธีที่สอง : 8, 7, 6, 5, 4

ในการหาค่าความแปรปรวนของข้อมูลทั้งหมด ทำได้โดยใช้สูตร

$$\sigma^2 = \frac{\Sigma x^2}{n}$$

หาค่าความแปรปรวนของข้อมูลทั้งหมด (σ^2) :

X	x	x ²
5	.5	.25
4	-.5	.25
3	-1.5	2.25
2	-2.5	6.25
1	-3.5	12.25
8	3.5	12.25
7	2.5	6.25
6	1.5	2.25
5	.5	.25
4	-.5	.25

\bar{X} 4.5

Σx^2 42.50

$$\therefore \sigma_{\infty}^2 = \frac{42.50}{10} = 4.25$$

หาค่าความแปรปรวนที่แท้จริง (σ_{∞}^2):

$$\text{ค่าเฉลี่ยของกลุ่มแรก} = \frac{5+4+3+2+1}{5} = 3$$

$$\text{ค่าเฉลี่ยของกลุ่มที่สอง} = \frac{8+7+6+5+4}{5} = 6$$

X	x	x ²
3	-1.5	2.25
6	1.5	2.25

$$\bar{X} \quad 4.5$$

$$\Sigma x^2 \quad \quad \quad 4.50$$

$$\therefore \sigma_{\infty}^2 = \frac{4.50}{2} = 2.25$$

หาค่าความแปรปรวนอันบังเอิญ (σ_e^2) :

กลุ่มแรก (A ₁)	x	x ²	กลุ่มหลัง (A ₂)	x	x ²
5	2	4	8	2	4
4	1	1	7	1	1
3	0	0	6	0	0
2	-1	1	5	-1	1
1	2	4	4	-2	4

$$\bar{X} \quad 3$$

$$6$$

$$\Sigma x^2$$

$$10$$

$$10$$

$$\sigma^2 A_1 = \frac{10}{5} = 2$$

$$\sigma^2 A_2 = \frac{10}{5} = 2$$

$$\therefore \sigma_e^2 = \frac{2+2}{2} = 2$$

จะเห็นได้ว่า เมื่อนำค่าความแปรปรวนที่แท้จริง คือ 2.25 รวมกับค่าความแปรปรวนอันบังเอิญคือ 2 จะได้เท่ากับ 4.25 ซึ่งเป็นค่าความแปรปรวนของข้อมูลทั้งหมด จึงเขียนเป็นสูตรได้ว่า

$$\sigma_t^2 = \sigma_{\infty}^2 + \sigma_e^2$$

ค่าความแปรปรวนที่แท้จริงอันเกิดจากอิทธิพลของตัวแปร ยังอาจแยกได้เป็นอีกสองส่วน คือ ค่าความแปรปรวนขององค์ประกอบร่วมกัน (common factor variance) และค่าความแปรปรวนเฉพาะ (specific variance)

ในบทที่ 3 ในหัวข้อเรื่องความสัมพันธ์ระหว่างค่าความแปรปรวนและสัมประสิทธิ์สหสัมพันธ์ ได้อธิบายไว้แล้วว่า เมื่อนำค่าสหสัมพันธ์มายกกำลังสองก็จะได้ผลลัพธ์เป็นค่าความแปรปรวนขององค์ประกอบร่วมกัน ค่าความแปรปรวนที่คำนวณได้โดยวิธีนี้ชื่อเรียกอีกอย่างหนึ่งว่า สัมประสิทธิ์เพื่อการวินิจฉัย (coefficient of determination)

ในทางตรงข้ามกับความแปรปรวนขององค์ประกอบร่วมกัน ข้อมูลกลุ่มใดกลุ่มหนึ่งภายในชุด อาจมีความแปรปรวนที่ไม่มีความสัมพันธ์ร่วมกับข้อมูลกลุ่มอื่น ๆ เลย ความแปรปรวนนี้เรียกว่า ความแปรปรวนเฉพาะ

การที่ความแปรปรวนที่แท้จริงแยกได้เป็นสองส่วนนี้ก็โดยที่ข้อมูลกลุ่มต่าง ๆ ในชุดอาจมีลักษณะความแปรปรวนร่วมกัน เพราะเป็นข้อมูลที่ได้มาจากเครื่องมือวัดที่มีลักษณะคล้าย ๆ กัน แต่เครื่องมือวัดข้อมูลที่คล้าย ๆ กัน ก็ยังอาจแยกได้เป็น 2 ชนิดคือ ชนิดแรกเป็นเครื่องมือวัดคนละอัน แต่คล้ายกันในเรื่องเนื้อหา และชนิดที่สองเป็นเครื่องมือวัดอันเดียวกันแต่จัดรูปแบบในลักษณะต่าง ๆ กัน ผลของการใช้เครื่องมือวัดชนิดแรกจะทำให้เกิดความแปรปรวนขององค์ประกอบร่วมกัน แต่ผลของการใช้เครื่องมือวัดชนิดที่สองทำให้เกิดความแปรปรวนเฉพาะ เช่น เครื่องมือวัดหลาย ๆ อัน ที่ต่างก็วัดทักษะทั่วไปในการปฏิบัติงาน จะทำให้ข้อมูลซึ่งรวบรวมได้จากเครื่องมือวัดเหล่านี้เกิดความแปรปรวนร่วมกัน ในขณะที่หากเครื่องมือวัดอันหนึ่งอันใดในบรรดาเครื่องมือวัดเหล่านี้ สามารถวัดทักษะพิเศษด้วย ข้อมูลที่รวบรวมได้ก็就会有ความแปรปรวนเฉพาะเกิดขึ้นได้ ความแปรปรวนที่เกิดในส่วนหลังนี้เป็นค่าความสัมพัทธ์ระหว่างข้อมูลที่รวบรวมโดยเครื่องมือวัดอันหนึ่งอันใดเหล่านี้ กับข้อมูลที่อาจรวบรวมได้โดยเครื่องมือวัดอันเดียวกันนี้ เมื่อเปลี่ยนรูปร่างให้มีลักษณะต่าง ๆ ออกไป การแยกความแปรปรวนที่แท้จริงออกเป็น 2 ส่วน ได้มีกล่าวไว้เป็นสูตรแล้วในบทที่ 5 ในหัวข้อเรื่องความเที่ยงตรงกับความแปรปรวน คือ

$$\sigma_o^2 = \sigma_{xy}^2 + \sigma_{sp}^2$$

โดย σ_{xy}^2 = ความแปรปรวนขององค์ประกอบร่วมกัน

σ_{sp}^2 = ความแปรปรวนเฉพาะ

สูตรเบื้องต้นที่ใช้ในทฤษฎีการวิเคราะห์องค์ประกอบ ในบทที่ 3 ในหัวข้อเรื่องตารางสัมพัทธ์แบบเอฟและแบบอาร์ ได้อธิบายความหมายขององค์ประกอบของเครื่องมือวัด (factor loading) และค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบ (communality) ไว้แล้ว ในที่นี้จะได้วิเคราะห์ความหมายดังกล่าวในเชิงคณิตศาสตร์ เพื่อให้เห็นชัดเจนยิ่งขึ้น

ถ้าหากค่าของข้อมูลทุก ๆ กลุ่มในชุดมีตั้งแต่กลุ่ม a ถึงกลุ่ม n และได้คำนวณเป็นค่ามาตรฐานแล้ว (standard score) ค่าของข้อมูลทั้งหมดก็อาจเขียนเป็นสูตรได้ดังนี้

$$Z_t = AZ_a + BZ_b + CZ_c + \dots + NZ_n + SZ_s + Z_e$$

โดยให้ Z_t = ค่ามาตรฐานของข้อมูลทั้งหมด

Z = ค่ามาตรฐานของข้อมูลแต่ละกลุ่ม

SZ_s = ค่ามาตรฐานขององค์ประกอบเฉพาะ

Z_e = ค่ามาตรฐานในส่วนที่ผิดพลาด

ถ้าหากสัมประสิทธิ์ของข้อมูลแต่ละกลุ่มไม่เป็นศูนย์ ค่าของข้อมูลทุก ๆ กลุ่มก็จะมีส่วนทำให้เกิดความแปรปรวนของข้อมูลทั้งหมด จึงจะทำให้เขียนสูตรที่ใช้ในทฤษฎีการวิเคราะห์องค์ประกอบได้ดังนี้

$$\sigma_t^2 = \sigma_a^2 + \sigma_b^2 + \sigma_c^2 + \dots + \sigma_n^2 + \sigma_s^2 + \sigma_e^2$$

ถ้าเอาความแปรปรวนทั้งหมดหารตลอด

$$\frac{\sigma_t^2}{\sigma_t^2} = \frac{\sigma_a^2}{\sigma_t^2} + \frac{\sigma_b^2}{\sigma_t^2} + \frac{\sigma_c^2}{\sigma_t^2} + \dots + \frac{\sigma_n^2}{\sigma_t^2} + \frac{\sigma_s^2}{\sigma_t^2} + \frac{\sigma_e^2}{\sigma_t^2}$$

$$\therefore 1.00 = \frac{\sigma_a^2}{\sigma_t^2} + \frac{\sigma_b^2}{\sigma_t^2} + \frac{\sigma_c^2}{\sigma_t^2} + \dots + \frac{\sigma_n^2}{\sigma_t^2} + \frac{\sigma_s^2}{\sigma_t^2} + \frac{\sigma_e^2}{\sigma_t^2}$$

สมมุติให้ $A_x^2 + B_x^2 + C_x^2 + \dots + N_x^2$ เป็นผลบวกของอัตราส่วนระหว่างความแปรปรวนของข้อมูลแต่ละกลุ่มกับความแปรปรวนของข้อมูลทั้งหมดตั้งแต่กลุ่ม a ถึงกลุ่ม n ให้ S_x^2 และ E_x^2 เป็นอัตราส่วนระหว่างค่าความแปรปรวนเฉพาะ และความแปรปรวนอันเนื่องมาจากความผิดพลาด กับความแปรปรวนของข้อมูลทั้งหมดตามลำดับ โดยที่ข้อมูลเหล่านี้รวบรวมด้วยเครื่องมือวัด x ก็อาจเขียนเป็นสูตรใหม่ได้ดังนี้

$$A_x^2 + B_x^2 + C_x^2 + \dots + N_x^2 + S_x^2 + E_x^2 = 1.00$$

$$\underbrace{\hspace{10em}}_{H_x^2}$$

ค่าอัตราส่วนระหว่างความแปรปรวนตั้งแต่กลุ่ม a ถึงกลุ่ม n คือ ค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบดังกล่าวได้เคยอธิบายมาแล้ว เมื่อพิจารณาจากคำอธิบายซึ่งได้เคยให้ไว้ว่า ค่าองค์ประกอบทุก ๆ ตัวของเครื่องมือวัดหนึ่ง ๆ เมื่อยกกำลังสองแล้วนำมารวมกันจะได้ผลลัพธ์เป็นค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบของเครื่องมือวัดอันหนึ่ง ๆ ก็จะเห็นว่า ค่า $A_x^2 + B_x^2 + C_x^2 + \dots + N_x^2$ คือค่าองค์ประกอบแต่ละตัวที่ยกกำลังสอง

$$\therefore R_h = A_x^2 + B_x^2 + C_x^2 + \dots + N_x^2 + S_x^2 \dots \text{(บทที่ 4)}$$

$$\therefore H^2 = R_h - S_x^2$$

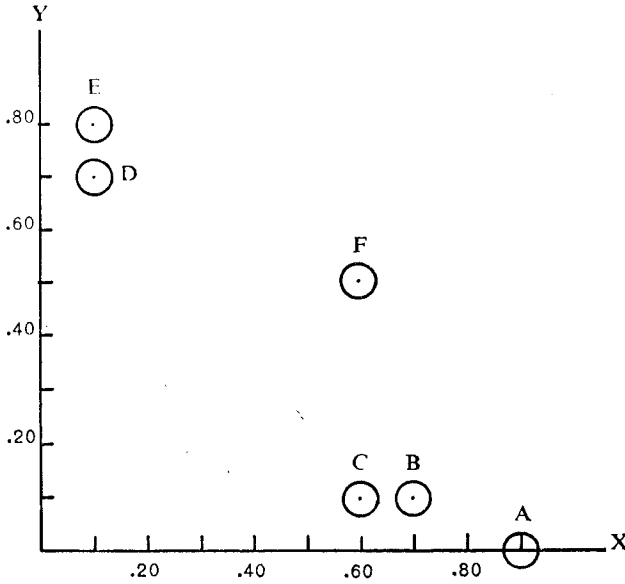
จากสูตรข้างบนนี้ ทำให้สามารถเห็นได้ว่า ค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบหรือค่าความแปรปรวนขององค์ประกอบร่วมกัน เกี่ยวโยงกับสัมประสิทธิ์ความเชื่อมั่นอย่างใกล้ชิด สัมประสิทธิ์แรกจะไม่มีทางสูงกว่าสัมประสิทธิ์หลัง สัมประสิทธิ์ความเชื่อมั่นจะต้องสูงกว่า หรืออย่างน้อยต้องเท่ากับค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบ

ในขณะเดียวกันค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบ อาจเขียนเป็นสูตรอื่นได้อีกคือ

$$H_x^2 = 1.00 - (S_x^2 + E_x^2)$$

ฉะนั้น ถ้าค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบสูง ค่าความแปรปรวนเฉพาะและค่าความแปรปรวนอันบังเอิญก็ย่อมต่ำ เช่น ในตารางที่ 4 ของบทที่ 3 ค่าสหสัมพันธ์ระหว่างองค์ประกอบของเครื่องมือวัดการเข้าใจคนมีปริมาณ 0.81 ค่าความแปรปรวนเฉพาะ และค่าความแปรปรวนอันบังเอิญ เมื่อรวมกันแล้วก็จะไม่มีโอกาสสูงกว่า 0.19 ความเป็นจริงเช่นนี้ ทำให้สามารถสร้างองค์ประกอบได้จากผลของการที่ข้อมูลมีค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบสูง

ลักษณะขององค์ประกอบในเชิงสถิติ การพิจารณาค่าองค์ประกอบเพื่อให้เห็นเด่นชัดขึ้น อาจอาศัยวิธีการทางสถิติประกอบ เช่น ในแผนภาพที่ 8



แผนภาพที่ 8 ลักษณะองค์ประกอบแสดงด้วยกราฟ

ในแผนภาพที่ 8 แสดงองค์ประกอบโดยใช้ข้อมูลของตารางที่ 4 ในบทที่ 3 องค์ประกอบ 2 ตัว คือ ความสามารถเข้าใจนามธรรม และความสามารถหาเหตุผลถูกกำหนดให้เป็นอักษร X และ Y ตามลำดับ (ในความเป็นจริง ผู้วิเคราะห์องค์ประกอบย่อมไม่ทราบว่าจะองค์ประกอบนั้นชื่อว่าอะไร จนกว่าจะได้วิเคราะห์ค่าองค์ประกอบและค่าสหสัมพันธ์ระหว่างองค์ประกอบเสียก่อน) แกน X และ Y มีชื่อเรียกว่า แกนอ้างอิง (reference axe) ค่าแต่ละค่าก็จะบันทึกบนแผ่นกราฟตามความสัมพันธ์ของค่าองค์ประกอบที่หามาได้ เช่น ค่าองค์ประกอบของเครื่องมือวัดอันแรก คือ เครื่องมือวัดการเข้าใจคนได้ค่า 0.90 กบ 0.00 ก็จะยาวไปตามแกน X เท่ากับ 0.90 และขึ้นไปบนแกน Y เท่ากับ 0.00 ผลที่ได้ก็มีวงกลมและอักษร A กำกับว่าเป็นเครื่องมือวัด

อันแรก ค่าองค์ประกอบของเครื่องมือวัดอื่นก็ได้กำหนดให้เป็นอักษร B, C และ D ตามลำดับ บนแผนกราฟจะเห็นได้ว่าเครื่องมือวัดแต่ละอันมีองค์ประกอบตัวหนึ่งตัวใดเป็นพิเศษ กล่าวคือ เครื่องมือวัด A, B และ C มีองค์ประกอบร่วมกัน และเครื่องมือวัด D ก็มีองค์ประกอบร่วมกับ E ถ้าหากเครื่องมือวัดไม่สามารถแยกแยะองค์ประกอบได้ ค่าที่คำนวณก็จะไม่เอียงไปข้างใดข้างหนึ่งอย่างเห็นได้ชัด ในแผนภาพนี้ได้สมมุติค่า F ซึ่งเป็นค่าของ 0.50 และ 0.60 ไว้เป็นตัวอย่าง

หลักคณิตศาสตร์ในการวิเคราะห์องค์ประกอบ เมื่อแทนค่าของข้อมูลชุดหนึ่ง ๆ ด้วย vector สัมประสิทธิ์สหสัมพันธ์ระหว่างข้อมูลสองชุด จะมีค่าเท่ากับผลคูณระหว่าง vectors ทั้งสอง และค่า cosine ของมุม ϕ ที่เกิดระหว่าง vectors ทั้งสอง จะเขียนเป็นสูตรได้ดังนี้

$$R_{12} = V_1 V_2 \cos \phi_{12}$$

$$R_{12} = \text{สัมประสิทธิ์สหสัมพันธ์ระหว่างข้อมูลสองชุด}$$

$$V_1 = \text{vector ที่แทนค่าข้อมูลชุดที่หนึ่ง}$$

$$V_2 = \text{vector ที่แทนค่าข้อมูลชุดที่สอง}$$

อันจะพิสูจน์ได้ดังต่อไปนี้

ให้ข้อมูลที่เป็ผลจากการวัดโดยเครื่องมือทดสอบสองชุด แทนค่าด้วย vectors ที่มีความยาวเป็นหนึ่งหน่วยเต็ม หากมุม ϕ มีค่าเท่ากับ 90° , 45° , 0° และ 120° ผลของการคำนวณสัมประสิทธิ์สหสัมพันธ์จะได้ค่าดังต่อไปนี้

$$\cos 90 = 0.0 \quad r = (1)(1)(0) = 0$$

$$\cos 45 = 0.707 \quad r = (1)(1)(.707) = .707$$

$$\cos 0 = 1.0 \quad r = (1)(1)(1) = 1$$

$$\cos 120 = -.50 \quad r = (1)(1)(-.50) = -.50$$

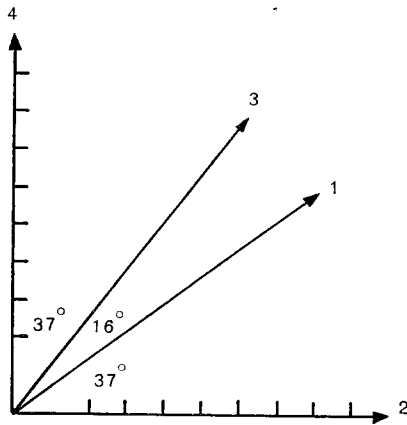
การที่ทราบความสัมพันธ์ระหว่าง vectors กับสัมประสิทธิ์สหสัมพันธ์เช่นนี้ ทำให้สามารถคำนวณมิติของความสัมพันธ์ระหว่าง vectors จากค่าสหสัมพันธ์ระหว่างเครื่องมือวัดหรือข้อมูลชุดต่าง ๆ ค่านี้แสดงในตารางสัมพันธ์แบบอาร์ซึ่งได้เคยกล่าวไว้แล้วในบทที่ 3 มิติของ vectors จึงมีชื่อเรียกเฉพาะอีกว่า ศักดิ์ของตารางสัมพันธ์ (rank of a matrix) ศักดิ์ของตารางสัมพันธ์ทำให้ทราบว่า ข้อมูลในชุดทั้งหลายเกี่ยวข้องกันเป็นจำนวนมิติเท่าใด หรือกล่าวอีกอย่างหนึ่งได้ว่า ในการที่ข้อมูลชุดต่าง ๆ สัมพันธ์

กันนั้น มุ่งเป้าประกอบ (ซึ่งได้แก่มิติแต่ละด้าน) ร่วมกันเป็นปริมาณเท่าใด การพิจารณาตีของความสัมพันธ์ระหว่าง vectors ทำได้ดังต่อไปนี้

ตัวอย่าง 1 ถ้าหากข้อมูล 4 ชุด หรือผลของการวัดโดยเครื่องมือวัด 4 ชุด มีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันดังต่อไปนี้

	1	2	3	4
1				
2	.80			
3	.96	.60		
4	.60	.00	.80	

ถึงแม้จะมีข้อมูลเป็นจำนวน 4 ชุด ซึ่งแปลว่าจะต้องมี vectors 4 เส้น ค่าสหสัมพันธ์ข้างบนนี้ก็อาจแสดงได้เพียง 2 มิติ โดยมีความสอดคล้องกับตัวเลขทั้งหมด กล่าวคือ ค่าสหสัมพันธ์ของข้อมูลชุดที่สองและที่สี่ แสดงว่า vectors ที่แทนค่าข้อมูลทั้งสองชุดจะต้องตั้งฉากกัน (เพราะว่า $\cos 90^\circ = 0$) ส่วนข้อมูลชุดที่หนึ่งก็มีค่าเท่ากับ .80 ของชุดที่สอง และ .60 ของชุดที่สี่ ข้อมูลชุดที่สามเท่ากับ .60 ของชุดที่สอง และ .80 ของชุดที่สี่ เมื่อลาก vectors ที่แทนข้อมูลชุดที่หนึ่งและที่สามให้สัมพันธ์กับชุดที่สองและชุดที่สี่แล้ว ค่ามุมระหว่าง vectors ต่าง ๆ จะมีผลรวมเท่ากับค่ามุมระหว่าง vectors ที่แทนข้อมูลชุดที่สองและที่สี่ ดังแสดงไว้ในแผนภาพที่ 9



แผนภาพที่ 9 vectors หลายเส้นที่มีเพียงสองมิติ

ค่าของ vector แต่ละเส้นมีความสัมพันธ์กับ vectors ของข้อมูลชุดที่สองและที่สี่ ดังนี้

$$V_1 = .80 V_2 + .60 V_4$$

$$V_2 = 1.00 V_2 + .00 V_4$$

$$V_3 = .60 V_2 + .80 V_4$$

$$V_4 = .00 V_2 + 1.00 V_4$$

การวิเคราะห์ข้อมูลทั้ง 4 ชุด ปรากฏผลว่ามีองค์ประกอบ 2 ตัว และอาจเขียนค่าความสัมพันธ์ระหว่างข้อมูลแต่ละชุดกับองค์ประกอบแต่ละตัว โดยใช้ค่า vectors ข้างบน ดังแสดงไว้ในตารางที่ 19

ตารางที่ 19

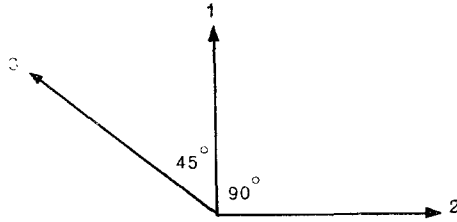
ตารางแบบ เอฟ สองมิติ

องค์ประกอบ ข้อมูลชุดที่	ตัวที่หนึ่ง	ตัวที่สอง
1	.80	.60
2	1.00	.00
3	.60	.80
4	.00	1.00

ตัวอย่างที่ 2 ถ้าหากข้อมูล 3 ชุด มีสัมประสิทธิ์สหสัมพันธ์ระหว่างกัน ดังต่อไปนี้

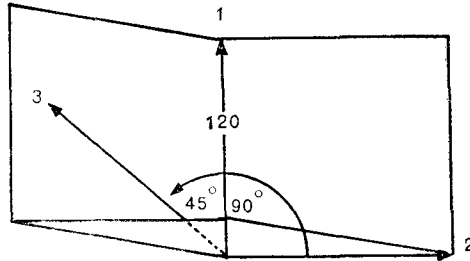
	1	2	3
1		.000	.707
2	.000		-.500
3	.707	-.500	

เมื่อลาก vectors ที่แทนข้อมูลชุดที่หนึ่งและที่สองให้ตั้งฉากกันแล้ว และลาก vector ที่แทนข้อมูลชุดที่ 3 ให้ทำมุมกับชุดที่หนึ่ง 45° ผลรวมระหว่างมุมทั้งสองจะไม่เท่ากับ 120° ซึ่งจะต้องเป็นค่ามุมระหว่าง vectors ที่แทนข้อมูลชุดที่ 2 และที่ 3 ดังแสดงไว้ในแผนภาพที่ 10



แผนภาพที่ 10 vectors ที่แสดงมิติที่ถูกตัด

การที่ vectors ซึ่งแทนข้อมูลชุดที่ 2 และชุดที่ 3 ไม่สัมพันธ์กันตามสัมประสิทธิ์สหสัมพันธ์ที่ระบุไว้ หมายความว่า การสร้าง vectors ทั้งหมดให้มีเพียงสองมิติไม่ถูกต้อง ฉะนั้นจึงต้องสร้างเป็นสามมิติ ดังแสดงในแผนภาพที่ 11

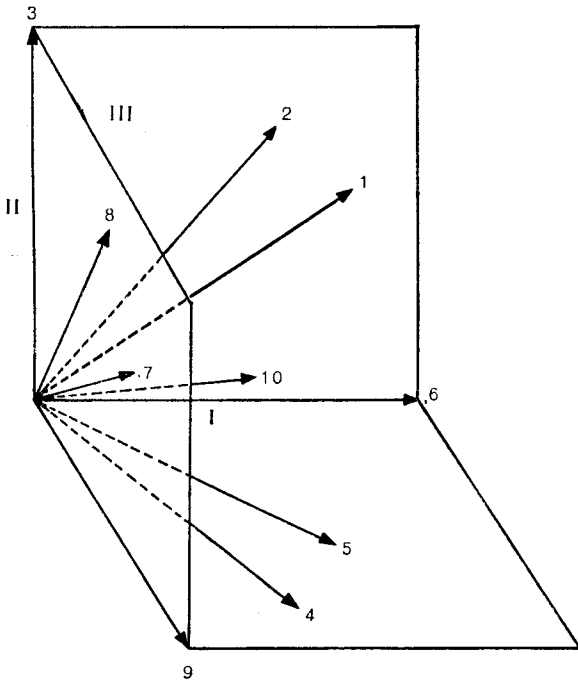


แผนภาพที่ 11 vectors สามเส้นที่มีสามมิติ

ตัวอย่างที่ 3 ถ้าหากข้อมูล 10 ชุด มีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันดังต่อไปนี้

	1	2	3	4	5	6	7	8	9	10
1		.96	.60	.48	.64	.80	.36	.48	.00	.70
2	.96		.80	.36	.48	.60	.48	.64	.00	.70
3	.60	.80		.00	.00	.00	.60	.80	.00	.50
4	.48	.36	.00		.96	.60	.64	.48	.80	.86
5	.64	.48	.00	.96		.80	.48	.36	.60	.82
6	.80	.60	.00	.60	.80		.00	.00	.00	.50
7	.36	.48	.60	.64	.48	.00		.96	.80	.86
8	.48	.64	.80	.48	.36	.00	.96		.60	.82
9	.00	.00	.00	.80	.60	.00	.80	.60		.70
10	.70	.70	.50	.86	.82	.50	.86	.82	.70	

จากตารางสัมพันธ์ หน้า 158 จะเห็นว่า ค่าของข้อมูลชุดที่สามมีค่าสัมพันธ์กับชุดที่สี่ ห้า หก และเก้า เป็นศูนย์ และในขณะเดียวกัน ค่าสหสัมพันธ์ระหว่างชุดที่หกและที่เก้าก็เป็นศูนย์ ฉะนั้น vectors ที่แทนค่าข้อมูลชุดที่สาม หก และเก้า จึงตั้งฉากซึ่งกันและกัน แต่ vectors ที่แทนค่าข้อมูลชุดที่สี่และที่ห้าจะเพียงแต่ตั้งฉาก vector ของข้อมูลชุดที่สามเส้นเดียว vectors ของข้อมูลชุดอื่นก็อาจหาได้โดยเทียบค่าที่สัมพันธ์กับ vectors หลัก กล่าวคือ vectors ที่หนึ่งและที่สองอยู่ในมิติระหว่าง vectors ที่สามและที่หก vectors ที่สี่และที่ห้าอยู่ในมิติระหว่าง vectors ที่หกและที่เก้า vectors ที่เจ็ดและที่แปดอยู่ในมิติระหว่าง vectors ที่สามและที่เก้า vector ที่สิบอยู่ในช่วงที่เกิดขึ้นระหว่างมิติทั้งสาม ดังแสดงไว้ในแผนภาพที่ 12



แผนภาพที่ 12 vectors หลายเส้นที่มีสามมิติ

เมื่อสร้าง vectors ได้สอดคล้องกับค่าในตารางสัมพันธ์อาร์ ได้แล้วเช่นนั้น ก็แปลว่า vectors สำหรับข้อมูลชุดที่สาม ที่หก และที่เก้า เป็น vectors หลัก จึงเรียกได้ว่าเป็นองค์ประกอบที่สอง ที่หนึ่ง และที่สาม ตามลำดับ ต่อไปก็อาจคำนวณค่าองค์ประกอบจากการที่ข้อมูลแต่ละชุดสัมพันธ์กับ vectors หลัก ดังแสดงไว้ในตารางที่ 20

ตารางที่ 20

ตารางแบบ เอฟสามมิติ (น้ำหนักตัวประกอบ)

องค์ประกอบ ข้อมูลชุด	ตัวที่หนึ่ง	ตัวที่สอง	ตัวที่สาม
1	.80	.60	.00
2	.60	.80	.00
3	.00	1.00	.00
4	.60	.00	.80
5	.80	.00	.60
6	1.00	.00	.00
7	.00	.60	.80
8	.00	.80	.60
9	.00	.00	1.00
10	.50	.50	.70

จากตัวอย่างทั้งสามจะเห็นได้ว่า ค่าสหสัมพันธ์ระหว่าง vectors ต่าง ๆ อาจแสดงเป็นมิติได้ต่าง ๆ กัน ทั้งนี้ขึ้นกับการหักมุม (angular separation) ระหว่าง vectors หากมีมิติเดียว vectors ก็จะทำกันบนเส้นตรงเส้นเดียว เรียกว่าศักดิ์ของตารางสัมพันธ์มีปริมาณเท่ากับสาม และก็มีองค์ประกอบสามตัว ศักดิ์ของตารางสัมพันธ์นี้อาจมีมากเท่าใดก็ได้ (ถึงแม้ว่าเราจะวาดภาพไม่ได้) ศักดิ์ของตารางสัมพันธ์เมื่อมีปริมาณถึง n ก็แปลว่า มีองค์ประกอบ n ตัว

ในหัวข้อเรื่องตารางสัมพันธ์แบบเอฟและแบบอาร์ ในบทที่ 3 ได้กล่าวถึงค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบ (communality) ซึ่งมีตัวย่อว่า h^2 เมื่อนำเอาคำอธิบายนั้นมาพิจารณากับข้อความที่กล่าวในหัวข้ออื่น ก็อาจสร้างความเข้าใจให้ชัดเจนยิ่งขึ้น ค่าองค์ประกอบแต่ละตัวทั้งในตารางที่ 19 และที่ 20 เมื่อยกกำลังสองแล้วรวมกันเฉพาะในบรรทัดเดียวกัน แต่ละบรรทัดจะได้ค่า h^2 ซึ่งมีปริมาณเท่ากับ 1.00 เหมือนกันหมด จากสูตรเบื้องต้นที่ใช้ในทฤษฎีการวิเคราะห์องค์ประกอบ ซึ่งกล่าวมาก่อนแล้วในบทนี้ ค่า h^2 จะมีโอกาสเป็น 1.00 ได้ก็ต่อเมื่อค่าความแปรปรวนเฉพาะและความแปรปรวนอันบังเอิญต่างก็มีค่าเท่ากับศูนย์

ในการสร้าง vectors หลัก ในตัวอย่างข้างต้น อาจสังเกตได้ว่าไม่มีสถานที่กำหนดตายตัวในการลาก vectors ค่าในตารางสัมพันธแบบอาร์ที่กำหนดให้ ไม่จําจกัว่าจะต้องลาก vectors หลักในทีใด ผู้สร้าง vectors อาจเลือกที่ต่าง ๆ เพื่อลาก vector ในลักษณะอื่น ๆ ได้อีกมากมาย ฉะนั้นจึงได้มีการคำนวณค่าว่า ถ้าหากลาก vectors ในลักษณะอื่น ๆ จะทำให้ได้ค่าองค์ประกอบต่าง ๆ กันไปอย่างไร โดยการหมุน vectors ไปรอบ ๆ ตัว²

ถ้าหากหมุน vectors ตามเข็มนาฬิกา ค่าองค์ประกอบจะเปลี่ยนไปดังนี้

$$A'_{j1} = A_{j1} \cos \phi - A_{j2} \sin \phi$$

$$A'_{j2} = A_{j1} \sin \phi + A_{j2} \cos \phi$$

A_{j1} = ค่าองค์ประกอบตัวที่หนึ่งของข้อมูลชุด j เมื่อหมุน vectors แล้ว

A_{j1} = ค่าองค์ประกอบตัวที่หนึ่งของข้อมูลชุด j ก่อนหมุน vectors

A_{j2} = ค่าองค์ประกอบตัวที่สองของข้อมูลชุด j ก่อนหมุน vectors

ϕ = ค่าของมุมที่ vectors ที่หมุนไปแล้วทำมุมกับ vectors ในลักษณะเดิม

A'_{j2} = ค่าองค์ประกอบตัวที่สองของข้อมูลชุด j เมื่อหมุน vectors แล้ว

แต่ถ้าหากหมุนทวนเข็มนาฬิกา ค่าองค์ประกอบจะเปลี่ยนใหม่ดังนี้

$$A'_{j1} = A_{j1} \cos \phi + A_{j2} \sin \phi$$

$$A'_{j2} = -A_{j1} \sin \phi + A_{j2} \cos \phi$$

จากสูตรการหมุน vectors เมื่อนำมาใช้กับข้อมูลในตัวอย่างที่ 1 โดยสมมติให้ vectors ต่าง ๆ หมุนตามนาฬิกาไปเป็นปริมาณ 50° จะทำให้ค่าองค์ประกอบที่เปลี่ยนไป คือ

ค่าองค์ประกอบของข้อมูลชุดแรกในตัวอย่างที่ 1

$$\sin 50^\circ = .766 \quad \text{และ} \quad \cos 50^\circ = .643$$

$$A'_{j1} = (.80) (.643) - (.60) (.766) = .0548$$

$$A'_{j2} = (.80) (.766) + (.60) (.643) = .9986$$

ค่าองค์ประกอบของข้อมูลชุดอื่น ๆ ในตัวอย่างเดียวกัน ก็อาจหาได้ในทำนองเดียวกัน และได้แสดงไว้ในตารางที่ 21

² Fruchter, Benjamin, *Introduction to Factor Analysis*, p. 39.

ตารางที่ 21

ค่าองค์ประกอบเมือหมุน vectors 50°

ข้อมูลชุด องค์ประกอบ	ตัวที่หนึ่ง	ตัวที่สอง	h^2
1	.054	.998	.9989
2	.643	.766	1.00
3	-.227	.974	1.00
4	-.766	.643	1.00

ค่าองค์ประกอบต่างๆ จะเปลี่ยนไป เมื่อหมุน vectors ข้อมูลชุดที่หนึ่งและที่สาม จะมีค่าองค์ประกอบตัวที่สองสูงกว่าข้อมูลชุดอื่น ๆ ทั้ง ๆ ที่ไม่ได้เป็นเช่นนั้นมาก่อน การเปลี่ยนแปลงเช่นนี้จะเกิดขึ้นทุก ๆ ระยะที่มีการเปลี่ยนการหักมุมระหว่าง vectors แต่ค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบจะไม่เปลี่ยนแปลง ในที่นี้จะเห็นว่ามีแต่ข้อมูลชุดที่หนึ่งเท่านั้นที่ค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบมีปริมาณไม่ถึง 1.00 แต่ก็อาจถือว่ามีค่าตัดเทียม เพราะต่างกันตรงที่มีได้บดเศษส่วน

รูปแบบในการวิเคราะห์องค์ประกอบ

แนวคิดในการวางรูปแบบเพื่อการวิเคราะห์องค์ประกอบ (factor analytic research designs) เป็นผลงานของนักวิชาการในการวัดผลหลายท่าน ในขั้นแรก Spearman เป็นผู้ริเริ่มเสนอทฤษฎีองค์ประกอบ 2 ตัว (two-factor theory) เมื่อปี ค.ศ. 1940 ต่อมา ในปี ค.ศ. 1946 Cattell ได้ขยายความอธิบายว่า การวิเคราะห์องค์ประกอบอาจวางรูปแบบได้ถึง 6 อย่าง คือ R, Q, O, P, S และ T techniques

ความหมายของทฤษฎีของ Spearman ได้มีอธิบายไว้แล้วในบทที่ 3 ในที่นี้จะได้พิสูจน์ข้อเท็จจริงของทฤษฎีนี้ตามหลักคณิตศาสตร์ ก่อนที่จะกล่าวถึงทฤษฎีของ Cattell ซึ่งเป็นที่ยอมรับกันในปัจจุบันต่อไป

รูปแบบของทฤษฎีองค์ประกอบสองตัว หากตารางสัมพัทธ์แบบอาร์แสดงว่า อัตราส่วนระหว่างสัมประสิทธิ์สหสัมพันธ์แต่ละคู่ในสดมภ์ (column) เดียวกันมีปริมาณเท่ากับ อัตราส่วนระหว่างสัมประสิทธิ์สหสัมพันธ์แต่ละคู่ในสดมภ์ถัดไปที่อยู่ใบบรรทัดเดียวกัน ทั้งนี้โดยยกเว้นไม่นับค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อมูลแต่ละชุดกับตัวของมันเอง

ค่าสหสัมพันธ์ระหว่างข้อมูลแต่ละกลุ่มซึ่งมีสี่ชุดในลักษณะเช่นนี้ จะแสดงว่าข้อมูลทั้งกลุ่มมีความแปรปรวนมาจากแหล่งเดียวกัน หรือกล่าวอีกนัยหนึ่งว่ามีองค์ประกอบร่วมกัน หลักันเรียกว่าหลักมาตรฐานวัดความเป็นสัดส่วน (criterion of proportionality)³ ซึ่งเขียนเป็นสูตรได้ว่า

$$\frac{R_{13}}{R_{23}} = \frac{R_{14}}{R_{24}} = \frac{R_{15}}{R_{25}} = \dots = \frac{R_{1n}}{R_{2n}}$$

สูตรข้างบนระบุเฉพาะอัตราส่วนของสัมประสิทธิ์สหสัมพันธ์แต่ละคู่ในบรรทัดที่หนึ่งและที่สอง อัตราส่วนระหว่างบรรทัดที่สองและที่สาม และคู่บรรทัดอื่น ๆ ก็อาจพิจารณาได้ในทำนองเดียวกัน

จากสูตรนี้ก็อาจพิจารณาหาความแตกต่างระหว่างข้อมูลแต่ละกลุ่มซึ่งมี 4 ชุด (tetrad difference) ได้ คือ

$$R_{13} R_{24} - R_{14} R_{23} = 0$$

$$R_{13} R_{25} - R_{15} R_{23} = 0$$

และอาจจับกลุ่มเช่นนี้ไปได้เรื่อย ๆ หากมีข้อมูล n ชุด ก็จะจับกลุ่มได้ $\frac{1}{2}(n-1)(n-2)(n-3)$ ชุด เช่น ถ้ามีข้อมูล 4 ชุด จะมี 3 กลุ่ม แต่ถ้ามี 10 ชุดจะมี 252 กลุ่ม

Spearman และคณะได้ทดลองใช้เครื่องมือวัดหลาย ๆ อัน ทดสอบความสามารถของบุคคล และค้นพบว่าเครื่องมือวัดเหล่านี้หลายอันเข้าหลักมาตรฐานวัดความเป็นสัดส่วน จึงเรียกองค์ประกอบร่วมระหว่างเครื่องมือวัดต่าง ๆ ที่วัดความสามารถของบุคคลว่า g (general-intellectual factor) คำอธิบายความหมายขององค์ประกอบตัวนี้มีกล่าวไว้แล้วในบทที่ 3 อย่างไรก็ตามที่ทฤษฎีนี้ได้นำมากล่าวก็เพื่อให้ผู้อ่านได้เข้าใจเนื้อหาในเรื่องการวิเคราะห์องค์ประกอบตั้งแต่ได้เริ่มเกิดขึ้นในตอนแรก ในปัจจุบันเป็นที่เชื่อกันว่าเครื่องมือวัดต่าง ๆ จะมีองค์ประกอบหลายตัว ดังได้อธิบายไว้ในหัวข้อแนววิเคราะห์ทางคณิตศาสตร์ เมื่อตอนต้นบทแล้ว

รูปแบบในการวิเคราะห์ตัวแปร 3 ตัว ในปี ค.ศ. 1952 Cattell ได้พิสูจน์ว่า การวิเคราะห์องค์ประกอบของข้อมูลหรือเครื่องมือวัดใด ๆ อาจทำได้ โดยหาความแปรปรวนร่วม (covariation) ระหว่างตัวแปร 3 ตัว⁴ คือ

³ Guilford, *op. cit.*, p. 473.

⁴ Cattell, Raymond B., The Three Basic Factor Analytic Research Designs-Their Interrelations and Derivatives, in Jackson, Douglas N., and Messick, Sammel, *Problems in Human Assessment*, pp. 300-304.

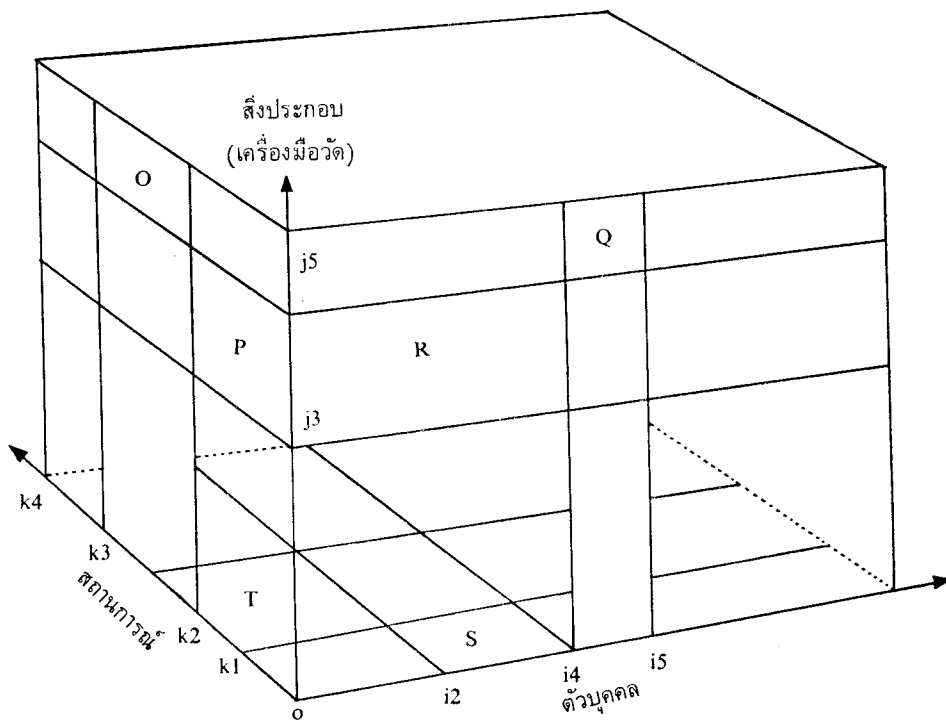
1. สภาพการณ์ (circumstances) หมายถึงสถานการณ์ต่างๆ ที่ทำให้เกิดการเร่งเร้า (stimulus situation) คือสิ่งต่างๆ ที่เป็นปัจจัยทำให้เกิดลักษณะประกอบ (attribute) ของบุคคลต่างๆ สิ่งเหล่านี้มีหลายอย่าง เช่น เวลา สถานที่ เรียกได้อีกอย่างหนึ่งว่าโอกาส (occasions)

2. ตัวบุคคล (persons) หมายถึงตัวตนของสิ่งมีชีวิต (organism) ซึ่งเป็นแหล่งที่มีสิ่งประกอบต่างๆ

3. สิ่งประกอบ (attributes) คือ สิ่งซึ่งนิยามได้ด้วยการสังเกต หรือการวัดส่วนต่างๆ ของการเร่งเร้า เรียกได้อีกอย่างหนึ่งว่า การสนองตอบ (response) เฉพาะในเรื่องนี้หมายถึงเครื่องมือวัดหนึ่ง ๆ ซึ่งสะท้อนลักษณะการสนองตอบของบุคคลในด้านต่างๆ เป็นค่าตัวเลข

ในแผนภาพที่ 13 ได้แสดงแผนภูมิของความแปรปรวนร่วมระหว่างตัวแปรทั้งสามตัว จะเห็นได้ว่า ในการวิเคราะห์องค์ประกอบ หากดำเนินการตามแนวของทฤษฎีองค์ประกอบสองตัว ก็ทำได้โดยหาค่าสหสัมพันธ์ระหว่างสิ่งประกอบสองสิ่งในกลุ่มบุคคลหลายๆ คน ในแผนภาพจะเห็นว่า เส้น j_3 และ j_5 แทนค่าสิ่งประกอบสองสิ่งขนานกันไปตามความยาวตั้งแต่บุคคลที่ศูนย์ (io) ถึงบุคคลที่ n (in) ในขณะที่เดียวกันยังอาจวิเคราะห์องค์ประกอบโดยเทียบค่าสหสัมพันธ์ ระหว่างสภาพการณ์ที่ศูนย์ (ko) ถึงที่ n (kn) ของบุคคลเพียงคนเดียวก็ได้ การวางรูปแบบในการวิเคราะห์องค์ประกอบจึงอาจทำได้เป็น 6 แนว คือ

1. แนว อาร์ (R technique) หาองค์ประกอบจากการวิเคราะห์สิ่งประกอบตั้งแต่สองสิ่งในกลุ่มบุคคลหลายๆ คน ในสถานการณ์หนึ่งสถานการณ์ใด องค์ประกอบที่วิเคราะห์ได้ จะเป็นลักษณะนิสัยร่วม (common traits) ที่จะแสดงความแตกต่างกันระหว่างบุคคลหนึ่ง ๆ ในการวิเคราะห์อาจใช้เครื่องมือวัดตั้งแต่สองอันขึ้นไปคำนวณหาค่าความแปรปรวนร่วมระหว่างกลุ่มของบุคคล ซึ่งได้สุ่มตัวอย่างแล้วในสภาพการณ์หนึ่ง ๆ ผลของความแปรปรวนร่วมระหว่างเครื่องมือวัดถือว่าแสดงสิ่งประกอบของตัวบุคคล



แผนภาพที่ 13 แผนภูมิความแปรปรวนร่วมของ Cattell

2. แนวคิว (Q technique) เป็นการวิเคราะห์ตารางสัมพันธ์สลับ (transpose)⁵ ของแนวอาร์ หาดองค์ประกอบโดยวิเคราะห์บุคคลตั้งแต่สองคนในกลุ่มสิ่งประกอบหลายๆ สิ่ง ในสถานการณ์หนึ่งสถานการณ์ใด เส้น i_4 และ i_5 แทนค่าตัวบุคคล 2 คน ขนานไปตามความยาวของเส้น j ซึ่งแทนค่าสิ่งประกอบ การวัดตัวบุคคล 2 คนนี้ อาจวัดตั้งแต่สิ่งประกอบ j_0 ไปจนถึง j_n ในสถานการณ์ k เพียงอันเดียว องค์ประกอบที่วิเคราะห์จะแสดงชนิดของสิ่งประกอบของตัวบุคคล การวิเคราะห์เช่นนี้เรียกว่าเป็นการปฏิบัติโดยเน้นตัวบุคคล (ipsative)

3. แนว พี (P technique) ถือว่า ค่าสหสัมพันธ์ระหว่างสิ่งประกอบตั้งแต่สองสิ่งของบุคคลคนเดียวกันในสถานการณ์ต่างๆ กัน แสดงลักษณะนิสัยเฉพาะอย่างหนึ่งๆ เส้น j_3 และ j_5 ขนานกันไปตามเส้น k ซึ่งแทนค่าสถานการณ์ตั้งแต่ k_0 ถึง k_n ในการ

⁵ ตารางสลับ คือ ตารางสัมพันธ์ที่จัดสลับสดมภ์ (column) กับบรรทัด (row) เสียใหม่ เมื่ออ่านค่าในแนวของตารางสลับจะได้ค่าตรงกันกับอ่านจากสดมภ์ของตารางสัมพันธ์.

วิเคราะห์อาจใช้เครื่องมือวัดหลาย ๆ อัน หาข้อมูลจากการวัดบุคคลเพียงคนเดียวซ้ำ ๆ กันหลาย ๆ ครั้ง

4. แนวโอ (O technique) เป็นการวิเคราะห์ตารางสัมพันธ์สลับของแนวพี หองค์ประกอบจากการวิเคราะห์สิ่งประกอบทั้งหมดของบุคคลเพียงคนเดียว ในสถานการณ์ตั้งแต่สองอย่างขึ้นไป เส้น k_2 และ k_4 แทนค่าสถานการณ์สองอย่าง หากคู่ขนานกันไป ตามเส้น j ตั้งแต่ j_0 ไปจนถึง j_n ในการวิเคราะห์ใช้เครื่องมือวัดหลาย ๆ อันหาข้อมูลโดยการวัดบุคคลเพียงคนเดียวในโอกาสต่าง ๆ องค์ประกอบที่ได้แสดงอิทธิพลของสถานการณ์ต่าง ๆ ที่มีต่อลักษณะนิสัยของบุคคล

5. แนวเอส (S technique) หองค์ประกอบจากการวิเคราะห์สิ่งประกอบสิ่งเดียวระหว่างบุคคลตั้งแต่ 2 คนขึ้นไป ในสถานการณ์ต่าง ๆ กัน องค์ประกอบที่วิเคราะห์ได้แสดงอิทธิพลของสถานการณ์ต่าง ๆ ที่มีต่อสิ่งประกอบหนึ่ง ๆ ระหว่างบุคคล เส้น i_2 และ i_4 แทนค่าตัวบุคคล 2 คน ขนานไปตามความยาวของเส้น k ซึ่งมีค่าตั้งแต่ k_0 ถึง k_n ในการวิเคราะห์ใช้ค่าสหสัมพันธ์ระหว่างบุคคลตั้งแต่ 2 คน ในเรื่องปฏิบัติการที่บุคคลทั้ง 2 คนนี้มีการเร่งเร้าซึ่งได้แก่เครื่องมือวัดที่สร้างขึ้น ค่าองค์ประกอบที่คำนวณได้ แสดงปริมาณที่บุคคลแต่ละคนสามารถรับบทบาททางสังคม (social roles) เช่น การศึกษาหาความรู้ และการปฏิบัติตนในโอกาสต่าง ๆ ในสังคม

6. แนวที (T technique) เป็นการวิเคราะห์ตารางสัมพันธ์สลับของแนวเอส หองค์ประกอบจากการวิเคราะห์สิ่งประกอบสิ่งเดียวในกลุ่มบุคคลหลาย ๆ คน ในสถานการณ์ตั้งแต่ 2 อย่างขึ้นไป ในการวิเคราะห์ใช้เครื่องมือวัดเพียงอันเดียว แล้ววัดบุคคลหลาย ๆ คนในโอกาสที่เปรียบเทียบกันได้ เส้น k_1 และ k_3 แทนค่าสถานการณ์ 2 อย่างที่ลากขนานไปกับเส้น i ซึ่งแทนค่าตัวบุคคล และมีค่าตั้งแต่ i_0 ถึง i_n แนวทีมีลักษณะคล้ายแนวโอ ในการที่ใช้ค่าสหสัมพันธ์ที่เกิดขึ้นระหว่างสถานการณ์ตั้งแต่ 2 อย่าง มาพิจารณาหองค์ประกอบ แต่ต่างกับแนวโอ ตรงที่มีการใช้เครื่องมือวัดอันเดิมกับบุคคลหลาย ๆ คน เช่นเดียวกับแนวอาร์ เมื่อพิจารณาแนวการวิเคราะห์หองค์ประกอบตามแนวทีกับวิธีการหาค่าความเชื่อมั่นในบทที่ 4 จะเห็นว่า รูปแบบของวิธีนี้เหมือนกับการหาสัมประสิทธิ์ของความเชื่อมั่นด้วยวิธีทดลองเครื่องมือวัดซ้ำ (test-retest reliability) ในกรณีที่เครื่องมือวัดที่ใช้ทดลองแสดงลักษณะนิสัยที่เปลี่ยนแปลงได้ (dynamic traits) การวิเคราะห์หองค์ประกอบด้วยวิธีนี้ย่อมทำให้ทราบได้ว่า โอกาสใดในลักษณะใดบ้างจะมีอิทธิพลต่อสิ่งประกอบของตัวบุคคล เครื่องมือวัดที่สร้างด้วยวิธีนี้ย่อมแสดงบรรยากาศ

ทางสังคม เช่น ความแตกต่างของความเห็นของประชาชนในเรื่องปัญหาต่าง ๆ ของประเทศในโอกาสต่าง ๆ กัน และระยะเวลาที่แสดงความเปลี่ยนแปลงของวงจรกิจ

วิธีการแปลความหมายขององค์ประกอบ

จากข้อความตั้งแต่ต้นมาจนถึงบัดนี้ ผู้อ่านอาจจะทราบแต่เพียงแนวคิดในการวิเคราะห์องค์ประกอบ แต่ก็ยังไม่เข้าใจว่าในการวิเคราะห์องค์ประกอบนั้นได้องค์ประกอบมาอย่างไร ตามธรรมดานักวิเคราะห์องค์ประกอบจะเริ่มต้นงานด้วยการสร้างเครื่องมือวัด เพื่อหาข้อมูลจากบุคคลต่าง ๆ โดยคัดเลือกบุคคลเหล่านั้นให้สอดคล้องกับรูปแบบในการวิจัย ดังได้กล่าวไว้ในหัวข้อเรื่องรูปแบบในการวิเคราะห์ตัวแปรสามตัวก่อนหน้านี้แล้ว เมื่อได้ข้อมูลจากเครื่องมือวัดชุดต่าง ๆ แล้ว ก็นำมาคำนวณสัมประสิทธิ์สหสัมพันธ์ แล้วลงหลักฐานผลการคำนวณไว้ในตารางสัมพันธ์แบบอาร์

ในตารางสัมพันธ์ จะมีช่องว่างเป็นช่วงทแยงหลัก (principal diagonal) ซึ่งเป็นช่องสำหรับค่าสหสัมพันธ์ระหว่างเครื่องมือวัดต่าง ๆ แต่ละอัน กับตัวของมันเอง และตามปกติจะปล่อยให้ว่างไว้ ในช่วงทแยงหลักมักจะเอาค่าหนึ่งค่าใดใน 3 ประเภทใส่ไว้ คือ ค่าความแปรปรวนขององค์ประกอบร่วมกัน หรือค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบ (h^2) สัมประสิทธิ์ความเชื่อมั่น และค่าความแปรปรวนของข้อมูลทั้งชุดแต่ละชุด ทั้งนี้ขึ้นอยู่กับวิธีที่จะใช้คำนวณค่าองค์ประกอบแต่ละตัว

วิธีคำนวณค่าองค์ประกอบนั้นมีหลายวิธี วิธีที่รู้จักกันแพร่หลายที่สุด คือ วิธีระบบจุดศูนย์กลาง (centroid method) ของ Thurstone วิธีคำนวณแกนต่าง ๆ โดยใช้ค่าเฉลี่ยของสัมประสิทธิ์สหสัมพันธ์ เพราะถ้าวัดค่าเฉลี่ยนี้จะแสดงจุดใจกลางเส้น vectors ที่ใช้แทนเครื่องมือวัด แต่ในที่นี้จะอธิบายเฉพาะวิธีเบื้องต้น คือ วิธีช่วงทแยง (diagonal method) เพราะเป็นวิธีที่เข้าใจได้โดยง่ายที่สุด และอาจใช้เป็นพื้นฐานสำหรับผู้ต้องการศึกษาวิธีอื่น ๆ ต่อไป อย่างไรก็ตาม การใช้วิธีนี้มีข้อแม้ว่า จะต้องทราบค่าสหสัมพันธ์ร่วมระหว่างองค์ประกอบที่ถูกต้อง จึงจะคำนวณค่าองค์ประกอบได้ แต่ตามธรรมดาแล้วย่อมเป็นการยากที่จะทราบค่าความแปรปรวนขององค์ประกอบร่วมกันล่วงหน้า

วิธีช่วงทแยง ในที่นี้จะอธิบายโดยใช้ตัวอย่างของ Fruchter⁶ สมมุติว่าได้ค่าสหสัมพันธ์ระหว่างเครื่องมือวัด และค่าความแปรปรวนขององค์ประกอบร่วมกันตามตัวเลขในตารางแบบอาร์ข้างล่างนี้

⁶ Fruchter, *op. cit.*, pp. 52—59.

เครื่องมือวัด	1	2	3	4	5
1	(.52)	.48	.36	.40	.58
2	.48	(.64)	.00	.16	.72
3	.36	.00	(.81)	.63	.09
4	.40	.16	.63	(.53)	.25
5	.58	.72	.09	.25	(.82)

จากค่าข้างบนนี้ จะต้องคำนวณให้ได้ค่าเป็นตารางแบบ เอฟ ดังแสดงไว้ในตารางที่ 22

ตารางที่ 22

ตารางแบบเอฟ แสดงด้วยตัวอักษร

เครื่องมือวัด	องค์ประกอบ					h^2
	1	2	3	4	5	
1	A_{11}	0	0	0	0	h_1^2
2	A_{21}	A_{22}	0	0	0	h_2^2
3	A_{31}	A_{32}	A_{33}	0	0	h_3^2
4	A_{41}	A_{42}	A_{43}	A_{44}	0	h_4^2
5	A_{51}	A_{52}	A_{53}	A_{54}	A_{55}	h_5^2

ในการคำนวณค่าองค์ประกอบ จะถือว่า องค์ประกอบตัวที่หนึ่งเป็นเส้นที่ทับเส้น vector ที่แทนเครื่องมือวัดอันที่หนึ่ง ส่วนองค์ประกอบตัวอื่นๆ ทุกตัวนั้น เป็นเส้นที่ตั้งฉากกับ vector ที่หนึ่ง ฉะนั้น ความสัมพันธ์ระหว่างเครื่องมือวัดอันที่หนึ่งกับองค์ประกอบทุกตัว ยกเว้นตัวที่หนึ่งจึงเท่ากับศูนย์ (เพราะ $\cos 90^\circ = 0$) ในช่องต่างๆ ทุกช่องที่อยู่ทางขวามือของช่วงทแยง จะมีค่าเท่ากับศูนย์หมด

ในการคำนวณค่าองค์ประกอบตัวแรกของเครื่องมือวัดอันที่หนึ่ง :

$$\therefore h_1^2 = A_{11}^2 + A_{12}^2 + A_{13}^2 + A_{14}^2 + A_{15}^2$$

แต่ $A_{12}, A_{13}, A_{14}, A_{15}$ ต่างมีค่าเป็นศูนย์

$$\therefore A_{11} = \sqrt{h_1^2}$$

แทนค่า

$$A_{11} = \sqrt{.52} = .7211$$

ในการคำนวณค่าองค์ประกอบตัวแรกของเครื่องมือวัดอันอื่น ๆ จากสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดนั้น ๆ แต่ละอัน กับเครื่องมือวัดอันที่หนึ่ง กล่าวคือ ค่าสหสัมพันธ์ระหว่างเครื่องมือวัด หรือข้อมูลแต่ละคู่ มีค่าเท่ากับผลรวมของผลคูณของเครื่องมือวัด หรือข้อมูลแต่ละชุดกับองค์ประกอบแต่ละตัว คือ

$$\therefore R_{12} = A_{11} A_{21} + A_{12} A_{22} + A_{13} A_{23} + A_{14} A_{24} + A_{15} A_{25}$$

แต่ $A_{12} A_{22}$, $A_{13} A_{23}$, $A_{14} A_{24}$, $A_{15} A_{25}$ มีค่าเป็นศูนย์

$$\therefore A_{21} = \frac{R_{12}}{A_{11}}$$

แทนค่า

$$A_{21} = \frac{.48}{.7211} = .6656$$

โดยวิธีเดียวกันจะได้ค่า

$$A_{51} = \frac{R_{13}}{A_{11}} = .4992$$

$$A_{41} = \frac{R_{14}}{A_{11}} = .5547$$

$$A_{51} = \frac{R_{15}}{A_{11}} = .8043$$

ในการคำนวณค่าองค์ประกอบตัวที่สอง จะต้องถือว่าองค์ประกอบตัวที่สองอยู่บนเส้นที่ตั้งฉากกับองค์ประกอบตัวแรก ค่าองค์ประกอบทุกตัวของเครื่องมือวัดอันที่สอง ยกเว้นแต่องค์ประกอบตัวแรกและตัวที่สอง จึงมีค่าเท่ากับศูนย์ (ตามค่า $\cos 90^\circ$) ฉะนั้น

$$\therefore h_2^2 = A_{21}^2 + A_{22}^2 + A_{23}^2 + A_{24}^2 + A_{25}^2$$

แต่ A_{23}^2 , A_{24}^2 , A_{25}^2 มีค่าเป็นศูนย์

$$\therefore A_{22} = \sqrt{h_2^2 - A_{21}^2}$$

แทนค่า

$$A_{22} = \sqrt{.64 - (.6656)^2} = .4438$$

ในการคำนวณค่าองค์ประกอบตัวที่สองของเครื่องมือวัดอันที่สามที่สี่และที่ห้าก็ทำได้จากการเทียบค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างเครื่องมือวัดเหล่านี้แต่ละอันกับเครื่องมือวัดอันที่สอง คือ

$$\therefore R_{23} = A_{21} A_{31} + A_{22} A_{32} + A_{23} A_{33} + A_{24} A_{34} + A_{25} A_{35}$$

แต่ $A_{23} A_{33}$, $A_{24} A_{34}$, $A_{25} A_{35}$ มีค่าเป็นศูนย์

$$\therefore A_{32} = \frac{R_{23} - A_{21} A_{31}}{A_{22}}$$

แทนค่า

$$A_{32} = \frac{.00 - (.6656) (.4992)}{.4438} = - .7489$$

โดยวิธีเดียวกัน จะได้ค่า

$$A_{42} = \frac{R_{24} - A_{21} A_{41}}{A_{22}} = - .4714$$

$$A_{52} = \frac{R_{25} - A_{21} A_{51}}{A_{22}} = .4161$$

เมื่อกำหนดค่าองค์ประกอบตัวที่สาม

$$\therefore h_3^2 = A_{31}^2 + A_{32}^2 + A_{33}^2 + A_{34}^2 + A_{35}^2$$

แต่ A_{34}^2 , A_{35}^2 มีค่าเป็นศูนย์

$$\therefore A_{33} = \sqrt{h_3^2 - A_{31}^2 - A_{32}^2}$$

แทนค่า

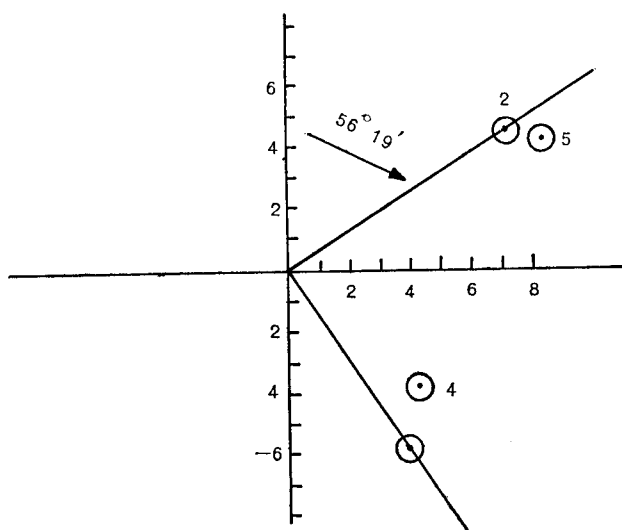
$$A_{33} = \sqrt{.81 - (.4992^2 + .7489^2)} = .0072$$

ค่าที่ได้นี้มีปริมาณเป็นเศษของ .00 (within rounding error of zero) หมายความว่าไม่อาจแยกองค์ประกอบตัวที่สามจากเครื่องมือวัดอันที่สาม และเมื่อทดลองเอาเครื่องมือวัดอันที่สี่และที่ห้ามาจัดลำดับใหม่ให้เป็นลำดับที่สาม ก็จะได้ค่าองค์ประกอบตัวที่สามที่มีปริมาณเป็นเศษของ .00 เช่นเดียวกัน ดังนั้นก็ถือได้ว่ามีองค์ประกอบเพียงตัวที่หนึ่งและตัวที่สองเท่านั้น ผลของการคำนวณองค์ประกอบของเครื่องมือวัดต่างๆ มีแสดงในตารางที่ 23

ตารางที่ 23
ค่าองค์ประกอบที่คำนวณโดยวิธีช่วงทแยง

เครื่องมือวัด	องค์ประกอบ					h ²
	1	2	3	4	5	
1	.7211	0	0	0	0	.52
2	.6656	.4438	0	0	0	.64
3	.4992	-.7489	.00	0	0	.81
4	.5547	-.4714	.00	.00	0	.53
5	.8043	.4161	.00	.00	.00	.82

จากตารางที่ 23 จะเห็นว่าค่าองค์ประกอบบางตัวติดลบ ในทางปฏิบัติเราจะไม่ทราบค่าองค์ประกอบที่แท้จริง ในการนี้ จึงได้มีหลักว่าจะต้องหมุนแกนขององค์ประกอบโดยใช้มาตรฐานค่าบวก (criterion of positive manifold)⁷ กล่าวคือ จะต้องหมุนแกนไปจนกว่าจะได้ค่าองค์ประกอบทั้งหมดเป็นค่าบวก เฉพาะค่าองค์ประกอบจากตัวอย่างข้างต้น ก็อาจหมุนแกนไปตามเข็มนาฬิกาเป็นมุม $56^{\circ} 19'$ ก็จะทำให้ได้ค่าทั้งหมดเป็นบวกโดยที่องค์ประกอบทุกตัวจะอยู่ภายในเสี้ยว (quadrant) เดียวกันของวงกลมดังแสดงในแผนภาพที่ 14



แผนภาพที่ 14 องค์ประกอบที่หมุนแกนแล้ว

⁷ Guilford, *op. cit.*, p. 485.

ค่าองค์ประกอบที่หมุนแกนแล้วอาจคำนวณได้ โดยใช้สูตรการหมุนไปรอบตัวตั้งได้ กล่าวไว้ในตอนต้นบท คือ

$$A'_{j_1} = A_{j_1} \cos \phi - A_{j_2} \sin \phi$$

$$A'_{j_2} = A_{j_1} \sin \phi + A_{j_2} \cos \phi$$

เนื่องจาก ค่า $\sin 56^\circ 19'$ เท่ากับ .8321 และ ค่า \cos ของมุมเดียวกันเท่ากับ .5546 ฉะนั้น

องค์ประกอบตัวแรกและตัวที่สองของเครื่องมือวัดอันที่หนึ่ง คือ

$$\begin{aligned} A'_{j_1} &= (.7211)(.5546) - (0)(.8321) \\ &= .3999 \end{aligned}$$

$$\begin{aligned} \text{และ } A'_{j_2} &= (.7211)(.8321) - (0)(.5546) \\ &= .6000 \end{aligned}$$

องค์ประกอบของเครื่องมือวัดทั้งสองตัวของเครื่องมือวัดอันนี้อาจหาได้โดยทำนองเดียวกัน ดังได้แสดงผลการคำนวณไว้ในตารางที่ 24

ตารางที่ 24

ค่าองค์ประกอบที่หมุนแกนแล้ว

เครื่องมือวัด	องค์ประกอบ		h^2
	1	2	
1	.4	.6	.52
2	.0	.8	.64
3	.9	.0	.81
4	.7	.2	.53
5	.1	.9	.82

การตั้งชื่อองค์ประกอบ เมื่อทราบค่าองค์ประกอบแล้ว การดำเนินงานขั้นต่อไปคือการพิจารณาว่า เครื่องมือวัดหรือข้อมูลชุดใด มีค่าองค์ประกอบร่วมกันสูง เมื่อเทียบกับค่าองค์ประกอบตัวเดียวกันของข้อมูลชุดอื่น ๆ เมื่อทราบเช่นนี้แล้วก็พิจารณาต่อไปว่า ลักษณะของกลุ่มเครื่องมือวัดที่มีค่าองค์ประกอบร่วมกันสูงนี้มีอะไรพิเศษแตกต่าง

จากกลุ่มเครื่องมือวัดที่มีค่าองค์ประกอบต่ำ องค์ประกอบตัวนี้ก็จะมิชื่อตามลักษณะพิเศษของกลุ่มเครื่องมือวัดที่มีค่าองค์ประกอบร่วมกันสูง

ในตัวอย่าง การตั้งชื่อองค์ประกอบของ Fructer⁸ ได้ชี้ว่า ในบรรดาเครื่องมือวัด 11 ชุด เครื่องมือวัดอันดับสาม ชื่อ การอ่านเพื่อความเข้าใจ และเครื่องมือวัดอันดับสี่ ชื่อ คุณพินิจในเชิงปฏิบัติ มีค่าองค์ประกอบตัวที่หนึ่ง .54 เท่ากัน และมีค่าองค์ประกอบตัวที่สี่ .42 และ .41 ตามลำดับ นับว่าสูงอย่างเห็นได้ชัดกว่าค่าองค์ประกอบตัวเดียวกันของเครื่องมือวัดอันดับอื่น ๆ และเนื่องจากเครื่องมือวัดอันดับสามมีเนื้อหาที่จะวัดความแตกต่างระหว่างบุคคลในการเข้าใจภาษา ส่วนเครื่องมือวัดอันดับสี่วัดความสามารถในการใช้ถ้อยคำเพื่ออธิบายสภาพการณ์ที่จะต้องมีการตัดสินใจ ส่วนเครื่องมือวัดที่มีค่าองค์ประกอบสองตัวต่ำ ไม่มีเนื้อหาที่จะวัดความเข้าใจในด้านภาษา ฉะนั้น จึงอาจตั้งชื่อองค์ประกอบตัวที่หนึ่งและตัวที่สี่ว่า องค์ประกอบความเข้าใจภาษา (verbal-comprehension factor) การตั้งชื่อองค์ประกอบตัวอื่น ๆ ก็กระทำในทำนองเดียวกัน เช่น องค์ประกอบตัวที่สอง ชื่อ ความสามารถในการใช้ตัวเลข เพราะเครื่องมือวัดหลายอันที่มีค่าองค์ประกอบตัวนี้สูง ต่างวัดความเร็วและความแม่นยำในทางเลขคณิต องค์ประกอบตัวที่สาม ชื่อ ความเร็วทางสัญชาน (perceptual speed) และองค์ประกอบตัวที่ห้าชื่อ ความสัมพันธ์ในเชิงว่าง (spatial relations) ดังนี้ เป็นต้น

⁸ Kerlinger, Fred N., *Foundations of Behavioral Research*, pp. 672—674.
Fruchter, *op. cit.*, pp. 120, 147, 149—150.

ทฤษฎีการวัดเชาว์

ในบทที่หนึ่งได้กล่าวมาแล้วว่า การวัดว่าบุคคลแต่ละคนสามารถทำอะไรได้บ้าง เป็นงานส่วนหนึ่งของการวัดผลและประเมินบุคคล ความสามารถในการทำอะไรได้นั้น แยกเป็น 2 ประเภทใหญ่ คือ ความถนัดทางธรรมชาติ ได้แก่ ความสามารถที่จะเรียนรู้ เพื่อไปปฏิบัติงานในระยะต่อไป และความสัมฤทธิ์ผล ได้แก่ ผลการเรียนรู้ที่จะนำไป ใช้ได้ในการปฏิบัติงาน เฉพาะในเรื่องความถนัดทางธรรมชาตินั้น ยังแยกออกไปได้อีก เป็น 2 อย่าง คือ ความสามารถทางสมอง (intellectual ability) และความถนัดทางธรรมชาติพิเศษ (special aptitude) ในบทนี้จะพิจารณาเฉพาะความสามารถทางสมอง ส่วน ความถนัดทางธรรมชาติพิเศษ จะแยกพิจารณาในบทต่อไป

เมื่อก้าวถึงความสามารถทางสมอง ก็มีศัพท์ที่ใช้เรียกลักษณะความสามารถชนิดนี้ หลายคำ เช่น เชาว์ (intellect) สติปัญญา (intelligence) ความเฉลียวฉลาด (sagacity) ความมีสมองดี (brains) เป็นต้น เมื่อก้าวถึงเชาว์ก็อาจนึกถึงความสามารถในการคิด หาเหตุผล และเข้าใจสิ่งต่างๆ ได้ในฉับพลัน สติปัญญาอาจทำให้นึกถึงความสามารถ ในการศึกษาหาความรู้และประสบการณ์ ความเฉลียวฉลาดอาจทำให้นึกถึงความสามารถ ในการหยั่งรู้ และการมองเห็นการณ์ไกลในการใช้ดุลพินิจในเรื่องต่างๆ และความมี สมองดี ก็อาจทำให้นึกถึงความสามารถทั่วไปที่ต้องใช้ความคิดและความเข้าใจซึ่งเป็น ศักยภาพของสมอง ศัพท์ทั้ง 4 คำที่ยกมาเป็นตัวอย่างนี้มีความแพร่หลายในการใช้ ในภาษาเขียนไม่เท่ากัน ศัพท์ที่ใช้แพร่หลายมากกว่าคำอื่น ๆ ก็คือเชาว์หรือสติปัญญา โดยเฉพาะในระเบียบและกฎข้อบังคับที่บังคับใช้ในประเทศนั้น ใช้คำว่า เชาว์ เป็นหลัก ส่วนในต่างประเทศนั้นก็ใช้คำ intelligence เป็นหลัก อย่างไรก็ตาม เพื่อ หลีกเลี่ยงการใช้ศัพท์ไม่ตรงกันในตำราการวัดผลและประเมินบุคคล ในปัจจุบันจึงใช้

คำว่า intellectual ability โดยให้หมายถึงความสามารถทั่วไปในการเรียนรู้ แต่ในตำราเล่มนี้จะเรียกความสามารถทางสมองนี้ว่า เซาว์น เพื่อให้ผู้อ่านได้เกิดมโนภาพ เพราะเป็นศัพท์ที่กะทัดรัดและรู้จักกันอย่างกว้างขวางอยู่แล้ว

ความเป็นมาของทฤษฎีการวัดเซาว์น

ก่อนที่จะอ่านหัวข้อนี้ ผู้อ่านควรจะได้เข้าใจประวัติการวัดผลบุคคล ซึ่งได้กล่าวไว้โดยสาระสำคัญแล้วในบทที่หนึ่ง เพราะการวัดผลบุคคลมีประวัติเกี่ยวพันกับการวัดเซาว์น หรือความสามารถทางสมองอย่างใกล้ชิด

Sir Francis Galton ชาวอังกฤษ เชื่อว่า ความฉลาดและบุคลิกภาพเป็นเรื่องของกรรมพันธุ์ จึงได้พยายามสร้างเครื่องมือที่จะวัดความฉลาดของบุคคล และได้คิดศัพท์เรียกเครื่องมือนี้ว่า แบบทดสอบความฉลาด (mental test) วิธีวัดความฉลาดตามแนวคิดของ Galton ทำได้โดยวัดความสามารถทางประสาท (sensory discrimination) เพราะถือตามแนวคิดของนักปรัชญา Locke ว่า บุคคลที่มีความสามารถทางประสาทสูงเป็นผู้ที่มีพรสวรรค์ดีและสามารถเรียนรู้ได้ดี ความสามารถทางประสาทที่วัดได้ได้แก่ความสามารถในการฟังระดับเสียง ความว่องไวของจักขุ ความสามารถในการแยกสี ความสามารถในการได้ยิน เป็นต้น¹ ผลงานของ Galton ก็ได้ประมวลวิเคราะห์ในเชิงสถิติ ทำให้ยืนยันได้ว่าบุคคลแต่ละคนมีความแตกต่างกัน แต่ก็มีได้กำหนดแบบมาตรฐานอันใดในการวัดความฉลาดของบุคคล ฉะนั้นในสหรัฐฯ จึงได้มีผู้นำเอาหลักสถิติของ Galton ไปทำการวิจัยความฉลาดของนักศึกษามหาวิทยาลัย Columbia University ผู้หนึ่ง คือ James McKeen Cattell² ในระหว่างปี 1890—96 Cattell ได้ทดลองแบบทดสอบความฉลาด ที่วัดความสามารถทางประสาทในด้านต่าง ๆ เช่น ช่วงระยะเวลาที่บุคคลมีปฏิกิริยาโต้ตอบ เวลาที่ใช้ในการตัดสินใจ ความสามารถในการฟังเสียงสูงต่ำ การพิจารณาความหนักของสิ่งของโดยไม่ต้องชั่ง ความสามารถจำอักษรได้ เป็นต้น แต่ผลปรากฏว่า แบบทดสอบเหล่านี้มีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันต่ำมากและยังไม่มีความสัมพันธ์กับผลการศึกษาศึกษาบุคคลอีกด้วย

อย่างไรก็ดี ในปี 1904 Charles Spearman ชาวอังกฤษก็ได้ชี้ให้เห็นว่า การที่สัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบความสามารถทางประสาทต่ำ ก็เพราะแบบทดสอบแต่ละชุดขาดความเชื่อมั่น หรือมีสัมประสิทธิ์ความเชื่อมั่นต่ำ เมื่อหักค่า

¹ Nunnally, Jum C., *Psychometric Theory*, pp. 428—429.

² Jackson, Douglas N., and Messick, Samuel, eds., *Problems in Human Assessment*, pp. 413—414.

ผลิตผลของการวัด (ดูเทียบกับบทที่ 4) ผลของการวัดด้วยเครื่องมือวัดความสามารถทางประสาทชุดต่าง ๆ จะมีค่าความแปรปรวนร่วมกันในระดับค่อนข้างสูง ทั้งยังมีความสัมพันธ์กับผลการศึกษาก็ด้วย ซึ่งทำให้ Spearman เชื่อว่า ในเรื่องความสามารถทางสมองจะต้องมีอะไรอย่างหนึ่งที่เป็นตัวกลาง จึงได้ทำการทดลองวัดผลการศึกษาในวิชาสาขาต่างๆ ของบุคคล เช่น ภาษาฝรั่งเศส ภาษาอังกฤษ คณิตศาสตร์ ดนตรี ก็ปรากฏว่า ผลการศึกษาวิชาที่ไม่เหมือนกันเหล่านี้ มีสิ่งที่เป็นตัวกลางเหมือนกันจริง (ดูเทียบกับการวิเคราะห์องค์ประกอบในบทที่ 7) ผลการวิจัยนี้สรุปได้ว่า เครื่องมือวัดความฉลาดของบุคคล จะมีตัวกลางที่เป็นเครื่องแสดงความสามารถทั่วไป (general ability) กำกับเป็นหลักอยู่ และจะมีลักษณะบางอย่างที่มีคุณสมบัติพิเศษเฉพาะเครื่องมือวัดแต่ละอันประกอบอยู่ด้วย ลักษณะพิเศษนี้เรียกว่า ความสามารถพิเศษเฉพาะ แนวคิดนี้คือหลักของทฤษฎีองค์ประกอบสองตัวของ Spearman ซึ่งได้อธิบายมาแล้วในบทที่ 3

ในตอนที่ Spearman พยายามค้นคว้าทฤษฎีการวัดบุคคล Alfred Binet ชาวฝรั่งเศส ก็ได้สร้างเครื่องมือวัดสติปัญญาสำเร็จ Binet ได้ร่วมมือกับ Theodore Simon สร้างแบบทดสอบ ซึ่งประกอบด้วยปัญหา 30 รายการ เสร็จในปี 1905 ตามแนวคิดในการวัดความสามารถทางสมองของบุคคลนั้นไม่จำเป็นต้องวัด “ส่วนต่างๆ” (elements) ของความสามารถ แต่ควรจะวัดผลการแสดงความสามารถ ซึ่งอาจเรียกว่า สติปัญญา (intelligence) คำถามในแบบทดสอบนี้จึงเกี่ยวกับ ความคิด ความเข้าใจ และการตัดสินใจในเรื่องต่างๆ ภายในสังคม เช่น ความสามารถในการอธิบายลักษณะของบ้านพ่อแม่ ซ้อนส้ม และม้า หรือความสามารถในการพูดซ้ำประโยคที่มีคำถึง 15 คำ เมื่อได้ยินเพียงครั้งเดียว ในปี 1908 Binet ได้ขยายแบบทดสอบให้มีความยากง่ายเป็นหลายระดับ โดยให้สอดคล้องกับเกณฑ์เฉลี่ยของผู้มีอายุต่าง ๆ กันไป Lewis M. Terman และ Maude A. Merrill ชาวอเมริกันได้นำแบบทดสอบของ Binet มาแปลเป็นภาษาอังกฤษ และปรับปรุงอีกหลายครั้งจนแบบทดสอบนี้เป็นที่แพร่หลาย และเป็นแม่บทของเครื่องมือวัดความสามารถทางสมองในปัจจุบัน กล่าวคือ ได้มีการแก้ไข ในปี 1916 ปี 1937 และปี 1960 และแบ่งสติปัญญาของบุคคลออกเป็น 20 ระดับ ใช้ทดสอบกับเด็กตั้งแต่อายุ 2 ขวบ ไปจนถึงผู้ใหญ่ แบบทดสอบที่ใช้กับเด็กเป็นคำถามง่าย ๆ ใช้รูปภาพและวัตถุเป็นหลัก ส่วนที่ใช้กับผู้ใหญ่เป็นคำถามเกี่ยวกับแนวความคิด และความหมายต่างๆ ในภาษา สารระสำคัญขอเนื้อหาในแบบทดสอบมีในตัวอย่างข้างล่าง³

³ Thorndike, Robert L., and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*. pp. 210-211.

Nunnally, *op. cit.*, p. 430.

ระดับอายุสองขวบครึ่ง :

1. การรู้จักสิ่งของเครื่องใช้ (มีสิ่งของ 6 อย่าง)
“จงชี้ว่าอันไหน สำหรับดื่ม” ฯลฯ
2. อวัยวะและส่วนต่างๆ ของร่างกาย (มีตุ๊กตากระดาดตัวใหญ่)
“ชี้ว่าผมอยู่ตรงไหน” ฯลฯ
3. การเรียกชื่อ (ของเล่น 4 อย่างเช่น เก้าอี้ รถยนต์ เป็นต้น)
“สิ่งนี้เรียกชื่อว่าอะไร” ฯลฯ
4. ศัพท์ (มีรูปภาพ 18 รูป เป็นรูปสิ่งต่างๆ ที่รู้จักกันโดยทั่วไป)
“รูปนี้ รูปอะไร” ฯลฯ
5. พุดซ้ำเลขสองตัว
“จงพุดตาม เลข 4 เลข 7” ฯลฯ

การจับของใส่ช่อง (มีแผ่นกระดานที่เจาะช่องหลาย ๆ ช่อง เป็นสี่เหลี่ยม วงกลม และสามเหลี่ยม พร้อมทั้งแท่งไม้ที่จะใส่ไปในช่องเหล่านี้ได้พอดี)
“ให้เอาแท่งไม้ใส่ช่อง”

ระดับผู้ใหญ่ :

1. คำศัพท์ (มีศัพท์ 45 คำ เรียงตามความยากง่าย)
“ให้อธิบายความหมายของศัพท์ว่า อวดหรา (flaunt)” ฯลฯ
2. การหาเหตุผล (2 ภาค)
“ให้เหตุผล 3 ประการว่า ทำไมบุคคลที่ก่ออาชญากรรมอย่างร้ายแรง ควรได้รับการลงโทษ”
3. พุดซ้ำเลขแปดตัว
4. ความหมายของสุภาษิต
“จงอธิบายความหมายของประโยคว่า สิบเบี้ยใกล้มือ (bird in the hand)”
5. ลักษณะที่คล้ายกันระหว่างสิ่งที่ต่างกัน (6 ภาค)
“ความหนักและเบาเหมือนกันอย่างไร” ฯลฯ
6. พุดข้อความซ้ำ
“เมื่ออ่านข้อความต่อไปจนจบแล้ว จงพุดซ้ำให้มากที่สุดเท่าที่จะทำได้.....”

ผู้ที่ทำแบบทดสอบได้ในระดับอายุของตน หรือระดับอื่นๆ ก็สามารรถำนวนผลสติปัญญาได้ตามวิธีของ William Stern เช่น สมมุติว่า นาย ก. อายุ 20 ปี ทำแบบทดสอบได้ดังต่อไปนี้

ทำได้ทั้งหมดในระดับอายุ 6 ขวบ = 6 ปี
 ทำได้อีก 1 ใน 2 ของระดับ 7 ขวบ = 6 เดือน
 ทำได้อีก 1 ใน 6 ของระดับ 8 ขวบ = 2 เดือน
 ทำไม่ได้เลยในระดับ 9 ขวบ = 0
 รวมจำนวนปี = 6 ปี 8 เดือน

นำจำนวนปีนมากำหนดเขาวนตามสูตร

$$I.Q. = \frac{M.A.}{C.A.} (100)$$

โดย I.Q. = อัตราร้อยละของเขาวน (intelligence quotient)

M.A. = จำนวนปีของสติปัญญาคิดเป็นเดือน (mental age)

C.A. = จำนวนปีของอายุจริงคิดเป็นเดือน (chronological age)

$$I.Q. = \frac{80}{240} (100)$$

$$= 33.34$$

การที่นาย ก. มี I.Q. เพียงเท่านั้น หมายความว่า นาย ก. มีระดับเขาวนต่ำ คะแนน I.Q. ที่จะได้จากแบบทดสอบ Binet นี้ อาจแยกตามจำนวนผู้ที่ควรจะได้ ตามอัตราดังต่อไปนี้

ค่า I.Q.	อัตราร้อยละของผู้ที่จะได้	อัตราร้อยละสะสม
140 ขึ้นไป	1.3	99.9
130-139	3.1	98.6
120-129	8.2	95.5
110-119	18.1	87.3
100-109	23.5	69.2
90- 99	23.0	45.7
80- 89	14.5	22.7
70- 79	5.6	8.2
60- 69	2.0	2.6
60	0.6	0.6

เครื่องมือวัดของ Binet สามารถแสดงผลของการมีความสามารถทางสมอง เช่น ความชำนาญด้านภาษา (language mastery) ความสามารถในการหาเหตุผล (reasoning) และความสามารถในการประมวลข้อเท็จจริง (assimilation of facts) เป็นต้น แต่แนว

วิเคราะห์ของ Binet มิได้ช่วยให้แก่นักวัดบุคคลทราบว่าเขาจะ หรือความสามารถทางสมองนี้ประกอบด้วยอะไรบ้าง ส่วนทฤษฎีองค์ประกอบสองตัวของ Spearman นั้น แม้ได้ช่วยให้ได้เครื่องมือวัดเขาแล้ว แต่ก็ช่วยปูพื้นฐานในการวิเคราะห์เขาแล้ว อันจะเป็นทางให้สามารถสร้างเครื่องมือวัดที่สมบูรณ์อย่างแท้จริงต่อไป

สหรัฐอเมริกาได้รับแนวความคิดในเรื่องทฤษฎีองค์ประกอบจากอังกฤษ และวิธีการวัดผลที่เกิดจากความสามารถทางสมองของฝรั่งเศสมาดำเนินการวิจัยต่ออย่างกว้างขวาง บุคคลสำคัญชาวอเมริกัน คือ L.L. Thurstone ได้ค้นคว้าทฤษฎีกลุ่มขององค์ประกอบในปี 1938 ซึ่งค่านทฤษฎีองค์ประกอบสองตัวว่าหาขาดเกินไป การค่านทฤษฎีเดิมอาจทดลองได้โดยแบบทดสอบที่ใช้ภาษา (verbal test) มา 2 กลุ่ม ๆ ละ 3 ชุด สัมประสิทธิ์สหสัมพันธ์ระหว่างชุดภายในกลุ่มจะสูงมากกว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างกลุ่ม เช่น สัมประสิทธิ์สหสัมพันธ์ระหว่างกลุ่มมีค่า 0.30 สัมประสิทธิ์สหสัมพันธ์ภายในกลุ่มอาจสูงถึง 0.70 การที่เป็นเช่นนั้น จะถือว่าเมื่อองค์ประกอบร่วมที่มีผลกระทบต่อเครื่องมือวัดทุกชุดอยู่หนึ่งตัว และเมื่อองค์ประกอบพิเศษอีก 6 ตัว ย่อมไม่เป็นเหตุผล Thurstone ได้พิสูจน์ว่า ความสามารถของบุคคลอาจแยกเป็นองค์ประกอบได้ถึง 9 ตัว รายละเอียดของทฤษฎีกลุ่มองค์ประกอบได้กล่าวไว้แล้วในบทที่ 3 จึงจะไม่กล่าวซ้ำในที่นี้อีก ต่อจาก Spearman ก็ได้มีการค้นคว้าองค์ประกอบของเขาเรื่อย ๆ ในปัจจุบันได้มีการค้นพบองค์ประกอบเป็นร้อยตัว⁴ จากการวิเคราะห์องค์ประกอบนี้จึงทำให้เกิดทฤษฎีที่ยึดถือในปัจจุบันอย่างแพร่หลาย คือ ทฤษฎีโครงสร้างของเขา (structure of intellect) ของ J.P. Guilford ชาวอเมริกัน

ทฤษฎีโครงสร้างของเขา

Guilford ได้เสนอทฤษฎีโครงสร้างของเขา โดยได้ไปปรากฏการณ์ที่มหาวิทยาลัย Stanford University เมื่อวันที่ 13 เมษายน ค.ศ. 1959 และได้พิมพ์เผยแพร่ทฤษฎีนี้ในวารสาร American Psychologist ในปีเดียวกัน และได้อธิบายความละเอียดของทฤษฎีอีกในวารสาร Psychological Review ในปี 1961 สารสำคัญของทฤษฎีนี้มีดังต่อไปนี้⁵

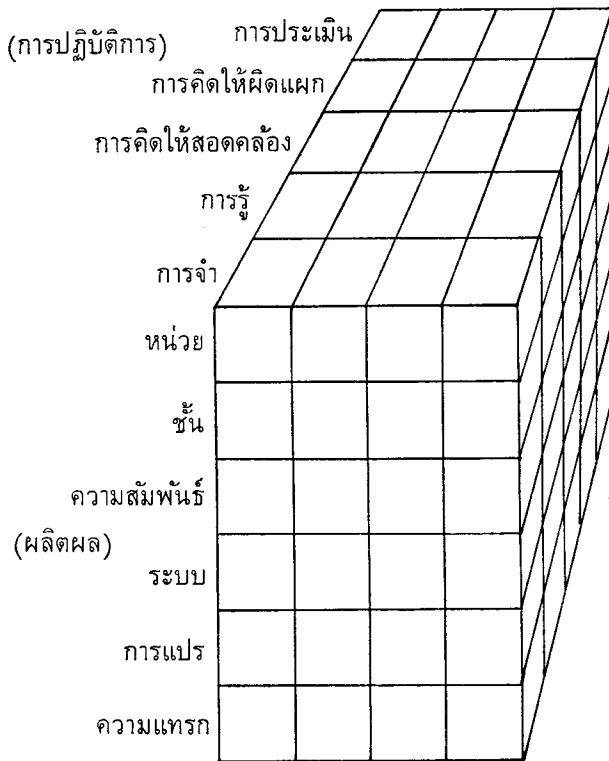
เขามีสถิติ ในการวิเคราะห์เขา อาจพิจารณาได้เป็น 3 แง่ คือ ในแง่ของสิ่งที่มาแรงเร้าให้เกิดปฏิกิริยาได้ตอบ ซึ่งเรียกได้ว่าเป็นเนื้อหา (content) ของสิ่งต่าง ๆ ในสังคมเป็นประการแรก ประการที่สองในแง่ลักษณะของขบวนการการแสดงออก

⁴ *Ibid.*, p. 442.

⁵ Guilford, J.P., "Three Faces of Intellect," in Jackson and Messick, eds., *op. cit.*, pp. 421-433.

ซึ่งเรียกได้ว่า การปฏิบัติการ (operations) ของผู้มีเชาวน์ และประการสุดท้าย ในแง่ของผลที่เกิดจากการแสดงออก ซึ่งเรียกได้ว่า ผลิตภัณฑ์ (product) ดังแสดงไว้ในแผนภาพที่ 15

เมื่อพิจารณาเชาวน์ในแง่เนื้อหา ก็อาจนำสิ่งต่างๆ ที่มีผลกระทบต่อความรู้สึกรวมของบุคคลมาแยกประเภท ทิ้งแยกได้เป็น 4 ประเภท คือ เนื้อหาที่เป็นรูป (figural content) หมายถึงวัตถุสิ่งของต่างๆ ที่ประสาทจะรับรู้ ได้แก่สิ่งที่มองเห็น หรือได้ยินหรือรู้สึกสัมผัสได้ เนื้อหาที่เป็นเครื่องหมาย (symbolic content) เครื่องหมายต่างๆ ที่ไม่มีความหมายในตัวเอง แต่มีความหมายแทนสิ่งต่างๆ เช่น ตัวเลข ตัวอักษร ซึ่งจัดไว้เป็นระบบแล้ว เนื้อหาที่เป็นภาษา (semantic content) หมายถึง ข้อความหรือข่าวสารที่แสดงแนวคิดให้เป็นที่เข้าใจระหว่างผู้พูดหรือผู้เขียน กับผู้ฟังหรือผู้อ่าน เนื้อหาที่เป็น



รูป เครื่องหมาย ภาษา พद्यติกรรม (เนื้อหา)

แผนภาพที่ 15 โครงสร้างของเชาวน์ที่มีสามมิติ

แหล่งที่มา : Guilford, J.P., "Three Faces of Intellect," in Jackson, Douglas N., and Messick, Samuel, eds., *Problems in Human Assessment*, p. 424.

พฤติกรรม (behavioral content) ได้แก่ สภาพการณ์ทางสังคมซึ่งบุคคลได้รับข่าวสารจากพฤติกรรมของผู้อื่น เช่น การเข้าใจ หรือการแปลความหมาย หรือแนวคิดจากการประพฤตินิติปฏิบัติของบุคคลต่างๆ ในเชิงปฏิบัติ เนื้อหาต่างๆ เหล่านี้สะท้อนอยู่ในสาระสำคัญของเครื่องมือวัด หรือแบบทดสอบที่ใช้

การพิจารณาเขาวงกตในแง่การปฏิบัติการ หมายความว่า การพิจารณาว่าบุคคลแต่ละคนได้มีการดำเนินการอย่างไรในการแสดงความสามารถทางสมอง ขบวนการที่เห็นได้ชัดก็คือ การจำ (memory) และการคิด (thinking) การจำหมายถึงการเก็บรักษาข่าวสารไว้ในตัว ส่วนการคิดเป็นการนำข่าวสารมาใช้หรือปฏิบัติการอย่างหนึ่งอย่างใด ในการทำการคิดนั้นยังแยกได้เป็นอาการ 3 อย่างใหญ่ๆ คือ การรู้ (recognition) การผลิต (production) และการประเมิน (evaluation) การรู้ คือ อาการที่บุคคลรู้จักหรือรับรู้แนวความคิดของสิ่งต่างๆ การรู้ทำให้บุคคลเข้าใจสภาพการณ์ต่างๆ ของสังคม การผลิตเป็นอาการที่บุคคลนำเอาสิ่งที่รู้ และเข้าใจแล้วมาคิดต่อเพื่อให้เกิดความรู้แตกแขนงออกไป อาการนี้แยกได้อีกเป็น 2 อย่าง คือ การคิดให้สอดคล้อง (convergent production of ideas) และการคิดให้ผิตแผก (divergent production of ideas) การคิดให้สอดคล้องเป็นการคิดตามข่าวสารที่มีเพื่อให้หาผลสรุปคำตอบที่ดีที่สุดตามแนวคิดที่เชื่อถือกันอยู่ในวงการหนึ่งๆ หรือสังคมหนึ่งๆ ส่วนการคิดให้ผิตแผกเป็นการคิดในทิศทางที่ต่างไปจากแนวคิดที่เชื่อถือกันอยู่ เพื่อให้สามารถค้นพบสิ่งแปลกๆ ไปจากเดิม ส่วนการประเมินนั้นเป็นอาการที่แต่ละบุคคลใช้ดุลพินิจของตนลงความเห็นว่า สิ่งหนึ่งๆ หรือเรื่องหนึ่งๆ มีความเพียงพอ หรือถูกต้อง หรือดีแล้วอย่างไร

ในการพิจารณาเขาวงกตในแง่ผลิตผลเป็นการพิจารณาว่า เมื่อมีเนื้อหาอย่างหนึ่ง และมีการปฏิบัติการอย่างหนึ่งแล้ว จะเกิดผลอย่างไร ไม่ว่าจะใช้เนื้อหาประเภทใด และมีการปฏิบัติการอย่างใดก็ตาม ผลที่เกิดจะอยู่ใน 6 ประเภท ดังต่อไปนี้ คือ ในประเภทขั้นต่ำที่สุด บุคคลจะเข้าใจหรือจำได้ว่า มีแนวความคิดตามจุดต่างๆ อย่างไร ผลิตผลในขั้นนี้เรียกว่า เป็นหน่วย (units) ถ้าบุคคลมีความสามารถทางสมองสูงกว่าระดับนี้ก็จะนำเอาความคิดที่เข้าใจหรือจำได้ในประเด็นต่างๆ นั้นมาจัดเป็นกลุ่มตามลักษณะที่คล้ายกันหรือต่างกัน เรียกผลิตผลนี้ได้ว่า ชั้น (classes) ถ้ามีความสามารถสูงขึ้นไป

ก็อาจจัดได้ว่าความคิดเหล่านั้นมีความสัมพันธ์ (relations) กันอย่างไร โดยพิจารณาจากลักษณะของตัวแปรที่มีอิทธิพลต่อแนวความคิดเหล่านั้น ยิ่งถ้าหากมีความสามารถสูงขึ้นไปอีก ก็อาจแยกแยะความสลับซับซ้อนของความคิดทั้งหมดให้เป็นระบบ (systems) โครงสร้างของความคิดเหล่านั้น ผู้มีความสามารถสูงกว่านี้ ก็อาจนำระบบทั้งหมดมาเปลี่ยนแปลงให้มีลักษณะที่จะเป็นประโยชน์ในการพิจารณาหรือวิเคราะห์เรื่องอื่นๆ ได้อีก อากาการในขั้นนี้เรียกว่า การแปร (transformations) ในขั้นสุดท้ายบุคคลที่มีความสามารถทางสมอง อาจพิจารณาไปจนเห็นความแทรก (implications) ของแนวความคิดที่ได้รับจากข่าวสารที่มีอยู่ โดยพิจารณาตัวแปรทั้งหมดในพิสัยที่ทราบ แล้วทำการคาดคะเนหรือหยั่งถึงผลที่อาจเกิดขึ้นได้

ตารางองค์ประกอบของเชาวน์ เมื่อพิจารณาเทียบลักษณะต่างๆ ของเชาวน์ตามแผนภาพที่ 15 จะเห็นได้ว่า การวัดเชาวน์ตามทฤษฎีนี้ จะสามารถแยกองค์ประกอบออกไปได้ถึง 120 ตัว แต่ละช่องสี่เหลี่ยมเล็กๆ ในแผนภาพนี้ แทนองค์ประกอบออกไปได้ถึง 120 ตัว แต่ละช่องสี่เหลี่ยมเล็กๆ ในแผนภาพนี้ แทนองค์ประกอบหนึ่งตัว ซึ่งคำรวมของเนื้อหา การปฏิบัติกร และผลิตผลของเชาวน์

ตัวอย่างเช่น หากมีการปฏิบัติกรในเรื่อง การรู้ (cognition) จากเนื้อหา ประเภท รูป เครื่องหมาย และภาษา โดยมีผลิตผลในทุกประเภท จะได้องค์ประกอบที่แสดงเชาวน์ ดังแสดงไว้ในตารางที่ 25

ในตารางที่ 25 แสดงองค์ประกอบไว้ 17 ตัว ความจริงควรจะต้องมีตัวที่ 18 คือองค์ประกอบอันเป็นคำรวมระหว่างเครื่องหมายกับการแปร การที่เว้นว่างไว้ก็เพราะยังไม่สามารถตั้งศัพท์เรียกชื่อองค์ประกอบที่ควรจะมีนี้ได้

ตารางที่ 25*

องค์ประกอบที่เกี่ยวข้องกับการรู้ของเขาวน

ผลิตผล	เนื้อหา		
	รูป	เครื่องหมาย	ภาษา
หน่วย	การรู้ทางประสาท	การรู้เครื่องหมาย	การเข้าใจภาษา
ชั้น	การแยกประเภทของรูป	การรู้จักกลุ่มต่างๆของเครื่องหมาย	การแยกประเภทแนวคิด
ความสัมพันธ์	การสรุปความสัมพันธ์ของรูป	การสรุปความสัมพันธ์ของเครื่องหมาย	การสรุปความสัมพันธ์ของภาษา
ระบบ	การเข้าใจช่วง	การสรุประบบของเครื่องหมาย	การคิดหาเหตุผลทั่วไป
การแปร	การวาดภาพในใจ	—	ความหยั่งรู้
ความแทรก	การวาดภาพในระยะต่อไป	การรู้ความแทรกของเครื่องหมาย	การมีมโนภาพทางแนวคิด

*แหล่งที่มา : Guion, Robert M., *Personnel Testing*, p. 70.

องค์ประกอบในตารางนี้มีความหมายดังต่อไปนี้ คือ

1. การรู้ทางประสาท (visual cognition and auditory cognition) ความสามารถที่จะรู้จักรูปที่เคยเห็นในสภาพการณ์ที่ลำบากต่อจักษุประสาท และการจำเสียงแบบต่างๆ ได้
2. การรู้เครื่องหมาย (symbolic cognition) ความสามารถที่จะรู้จักหน่วยของเครื่องหมาย เช่น คำแต่ละคำ
3. การเข้าใจภาษา (verbal comprehension) ความรู้ในความหมายของคำ
4. การแยกประเภทของรูป (figural classification) ความสามารถที่จะชี้ว่ารูปแต่ละรูปมีลักษณะทำให้ต่างประเภทกันอย่างไร
5. การรู้จักกลุ่มต่างๆ ของเครื่องหมาย (symbolic classification) ความสามารถที่จะรู้ว่าเครื่องหมายแต่ละอันมีลักษณะทำให้ต่างประเภทกันอย่างไร

6. การแยกประเภทแนวคิด (cognition of symbolic class) ความสามารถที่จะชี้ว่าแนวคิดที่ใช้ภาษาแต่ละตอนมีลักษณะทำให้ต่างประเภทกันอย่างไร

7. การสรุปความสัมพันธ์ของรูป (eduction of figural relations) ความสามารถที่ทำให้ค้นพบว่ารูปต่างเกี่ยวข้องกันอย่างไร

8. การสรุปความสัมพันธ์ของเครื่องหมาย (eduction of symbolic relations) ความสามารถที่ทำให้ค้นพบว่าเครื่องหมายต่าง ๆ เกี่ยวข้องกันอย่างไร

9. การสรุปความสัมพันธ์ของภาษา (eduction of semantic relations) ความสามารถที่ทำให้ค้นพบว่า แนวคิดต่าง ๆ ในภาษาเกี่ยวข้องกันอย่างไร

10. การเข้าใจช่วงในมิติ (spatial orientation) การรู้จักและเข้าใจช่วงหรือมิติต่าง ๆ โดยเอาตนเองเป็นจุดตั้งต้น

11. ความหยั่งรู้ (penetration) ความสามารถเข้าใจสาระสำคัญของแนวความคิดแต่ละอันได้อย่างลึกซึ้ง

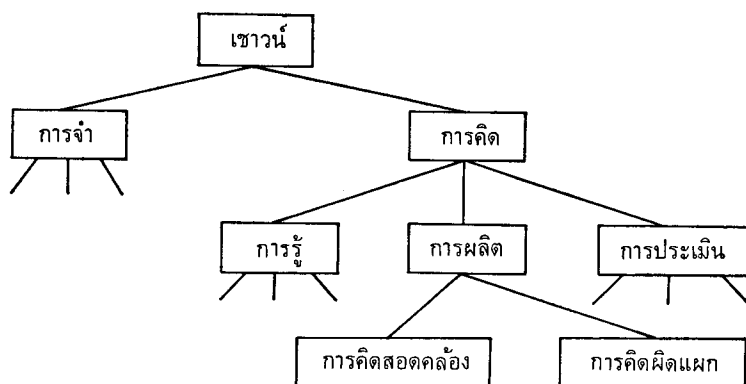
12. การวาดภาพในระยะต่อไป (perceptual foresight) ความสามารถที่จะนึกเห็นภาพของการปฏิบัติในแนวต่าง ๆ หากได้มีการกระทำขึ้นในระยะต่อไป

13. การรู้ความแทรกของเครื่องหมาย (symbolic perception) ความสามารถในการวางโครงสร้างของเครื่องหมายต่าง ๆ

14. การมีมโนภาพทางแนวคิด (conceptual foresight) ความสามารถมองเห็นล่วงหน้าได้ว่า แนวความคิดต่าง ๆ นั้นมีความหมายซ้อนแทรกอยู่อย่างไร

องค์ประกอบอื่นๆ ก็อาจพิจารณาได้ในทำนองเดียวกัน เช่น องค์ประกอบที่เป็นค่ารวมระหว่าง การจำ กับความสัมพันธ์และเครื่องหมาย คือ ความจำแบบนกแก้ว นกขุนทอง (rote memory) การจำกับความแทรกและเครื่องหมาย คือ ความคล่องแคล่วในเรื่องตัวเลข (numerical fluency) การคิดให้สอดคล้องกับความแทรกและภาษา คือ ความสามารถในการอนุมาน (deduction) ส่วนการคิดให้พิศมัยกับความแทรกและภาษา คือ ความสามารถขยายความ (semantic elaboration) เป็นต้น

ลำดับลดหลั่นขององค์ประกอบของเชวอร์ เนื่องจากองค์ประกอบที่แยกตามทฤษฎีนี้มีนับเป็นร้อยตัว จึงทำให้การที่จะสร้างเครื่องมือวัดตามทฤษฎีนี้เป็นไปด้วยความลำบาก ฉะนั้น จึงได้มีการกำหนดลำดับลดหลั่น (hierarchy) ขององค์ประกอบที่ทดลองได้มา เพื่อให้ทราบความสัมพันธ์ที่แท้จริงระหว่างองค์ประกอบแต่ละตัว ในแผนภาพที่ 16 แสดงลำดับลดหลั่นขององค์ประกอบในด้านการผลิต

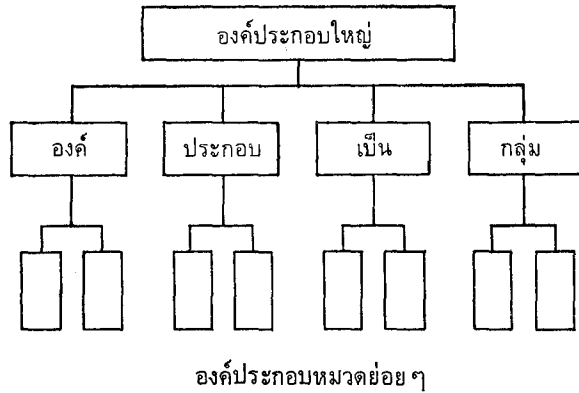


แผนภาพที่ 16 ลำดับลดหลั่นขององค์ประกอบของเชาวน์

แหล่งที่มา : Nunnally, Jum C., *Psychometric Theory*, p. 467.

แผนภาพที่ 16 แสดงว่า เมื่อพิจารณาเชาวน์ในแง่การปฏิบัติการ องค์ประกอบที่อยู่ลำดับชั้นแรก คือ ความจำ และการคิด ดังได้เคยอธิบายแล้ว ส่วนการรู้ การผลิตและการประเมินเป็นองค์ประกอบในลำดับชั้นถัดลงไป ในแผนภาพนี้จะเห็นได้ว่าลดหลั่นตามลำดับขององค์ประกอบในด้านการผลิตได้แสดงไว้โดยสมบูรณ์ที่สุด เพราะได้แยกแยะไปอีกว่าถัดจากองค์ประกอบในชั้นการผลิตไปแล้ว ยังมีองค์ประกอบอะไรอีกบ้าง จะเห็นได้ว่า หากพิจารณาแค่เพียงแผนภาพที่ 15 จะไม่ทราบลดหลั่นลำดับขององค์ประกอบต่าง ๆ เลย

การที่การกำหนดลำดับลดหลั่นองค์ประกอบช่วยให้ทราบความสัมพันธ์ระหว่างองค์ประกอบนั้น เป็นการอำนวยความสะดวกในการสร้างเครื่องมือวัด เพราะการสร้างเครื่องมือวัดให้ครอบคลุมองค์ประกอบทุกตัวโดยไม่มีขอบเขตจำกัดว่า จะเน้นหนักองค์ประกอบตัวใดตัวหนึ่งเป็นหลัก ย่อมทำได้ยาก นักจิตบุคลลชาวอังกฤษเช่น Cyril Burt และ Philip E. Vernon ได้สนับสนุนการกำหนดลำดับลดหลั่นขององค์ประกอบไว้ในที่หลายแห่ง อย่างไรก็ตามก็ดีกว่าเรื่องลำดับลดหลั่นขององค์ประกอบนี้ยังไม่เป็นที่ยุติ เพราะผลการวิจัยในเรื่องความสามารถของบุคคลนั้นยังไม่เป็นที่สิ้นสุด ดังนั้นในที่นี้จึงจะเสนอแต่เพียงโครงแบบ (model) ของลำดับลดหลั่นที่ควรจะเป็นตามแนวคิดของ Humphreys



แผนภาพที่ 17 โครงแบบของลำดับลดหลั่นองค์ประกอบสามชั้น

แหล่งที่มา : Humphreys, Lloyd G., "The Organization of Human Abilities," in Jackson and Messick eds., *Problems in Human Assessment*, p. 453.

วิธีการวัดเชาวน์

นับตั้งแต่ Binet ได้สร้างแบบทดสอบเชาวน์ ก็ได้มีผู้สร้างแบบทดสอบเชาวน์ขึ้นอีกมากมาย แบบทดสอบบางแบบใช้เวลาทำเพียง 15 นาที บางแบบก็ต้องใช้เวลาถึงกว่าชั่วโมง แต่แบบทดสอบทั่วไป ก็มีได้วัดเชาวน์อย่างแท้จริง⁶ เพราะดังที่ได้อธิบายไว้แล้วในทฤษฎีโครงสร้างของเชาวน์ การวัดเชาวน์จริงๆ นั้น ทำได้ด้วยความลำบากเพราะเชาวน์มีองค์ประกอบมากมายจนไม่สามารถทราบลำดับลดหลั่นขององค์ประกอบ แม้แบบทดสอบของ Binet ที่วัดผลมาได้เป็น I.Q. นั้น ก็เป็นเพียงการวัดผลของการที่บุคคลที่มีความสามารถทางสมอง ไม่ใช่เป็นการวัดความสามารถทางสมองของบุคคลโดยตรง ในการวัดผลของการมีความสามารถทางสมองนี้ ก็วัดได้เพียงว่า มีความสามารถในด้านต่างๆ เช่น ความสามารถในการเข้าใจภาษา ความสามารถในการหาเหตุผล ความสามารถในการจำความคล่องในเรื่องตัวเลข เป็นต้น ยังไม่สามารถวัดผลในระดับสูงของความสามารถทางสมองที่ได้แก่ ความคิดริเริ่ม ความสามารถในการวางแผน หรือ การมีดุลพินิจอันสุขุมได้

ตัวอย่างสิ่งที่น่าสนใจใช้วัดเชาวน์ บุคคลต่างๆ ที่ได้กล่าวนามมาแล้วในบทนี้ ต่างพยายามสร้างแบบทดสอบมาตรฐานเพื่อวัดเชาวน์ตามแนวความคิดของตน ในบรรดา

⁶ Guion, Robert M., *Personnel Testing*, p. 238.
Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, pp. 83-84.

แบบทดสอบเรื่องนั้นนี้อาจนำมายกตัวอย่างเพื่อให้เห็นว่า ถ้าจะวัดเชาวน์แล้ว จะวัดกันอย่างไร ดังต่อไปนี้⁷

ในเรื่องศัพท์ (vocabulary)

1. จงเลือกคำที่มีความหมายใกล้เคียงที่สุดกับคำว่า ลำสัน (คำตอบที่ถูกต้องคือคำที่พิมพ์ตัวหนัก)

ก. ร่าเริง

ข. แข็งแรง

ค. อ้วน

ง. เล็ก

จ. รวย

2. ในเรื่องความหมายในภาษา (verbal analogies)

กิ่งไม้มีความสัมพันธ์กับต้นไม้ เหมือนกับลำธารมีความสัมพันธ์กับ

ก. น้ำ

ข. ราก

ค. ริมฝั่งน้ำ

ง. แม่น้ำ

จ. การแตกกระจายของน้ำ

3. ในเรื่องการทำเหตุผลทางคณิตศาสตร์ (arithmetic reasoning)

เด็กชายคนหนึ่งซื้อลูกอมมากล่องละ 90 สตางค์ ในกล่องมีลูกอม 24 เม็ด เขานำลูกอมไปขายเม็ดละ 5 สตางค์ ถ้าถามว่าเขาได้กำไรเม็ดละกี่สตางค์

ก. 30 สตางค์

ข. $3 \frac{3}{4}$ สตางค์

ค. $1 \frac{1}{4}$ สตางค์

ง. $\frac{4}{4}$ สตางค์

จ. ไม่มีข้อใดถูก

4. ในเรื่องการเรียงลำดับตัวเลข (number series)

เลขอะไรควรต่อจากลำดับเลขที่จัดไว้แล้วดังนี้ คือ

1 2 4 7 11

ก. 14

ข. 15

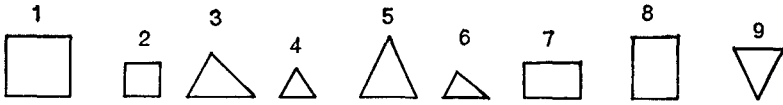
ค. 16

ง. 18

จ. 22

⁷ Thorndike and Hagen, *op. cit.*, pp. 205–207.

ความหมายของรูป



รูปที่ 1 เทียบกับรูปที่ 2 ได้เหมือนกับรูปที่ 3 เทียบกับรูปที่ 4, 5, 6, 7, 8, 9

แนวคิดเรื่องลักษณะต่าง ๆ ของแบบทดสอบ แบบทดสอบที่ได้สร้างขึ้นเป็นเครื่องมือวัดเชาวน์มีลักษณะที่ต่าง ๆ กันไป อันควรรำนำมากล่าวไว้ในที่นี้คือ ประการแรก แบบทดสอบนี้ทั้งใช้ภาษา (verbal tests) และที่ไม่ใช้ภาษา (non-verbal tests) ประการที่สอง มีแบบทดสอบบางชุดที่ปลอดจากอิทธิพลของวัฒนธรรม (culture-free tests) และประการที่สาม แบบทดสอบเชาวน์ อาจแยกประเภทตามลักษณะของการควบคุมการทดสอบก็ได้ คือ แยกประเภทเป็น แบบทดสอบกลุ่ม (group tests) และแบบทดสอบรายตัว (individual tests)⁸

โดยทั่วไปแบบทดสอบก็สร้างขึ้นเป็นข้อเขียนโดยประสงค์ให้ผู้เข้าสอบตอบคำถามที่ปรากฏบนกระดาษ แต่เนื่องจากมีบุคคลบางกลุ่มที่จะเสียเปรียบในการตอบแบบทดสอบ เพราะมีความชำนาญทางภาษาดีกว่าบุคคลทั่วไป เช่น ชาวต่างประเทศที่ไม่ได้ใช้ภาษาเดียวกับแบบทดสอบ อีกประการหนึ่งในบรรดาบุคคลที่ยังเยาว์วัยอยู่ต่างก็มีอัตราความเร็วในการเรียนภาษาไม่เท่ากัน แบบทดสอบที่เป็นข้อเขียนจึงทำให้บุคคลเหล่านี้บางคนต้องเสียเปรียบในการตอบ ทำให้ค่าคะแนนอาจต่ำกว่าที่ควรได้ตามระดับของเชาวน์ที่มีอยู่จริง จึงได้มีการสร้างแบบทดสอบเชาวน์ที่ไม่ต้องมีความรู้ในเรื่องภาษา (non-language tests) เช่น Lorge-Thorndike Intelligence Test, Non-verbal Series และ Pintner Non-language Test คำถามในแบบทดสอบประเภทนี้เป็นปัญหาเกี่ยวกับความหมายของรูปหรือลำดับตัวเลข

นอกจากปัญหาในการที่ความสามารถทางภาษาทำให้ไม่ได้ค่าคะแนนที่แท้จริงของเชาวน์แล้ว ยังมีปัญหาในเรื่องเกี่ยวกับอิทธิพลของวัฒนธรรมที่ทำให้ผู้ตอบแบบทดสอบบางกลุ่มต้องได้คะแนนต่ำกว่าที่ควรได้ตามระดับของเชาวน์ที่มีอยู่จริง กล่าวคือ เนื้อหาของแบบทดสอบตามธรรมเนียมสะสมสะท้อนสภาพการณ์ทางวัฒนธรรม ซึ่งทำให้บุคคลทั่วไปไม่มีความคุ้นเคยเพราะมีประสบการณ์จากการครองชีวิตในสังคมหนึ่ง ๆ แต่ทำให้บุคคล

⁸ Guion, *op. cit.*, pp. 491—493.

ที่ไม่เคยใช้ชีวิตในสังคมเดียวกันนี้มีความลำบากในการที่จะเข้าใจ เช่น คำถามที่เกี่ยวกับหิมะย่อมทำให้บุคคลในประเทศไทยทั่ว ๆ ไปตอบได้ยากเพราะไม่เคยเห็น หรือ คำถามที่เกี่ยวกับต้นหมาก ก็อาจทำให้ชาวอเมริกันไม่มีมโนภาพในการนึกตอบเพราะไม่เคยเห็นเช่นเดียวกัน ภาษาก็เป็นส่วนหนึ่งของวัฒนธรรม แบบทดสอบที่เป็นข้อเขียนก็มีส่วนที่ทำให้ปัญหาในเรื่องวัฒนธรรมมาบิดเบือนค่าคะแนนของผู้ตอบแบบทดสอบ ฉะนั้น จึงได้มีผู้พยายามสร้างแบบทดสอบที่ปลอดจากอิทธิพลของวัฒนธรรม (culture-free tests) แต่ก็การสร้างแบบทดสอบที่ปลอดจากอิทธิพลของวัฒนธรรมจริง ๆ ทำได้ยาก จึงมีชื่อเรียกแบบทดสอบที่พยายามขจัดแต่ไม่ปลอดจากอิทธิพลของวัฒนธรรมเสียทีเดียวว่า แบบทดสอบที่ไม่ลำเอียงในแง่วัฒนธรรม (culture-fair tests) เช่น Cattell Culture Free Intelligence Test และ Miller Analogies Test คำถามในแบบทดสอบประเภทนี้มักเป็นรูปภาพ

แบบทดสอบกลุ่ม (group tests) ต่างกับแบบทดสอบรายตัว (individual tests) ประการใหญ่ ๆ คือ ประการแรก คำถามในแบบทดสอบกลุ่มเป็นข้อเขียนเหมือนข้อสอบทั่ว ๆ ไป ผู้ตอบแบบทดสอบต้องอ่านคำถามเอง แต่แบบทดสอบรายตัวนั้น ผู้ควบคุมการสอบอ่านให้ผู้ตอบแบบทดสอบฟัง ประการที่สอง ปัญหาทั้งหมดในแบบทดสอบกลุ่มกำหนดเวลารวม ๆ ที่จะต้องตอบไว้โดยไม่จำกัดว่า ผู้ตอบแบบทดสอบคนใดจะใช้เวลาจะตอบปัญหาแต่ละข้อเป็นปริมาณเท่า ๆ กัน ส่วนแบบทดสอบรายตัว ผู้ตอบมีเวลาตอบปัญหาเรียงไปที่ละข้อโดยไม่จำกัดเวลา และประการสุดท้าย แบบทดสอบกลุ่มมักจะกำหนดคำตอบไว้ให้ผู้ตอบเลือก ส่วนผู้ตอบแบบทดสอบรายตัวนั้นจะคิดคำตอบอย่างไรก็ได้ ความแตกต่างเช่นนี้หมายความว่า ผู้ตอบแบบทดสอบรายตัวไม่จำเป็นต้องใช้ความสามารถในการอ่าน แต่ผู้ควบคุมการทดสอบแบบนี้ก็ต้องใช้เวลามากกว่าการทดสอบแบบกลุ่ม อย่างไรก็ตาม การที่ผู้ตอบสามารถเลือกคำตอบอย่างไรก็ได้ ทำให้การทดสอบช่วยให้ทราบลักษณะของพฤติกรรมต่าง ๆ ได้กว้างขวางกว่าการทดสอบแบบกลุ่ม แบบทดสอบรายตัวจึงมีประโยชน์ในการทดสอบบุคคลรุ่นเยาว์มากกว่าบุคคลที่มีวัยสูงแล้ว

Wechsler Adult Intelligence Scale (WAIS)⁹ แบบทดสอบนี้เป็นแบบทดสอบเขาวนั้ที่มีชื่อเสียงมากจึงนำมากล่าวถึงเป็นตัวอย่างไว้ในที่นี้ แบบทดสอบฉบับนี้ปรับปรุงแก้ไขเสร็จเมื่อปี 1955 มี 11 ภาค จากผลการวิจัยโดยแบบทดสอบนี้ปรากฏว่า สามารถวัด

⁹ *Ibid.*, pp. 491 – 493.

องค์ประกอบได้ 3 ตัว คือ ความเข้าใจภาษา (verbal comprehension) ซึ่งได้แก่ ศัพท์ ความเข้าใจข่าวสาร และเรื่องต่างๆ ในภาษา องค์ประกอบตัวที่สอง คือ ผลรวมของ สัญชาตญาณ (perceptual organization) ซึ่งได้แก่ ความสามารถในการออกแบบ การประมวลสิ่งของ และการจัดรูปภาพ องค์ประกอบตัวที่สาม คือ การจำ (memory) ซึ่งได้แก่ ตัวเลขและคณิตศาสตร์ต่างๆ แต่ถึงแม้ว่าแบบทดสอบนี้เป็นที่ยกย่องกันมากว่าสามารถวัดเชาวน์ได้ดี แต่ก็ไม่ค่อยเหมาะในการวัดผลสำหรับงานบุคคล เพราะแบบทดสอบฉบับนี้เป็นแบบทดสอบรายตัว จึงต้องเปลืองเวลามากในการควบคุมการสอบ ในปี 1956 ได้มีการทดลองใช้แบบทดสอบนี้ในงานบุคคล ปรากฏว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าคะแนนของผู้ตอบแบบทดสอบนี้กับค่าคะแนนผลการปฏิบัติงานของบุคคลกลุ่มเดียวกันอยู่ในระดับค่อนข้างต่ำ สาเหตุสำคัญของแบบทดสอบฉบับนี้ในแต่ละภาคมีดังต่อไปนี้

1. ข่าวสารทั่วไป (general information) คำถามที่ให้ผู้ตอบเติมคำในช่องว่างเกี่ยวกับเรื่องราวทั่วไป ที่ทุกคนอาจพบเห็นได้
2. ความเข้าใจ (comprehension) มีปัญหา 10 รายการ ถามถึงเหตุผลที่จำเป็นต้องมีกฎเกณฑ์ต่างๆ ทางสังคมเพื่อแก้ไขปัญหาต่างๆ ไปในสังคม
3. การหาเหตุผลทางคณิตศาสตร์ (arithmetic reasoning) เป็นคำถามปากเปล่า ผู้ตอบจะต้องตอบให้ทันเวลา
4. ตัวเลขต่างๆ (digit span) ผู้เข้าสอบฟังตัวเลขหลายๆ ตัว และว่าตามให้ถูกต้อง มีตัวเลขตั้งแต่สามตัวถึงเก้าตัว
5. ความคล้าย (similarities) แต่ละคำถามมีคำเป็นคู่ ผู้เข้าสอบต้องอธิบายลักษณะที่คล้ายกันของแต่ละคู่
6. ศัพท์ (vocabulary) ผู้เข้าสอบต้องนิยามศัพท์ 42 ตัว มีทั้งศัพท์ที่ใช้ทั่วไปประจำวัน และศัพท์ที่ยากๆ
7. การเติมรูปให้สมบูรณ์ (picture completion) ผู้เข้าสอบต้องชี้ว่า แต่ละภาพในปัญหาข้างขาดส่วนใดบ้าง
8. การจัดรูป (picture arrangement) มีรูปภาพ 8 ชุด ที่ต้องเรียงให้ได้ความเป็นเรื่องราว
9. การจัดสิ่งของ (object assembly) มีรูปของสิ่งของ 4 อย่าง ที่ตัดไว้เป็นชิ้น เพื่อให้ต่อให้สมบูรณ์ภายในเวลาที่กำหนด

10. การออกแบบ (block design) มีแท่งสีขาวและสีแดงที่จะต้องนำมาต่อให้เป็นรูปแบบ 10 อย่าง

11. เครื่องหมายตัวเลข (digit symbol) กำหนดเครื่องหมาย 9 อย่าง สำหรับตัวเลขแต่ละตัว ผู้เข้าสอบจะต้องนำเอาโจทย์ซึ่งมีเครื่องหมายต่าง ๆ มาแปลให้เป็นตัวเลขตามเกณฑ์ที่กำหนดไว้

Miller Analogies Test (MAT)¹⁰ แบบทดสอบนี้เชื่อกันว่าใช้ได้ผลในการวัดบุคคลได้มีการวิจัยแล้วว่า ค่าคะแนนของผู้เข้าสอบมีสัมประสิทธิ์สหสัมพันธ์กับผลการปฏิบัติงานในหน้าที่ผู้จัดการฝ่ายขาย และเจ้าหน้าที่เทคนิคในห้องปฏิบัติการวิทยาศาสตร์ แบบทดสอบฉบับนี้แบ่งเป็น 4 ชุด แต่ละชุดมีคำถาม 100 ข้อ เป็นปัญหาเกี่ยวกับความรู้ต่าง ๆ อย่างกว้างขวาง เนื้อหาของแบบทดสอบมีสัมประสิทธิ์ความเชื่อมั่นในระดับสูง จากผลการวิจัยปรากฏว่า สัมประสิทธิ์ความเชื่อมั่น ตามวิธีเทียบเครื่องมือวัดนี้กับเครื่องมือวัดอื่น ๆ ได้ค่าระหว่าง 0.85 ถึง 0.89 แต่การที่จะใช้แบบทดสอบนี้วัดบุคคลหน่วยงานที่ประสงค์จะต้องส่งบุคคลไปรับการทดสอบ ณ ที่ทำการของหน่วยการทดสอบที่ได้อนุญาตเฉพาะ (licensed testing centers) เท่านั้น ผลการทดสอบจะแจ้งให้ทราบแต่เฉพาะบริษัทธุรกิจหรือหน่วยงานที่ได้รับการพิจารณาแล้วเท่านั้น

เชาวน์กับการปฏิบัติงาน

หากจะวัดเชาวน์ให้ได้ประโยชน์เต็มที่ในการบริหารงานบุคคล ก็จะต้องมีเครื่องมือวัดที่ช่วยให้ได้ค่าคะแนนที่สอดคล้องกับความสำเร็จในการปฏิบัติงานตามทฤษฎีของ Guilford สิ่ง que แสดงว่าบุคคลหนึ่ง ๆ มีเชาวน์ในระดับสูงที่จะเป็นประโยชน์ต่อการปฏิบัติงานแยกได้เป็น 3 ประการ คือ ความสร้างสรรค์ (creativity) ความสามารถแก้ปัญหา (problem solving) และความสามารถในการวางแผน (planning)¹¹

ความสร้างสรรค์ การที่กล่าวว่า เชาวน์มีความสำคัญในแง่ที่ช่วยให้บุคคลมีความสามารถในการสร้างสรรค์ในการปฏิบัติงานนั้น ก็เพราะว่า บุคคลที่มีความสามารถที่จะคิดในเรื่องต่าง ๆ ได้แตกแขนงออกไป (divergent thinking ability) ย่อมมีโอกาสที่จะผลิตผลงานในแง่การสร้างสรรค์มากกว่าผู้อื่น ความจริงแล้วความสร้างสรรค์ของบุคคลในการปฏิบัติงานจะมีขึ้นได้โดยอาศัยแนว 3 ประการ คือ ความคล่อง ความยืดหยุ่น

¹⁰ *Ibid.*, pp. 227–228.

¹¹ Mackinnon, Donald W., "The Nature and Nurture of Creative Talent," in Jackson and Mesick, eds., *op. cit.*, pp. 491–493.

และความเป็นตัวของตัวเอง อย่างไรก็ตาม ในเรื่องนี้ได้มีผู้ทำวิจัยแล้ว แต่ผลงานวิจัยก็ยังเป็นที่อภิปรายกันอยู่ คือ M.S. Allen และคณะได้พิมพ์ผลงานวิจัยใน ค.ศ. 1960 กล่าวถึงผลการสำรวจความเห็นของนักวิจัยวิทยาศาสตร์ธรรมชาติจำนวนหนึ่ง ปรากฏว่านักวิจัยวิทยาศาสตร์ที่ถูกสำรวจลงความเห็นว่างค์ประกอบที่สำคัญที่สุดในเรื่องการสร้างสรรค คือ ความยืดหยุ่น ความหยั่งรู้ (penetration) ความสามารถในการหาเหตุผลทั่วไปและความสามารถในการนิยาม นอกจากนี้ก็มีองค์ประกอบอีก 24 ตัว ที่มีความสำคัญรองลงไป การวิจัยเรื่องนี้ ทำให้นักวิจัยบางคนลงความเห็นได้ว่าการที่จะทดสอบเขาวนเพื่อให้ทราบว่า ผู้ปฏิบัติงานจะมีความสามารถในการสร้างสรรคหรือไม่นั้น จะต้องเป็นการทดสอบเขาวนในลักษณะทั่วไป ไม่ใช่เป็นการวัดความฉลาดในเรื่องใดเรื่องหนึ่งโดยเฉพาะ เช่น ความฉลาดในเรื่องการใช้ภาษา หรือความฉลาดในเรื่องการคำนวณ

ความสามารถแก้ปัญหา ในปี ค.ศ. 1962 P.R. Merrifield และคณะได้พิมพ์ผลงานวิจัยที่ศึกษาว่า การที่บุคคลจะมีความสามารถแก้ปัญหาานั้นจะต้องมีความสามารถในด้านใดบ้าง ปรากฏว่า ผลงานวิจัยเรื่องนี้พิสูจน์ได้ว่า จะต้องมีความสามารถ 4 ประการ คือ ความสามารถในการคิดหยั่งรู้เหตุการณ์ล่วงหน้า ความเป็นตัวของตัวเอง และความสามารถที่จะรู้จักปัญหา ความสามารถสองประการแรกนับได้ว่าเป็นความสามารถเกี่ยวกับการเรียนรู้ ความสามารถประการที่สาม เป็นความสามารถในการที่จะคิดเรื่องต่าง ๆ ให้แตกแขนงออกไป และความสามารถประการสุดท้าย เป็นความสามารถในแง่การประเมินค่า อย่างไรก็ตาม ผลงานวิจัยเรื่องนี้ยังนับไม่ได้ว่ามีค่าสมบูรณ์ในการวิเคราะห์ความสามารถในเรื่องนี้ เพราะเครื่องมือที่ใช้ในการวิจัยเป็นเอกสารที่ใช้ภาษาเขียนอธิบายให้ผู้ถูกวิจัยเข้าใจข้อเท็จจริงและนโยบายต่าง ๆ ของการแก้ปัญหา ทำให้ผู้ถูกวิจัยไม่จำเป็นต้องแสดงพฤติกรรมในด้านการเตรียมการ และในด้านการวิเคราะห์

ความสามารถในการวางแผน งานส่วนใหญ่จำเป็นต้องมีความสามารถในการวางแผน อย่างไรก็ตาม ในเรื่องนี้ก็ยังต้องการผู้ที่จะทำการวิจัยให้ชัดเจนลงไปว่า บุคคลจะต้องมีความสามารถอะไรบ้าง จึงจะเรียกได้ว่ามีความสามารถในการวางแผน ในขั้นนี้เป็นที่เชื่อกันว่า ผู้ที่จะมีความสามารถในการวางแผนก็คือ ผู้ที่มีความสามารถในการคิดหยั่งรู้เหตุการณ์ล่วงหน้า มีความสามารถที่จะจัดลำดับ และมีความสามารถที่จะคิดขยายความต่าง ๆ ได้

หากพิจารณาตามแนวที่กล่าวมานี้ เครื่องมือวัดที่มีอยู่ในปัจจุบันยังไม่สามารถที่จะวัดเขาวนในลักษณะที่ครบถ้วนได้ กล่าวได้ว่า เครื่องมือวัดเขาวนต่าง ๆ วัดองค์ประกอบ

บางตัวของเขาวนได้ เช่น ความเข้าใจในเรื่องภาษา การแยกแนวคิดในภาษา การสรุปความสัมพันธ์ของภาษา องค์ประกอบในเรื่องความจำบางตัว ส่วนที่จะวัดว่าบุคคลมีความคิดริเริ่ม หรือมีความสามารถในการวางแผนอย่างใดนั้น ยังทำไม่ได้

ถึงแม้เครื่องมือวัดเขาวนจะยังขาดความสมบูรณ์ตามแนวดังกล่าว ลักษณะของเขาวนที่วัดได้ตามเครื่องมือวัดที่มีอยู่ในปัจจุบัน ก็มีความสัมพันธ์กับความสำเร็จในการปฏิบัติงานในระดับที่น่าพอใจ โดยเฉพาะในเรื่องการคัดเลือกคนเพื่อรับการฝึกอบรมก่อนการปฏิบัติงาน

เมื่อเทียบบุคคลในอาชีพต่างๆ แล้ว ปรากฏว่า บุคคลที่มีอาชีพต่างๆ กัน มีระดับของเขาวนต่างกัน จากผลการวิจัยทหารเกณฑ์ในสหรัฐอเมริกาในสมัยสงครามโลกครั้งที่ 2 ปรากฏว่า อาชีพต่างๆ มีส่วนสัมพันธ์กับเขาวน ดังแสดงไว้ในตารางที่ 26

ตารางที่ 26*

ความสัมพันธ์ระหว่างเขาวนกับอาชีพ

อาชีพ	เปอร์เซ็นต์	10	25	50	75	90
นักบัญชี		114	121	129	136	143
ครู		110	117	124	132	140
ทนายความ		112	118	124	132	141
พนักงานบัญชี		108	114	122	129	138
หัวหน้าเสมียน		107	114	122	131	141
ช่างเขียนแบบ		99	109	120	127	137
เสมียนไปรษณีย์		100	109	119	126	136
ช่างแก้ววิทย์		97	108	117	125	136
พนักงานเดินตลาด		94	107	115	125	133
ผู้จัดการร้าน		91	104	115	124	133
ช่างทำเครื่องมือ		92	101	112	123	129
ช่างประจำเครื่องกล		86	99	110	120	127
ตำรวจ		86	96	109	118	128
ช่างไฟฟ้า		83	96	109	118	124

อาชีพ	เปอร์เซ็นต์	10	25	50	75	90
คนฆ่าเหาะเนื้อ		80	94	108	117	126
ช่างประจำเครื่องมือกล		77	89	103	114	123
ช่างเครื่องรถยนต์		75	89	102	114	122
ช่างไม้ทั่วไป		73	86	101	113	123
คนทำขนมปัง		69	83	99	113	123
คนขับรถบรรทุกขนาดใหญ่		71	83	98	111	120
กรรมกรทั่วไป		65	76	93	108	119
ช่างตัดผม		66	79	93	109	120
กรรมกรเหมืองแร่		67	75	87	103	119
ลูกจ้างทำนา		61	70	86	103	115
คนตัดไม้ประจำโรงเลื่อย		60	70	85	100	116

* แหล่งที่มา : Thorndike, Robert L., and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education*, p. 233.

ผลการวิจัยในตารางที่ 26 เป็นค่าเปอร์เซ็นต์ที่ได้จากการทดลองใช้แบบทดสอบ Army General Classification Test เมื่อเทียบระหว่างอาชีพพนักงานบัญชีหรือครู กับอาชีพกรรมกร หรือลูกจ้างทำนา จะเห็นว่ามีความแตกต่างกันอย่างเห็นได้ชัด องค์กรใดที่บุคคลที่ดีที่สุดของอาชีพที่รอง ๆ ลงมาก็มีระดับเขาวนที่ไม่แพ้บุคคลในอาชีพสูง ๆ ขึ้นไป หรืออาจจะดีกว่าด้วยซ้ำ เช่น ช่างไม้ หรือคนขับรถบรรทุกขนาดใหญ่ที่ดีที่สุด 10% แรก มีเขาวนในระดับเดียวกับพนักงานบัญชี หรือช่างเขียนแบบที่อยู่ในระดับปานกลาง และมีเขาวนสูงกว่าทนายความหรือครูที่อยู่ในประเภท 10% ที่ย่ำสุดเสียอีก

Ghiselli และ Brown ได้ทำการค้นคว้าความเที่ยงตรงของเครื่องมือวัดเขาวนชุดต่าง ๆ ในแง่การคาดคะเน (predictive validity) ปรากฏผลว่า เครื่องมือวัดเขาวนสามารถคาดคะเนในเรื่องการคัดเลือกบุคคลเพื่อรับการฝึกอบรมก่อน การปฏิบัติงานดีกว่าในเรื่องผลสำเร็จในการปฏิบัติงาน สรุปผลการค้นคว้าในเรื่องนี้แสดงไว้ในตารางที่ 27

ตารางที่ 27*

เกณฑ์เฉลี่ยของความสามารถในการคะแนนของแบบทดสอบชาวน์

อาชีพ	เกณฑ์เฉลี่ยของสัมประสิทธิ์ ความเที่ยงตรง	การฝึกอบรม	ความสำเร็จ ในงาน
ช่างประจำเครื่องมือกล		—	.16
ช่างประกอบเครื่อง		.02	.22
นายตรวจงาน		.19	.35
พนักงานห่อ		.22	.13
ช่างซ่อมเครื่องกล		.38	.04
พนักงานเครื่องไฟฟ้า		.43	.47
ช่างประจำเครื่องมือกลขั้นสูง		.34	.28
พนักงานเครื่องกล		.30	.08
ผู้ปฏิบัติงานในหน้าที่บำรุงรักษา		.46	.26
ผู้ปฏิบัติงานในด้านบริการ		.50	.07
พนักงานขับรถ		.18	.14
พนักงานขาย		—	.32
เสมียนขาย		—	.10
เสมียนเครื่องจักรกลคณนา		.23	.16
เจ้าหน้าที่ฝ่ายจัดการ		—	.37
หัวหน้างานชั้นต้น		—	.27

* แหล่งที่มา : Guion, Robert M., *Personnel Testing*, p. 236.

จากตารางที่ 27 จะเห็นว่า ค่าในช่องการฝึกอบรมมีแนวโน้มสูงกว่าในช่องความสำเร็จของงาน ค่านี้คือสัมประสิทธิ์ความเที่ยงตรงในแง่คาดคะเน จำนวนที่ได้โดยนำค่าคะแนนของการวัดชาวน์โดยแบบทดสอบต่าง ๆ ไปเทียบกับค่าที่ประเมินเพื่อรับการฝึกอบรม และค่าที่ประเมินผลงานตามลำดับ แล้วนำมาเฉลี่ยเฉพาะในตำแหน่งในฝ่ายจัดการและพนักงานขายนั้น จะเห็นว่า ยังลงความเห็นเป็นที่ยุติไม่ได้ว่า ชาวน์ที่วัดได้

โดยแบบทดสอบต่าง ๆ นั้นจะช่วยในการคาดคะเนความสำเร็จในการปฏิบัติงานได้ตามสมควร ทั้งนี้จะสังเกตได้จากสัมประสิทธิ์ .37 และ .32 ตามลำดับ

จากข้อเท็จจริงดังกล่าวนี้ Guion จึงสรุปว่า เนื่องจากการที่จะปฏิบัติงานในหน้าที่ต่าง ๆ ได้นั้น บุคคลจะต้องมีความสามารถทางสมองที่จะเรียนรู้สิ่งต่าง ๆ ที่เกี่ยวข้องกับงาน ฉะนั้นแม้ว่าแบบทดสอบต่าง ๆ จะยังไม่สามารถวัด “ตัวเชาวน์” อย่างแท้จริง แบบทดสอบเหล่านี้ก็ช่วยในการคาดคะเนในงานบุคคลอย่างน่าพอใจ กล่าวคือนอกจากจะช่วยในการพิจารณาว่าควรรับบุคคลใดเข้าฝึกอบรมเพื่อปฏิบัติงานต่อไป บุคคลใดจะปฏิบัติงานได้สำเร็จ บุคคลใดจะปฏิบัติงานได้นานเท่าใด และบุคคลใดสมควรได้รับการเลื่อนตำแหน่ง เพราะการที่บุคคลจะฝึกอบรมด้วยดี หรือปฏิบัติงานด้วยดี หรือจะมีกำลังใจปฏิบัติงานต่อไป หรือจะสามารถปฏิบัติงานในตำแหน่งสูงขึ้นไป จะต้องอาศัยความสามารถทางสมอง แบบทดสอบต่าง ๆ ที่มีอยู่ในปัจจุบันก็สามารถวัดผลของการใช้ความสามารถทางสมองของบุคคลได้กว้างขวางพอสมควร

บทที่ 9

ความถนัดพิเศษและทักษะทางกาย

นอกจากสติปัญญาแล้ว บุคคลยังต้องอาศัยความสามารถพิเศษบางอย่างเพื่อให้สามารถปฏิบัติงานหนึ่ง ๆ ได้ ในการวัดความสามารถที่จะทำงานแต่ละอย่างโดยเฉพาะจะต้องวิเคราะห์แนวปฏิบัติของผู้ปฏิบัติงานจริง ๆ เพื่อให้ทราบว่า การปฏิบัติงานนั้น ๆ ต้องทำอะไร อย่างไร และมีปัญหาอะไรบ้าง แล้วนำมาตั้งสมมติฐาน เพื่อสร้างเครื่องมือวัดความสามารถพิเศษเฉพาะเหล่านั้น ความสามารถนั้นแยกได้เป็น 2 ประการ คือ ความถนัดพิเศษ (special aptitudes) และ ทักษะทางกาย (physical skills)

ความถนัดพิเศษที่ได้พยายามวัดกันในงานบุคคลที่จะนำมากล่าวในที่นี้มีอยู่ 2 อย่าง คือ ความถนัดในงานสารบรรณ และความถนัดด้านเครื่องจักรกล ส่วนทักษะทางกายนั้นจะกล่าวถึงความคล่องของกล้ามเนื้อ และความสามารถของประสาทตา

การที่บุคคลต้องอาศัยความสามารถพิเศษประจำวันนอกเหนือไปจากสติปัญญาในการปฏิบัติงานให้ได้ผลดีนั้น อาจเป็นข้อสงสัยสำหรับบางคนที่จะเห็นว่า สติปัญญาย่อมเป็นหลักในการช่วยให้ปฏิบัติงานต่าง ๆ ได้ เรื่องความสามารถพิเศษประจำวัน ถ้าพิจารณาเฉพาะด้านทักษะทางกาย ก็ยังอาจไม่เห็นเป็นข้อคัดค้านทีเดียว เพราะโดยสามัญสำนึกก็อาจเห็นได้ว่า ในการปฏิบัติงานต่าง ๆ ใช้ความแคล่วคล่องว่องไวในเรื่องประสาทและส่วนต่าง ๆ ของร่างกายไม่เหมือนกัน ลักษณะการใช้สายตาของพนักงานคุมเครื่องจักรย่อมต่างกับภารโรงดูแลความสะอาด หรือพนักงานแบกของ ความเคลื่อนไหวของกล้ามเนื้อของพนักงานคุมปั้นจั่น ย่อมต่างกับพนักงานพิมพ์ดีด เหล่านี้เป็นต้น แต่เมื่อพิจารณาในแง่ความถนัดพิเศษ ก็อาจจะเห็นเป็นข้อโต้แย้งสำหรับบางคนได้ว่า ความถนัดพิเศษอันใดก็ตามเป็นเรื่องของสติปัญญาทั้งสิ้น

แท้ที่จริงการที่กล่าวว่า ความถนัดพิเศษเป็นเรื่องของสติปัญญา ก็มีส่วนถูกอยู่มาก กล่าวคือ การที่มีชื่อเรียกความถนัดพิเศษต่างหากไปจากสติปัญญาหรือเชาวน์ก็เพราะนักวัดบุคคลได้พยายามสร้างเครื่องมือวัดความสามารถซึ่งเชื่อกันว่ามีเฉพาะผู้ปฏิบัติงานในตำแหน่งหนึ่ง ๆ ไม่ซ้ำกับตำแหน่งอื่น ทั้งนี้เพื่อที่จะวัดให้ได้ว่าผู้ใดจะปฏิบัติงานได้ดีต่อไป ความถนัดพิเศษ จึงถูกตั้งชื่อเพื่อเรียกลักษณะนิสัย (trait) เฉพาะอย่างหนึ่งในการปฏิบัติงานของบุคคล ต่อมาเมื่อมีการศึกษาการวิเคราะห์องค์ประกอบ (factor analysis) อย่างกว้างขวาง และได้มีการนำเอาแบบทดสอบทั้งหมดมาแยกองค์ประกอบก็ทำให้ทราบว่า ความถนัดพิเศษทั้งในเรื่องงานสารบรรณ และงานเครื่องจักรกลประกอบด้วยสติปัญญา หรือความสามารถทางสมองอยู่ไม่น้อย¹ อย่างไรก็ตามแบบทดสอบที่ได้สร้างขึ้นไว้แล้วเพื่อการวัดความสามารถพิเศษประจำตัวนี้ ก็ยังใช้กันอยู่และก็ยังไม่มีเครื่องมือวัดความสามารถทางสมองทุก ๆ ด้านครบตามองค์ประกอบทุกตัว จึงยังมีความจำเป็นต้องศึกษาลักษณะของเครื่องมือวัดเหล่านี้ ในแง่การวัดความสามารถเฉพาะตัวต่อไป

ความถนัดในงานสารบรรณ

ความสามารถนี้แยกเป็นองค์ประกอบได้ 2 ประเภท คือ ประเภทภาพ (figural) และประเภทเครื่องหมาย (symbolic) องค์ประกอบ 2 ประเภททำให้เกิดความสามารถที่เรียกกันว่า ความคล่องทางสัญชาน (perceptual speed) ซึ่งนำมาตั้งชื่อใหม่ในวงการวัดผลบุคคลว่า ความถนัดทางงานสารบรรณ (clerical aptitude)

โดยความเป็นจริง ความถนัดทางงานสารบรรณ เป็นชื่อที่ไม่เหมาะสม² เพราะการที่บุคคลจะเรียกได้ว่ามีความถนัดในงานสารบรรณ ควรจะต้องมีความสามารถหลายอย่าง ไม่เพียงแต่มีความคล่องทางสัญชานเท่านั้น อาทิเช่น ต้องมีความรู้ในการสะกดตัวอักษรในการพิมพ์ เป็นต้น ทั้งยังจะต้องมีคุณสมบัติในด้านบุคลิกภาพอีกด้วย เช่น เป็นคนมีความอดทนในการทำงานที่มีรายละเอียดจุกจิก หรือเป็นคนที่มีจักรึกษาความลับของหน่วยงาน เป็นต้น การที่ได้มีการเรียกความสามารถนี้ว่าเป็นความถนัดทางงานสารบรรณ ก็เพราะได้มีการสร้างเครื่องมือวัดกันใช้แพร่หลายก่อนที่ Guilford จะได้วิเคราะห์องค์ประกอบในเรื่องสติปัญญาดังกล่าวไว้ในบทก่อน

¹ Guion, Robert M., *Personnel Testing*, p. 275.

² *Ibid.*, p. 240.

ความคล่องทางสัญชาตญาณเป็นเรื่องของความไวและความแม่นยำ ในเรื่องภาพและเครื่องหมาย เช่น ในการอ่านทานข้อความ ผู้ปฏิบัติงานต้องสามารถค้นหาคำผิดหรือตัวที่เรียงพิมพ์สลับได้โดยไม่ชักช้า จริงอยู่ ข้อความที่อ่านเป็นเรื่องภาษา แต่ผู้ปฏิบัติงานก็มีได้อ่านเพื่อที่จะค้นหาข้อบกพร่องของข้อความในภาษา จึงกล่าวว่างค์ประกอบของความสามารถนี้ได้แก่ภาพและเครื่องหมายเท่านั้น มิได้หมายรวมไปถึงภาษาด้วย เมื่อพิจารณาโดยทั่วไปแล้ว ความสามารถนี้อาจเป็นประโยชน์ต่อการปฏิบัติงานอื่น ๆ ที่มีใช้งานสารบรรณด้วยก็ได้ อย่างไรก็ตาม ในที่นี้จะได้กล่าวถึงแบบทดสอบที่แพร่หลายในด้านนี้ เพื่อให้เห็นว่า การวัดความถนัดในงานสารบรรณนั้น ดำเนินการกันอย่างไร

Minnesota Clerical Test เป็นแบบทดสอบความถนัดในงานสารบรรณที่แพร่หลายที่สุด สร้างในปี 1931 ชื่อเดิมว่า Minnesota Vocational Test for Clerical Workers แบบทดสอบนี้แบ่งเป็น 2 ภาค ภาคแรกเปรียบเทียบตัวเลข ภาคที่สองเปรียบเทียบชื่อ ผู้เข้าทดสอบจะต้องตรวจดูว่า ตัวเลขและชื่อที่ให้มาแต่ละคู่เหมือนกันหรือไม่ ถ้าหากเหมือนกันก็ให้ใส่เครื่องหมายลงไป ถ้าหากต่างก็ปล่อยเว้นไว้ คะแนนตามแบบทดสอบนี้คำนวณโดยเอากิ่งหนึ่งของจำนวนข้อที่ทำผิดไปหักออกจากจำนวนข้อที่ทำถูกทั้งหมด ตัวอย่างของแบบทดสอบนี้มีแสดงในแผนภาพที่ 18

Psychological Corporation General Clerical Test แบ่งเป็น 9 ภาค คะแนนที่คำนวณได้จากแบบทดสอบนี้ แยกได้เป็น 3 ชนิด คือ ความถนัดในงานสารบรรณ ความสามารถทางตัวเลข และความสามารถทางภาษา ภาคแรก เป็นคำถามเกี่ยวกับตัวเลขและชื่อ ผู้เข้าสอบจะต้องเทียบ รายชื่อ รายการที่อยู่ และจำนวนตัวเลขที่เรียงไว้ในครั้งหน้าหลังให้ตรงกับที่เรียงไว้ในครั้งหน้าแรก ภาคที่สอง แสดงชั้นเอกสารที่เรียงตัวอักษรไว้ให้ คำถามจะกำหนดให้ตอบว่า ถ้ามีเอกสารตามรายการจะต้องนำเก็บในช่องใด ภาคที่สาม เป็นโจทย์เลขคณิต ภาคที่สี่ มีโจทย์เลขคณิตและคำถามที่ต้องตอบว่าตัวเลขอันใดผิดบ้าง ภาคที่ห้า มีปัญหาเกี่ยวกับตัวเลขต่างๆ สี่ภาคสุดท้ายเป็นโจทย์เกี่ยวกับการสะกดตัวอักษร การอ่านเพื่อความเข้าใจ ศัพท์ และปัญหาทางไวยากรณ์ สองภาคแรกนั้นนับว่าเป็นการวัดความถนัดในงานสารบรรณโดยเฉพาะ ภาคสามถึงภาคห้าเป็นการวัดความสามารถทางตัวเลข ส่วนสี่ภาคสุดท้ายเป็นการวัดความสามารถทางภาษา

ข้อสอบของสำนักงาน ก.พ. สำนักงานคณะกรรมการข้าราชการพลเรือนได้พยายามสร้างแบบทดสอบ เพื่อใช้ในการสอบแข่งขันวิชาข้าราชการพลเรือนชั้นจัตวาและชั้นตรี ถึงแม้ว่าข้อสอบฉบับนี้จะอยู่ในระหว่างการดำเนินการทำให้เป็นมาตรฐาน ตัวอย่างจาก

MINNESOTA CLERICAL TEST

(formerly the Minnesota Vocational Test for Clerical Workers)
by Dorothy M. Andrew, Donald G. Paterson, and Howard P. Longstaff

Name _____ Date _____

TEST 1-Number Comparison

Number Right _____

Number Wrong _____

Score = R-W _____

Percentile Rating _____

Norms Used _____

TEST 2-Name Comparison

Number Right _____

Number Wrong _____

Score = R-W _____

Percentile Rating _____

Norms Used _____

Instructions

On the inside pages there are two tests. One of the tests consists of pairs of names and the other of pairs of numbers. If the two names or the two numbers of a pair are exactly the same make a check mark (✓) on the line between them: if they are different, make no mark on that line. When the examiner says "Stop" draw a line under the last pair at which you have looked.

Samples done correctly of pairs of Numbers

79542 _____ 79524
5794367 _____ 5794367

Samples done correctly of pairs of Names

John C. Linder _____ John C. Lender
Investors Syndicate _____ Investors Syndicate

Now try the samples below.

66273894 _____ 66273984
527384578 _____ 527384578
New York World _____ New York World
Cargill Grain Co _____ Cargill Grain Co

This is a test for Speed and Accuracy. Work as fast as you can without making mistakes. Do not turn this page until you are told to begin.

The Psychological Corporation, New York 17, N.Y.

แผนภาพที่ 18 ตัวอย่าง Minnesota Clerical Test

แหล่งที่มา: Guion, Robert M., *Personnel Testing*, p. 241.

ข้อสอบก็อาจเป็นประโยชน์ต่อผู้ศึกษาวิชานี้โดยตรง เฉพาะในการวัดความถนัดในงาน
สารบรรณนั้น ข้อสอบฉบับนี้ได้พยายามวัดทั้งในด้านความคล่องตัวไวและความ
ละเอียดแม่นยำในการปฏิบัติงานด้านธุรการ และความละเอียดแม่นยำและความรวดเร็ว
ในการปฏิบัติงาน เกี่ยวกับตัวหนังสือและตัวเลข

ตัวอย่างข้อทดสอบความถนัดในงานสารบรรณ³

ทดสอบความแคล่วคล่องว่องไวและความละเอียดแม่นยำในการปฏิบัติงานด้านธุรการ

27. ข้อต่อไปนี้เป็นสมุดโทรศัพท์ ข้อใดจะมาเป็นลำดับที่สอง

ก. ทวี ประยงค์รัตน์

ข. โฆสิต ชาติกำแหง

ค. สุพจน์ ม่วงศิริ

ง. บุญปลูก ชายเกตุ

(“ทวี ประยงค์รัตน์” จะต้องมาเป็นลำดับที่ 2 ดังนั้น ข้อที่ถูกจึงเป็นข้อ ก.)

ทดสอบความละเอียดแม่นยำและความรวดเร็วในการปฏิบัติงานเกี่ยวกับตัวหนังสือและตัวเลข

คำสั่ง ข้อต่อไปนี้เป็นข้อสอบเกี่ยวกับการสังเกต แต่ละข้อจะมีตัวเลขหรือตัวอักษรกำหนดให้ ท่านจงพิจารณาว่าในคำตอบทั้ง 4 นั้น มีคำตอบใดที่มีตัวเลขหรือตัวอักษรไม่เหมือนกับตัวเลขหรือตัวอักษรที่กำหนดให้บ้าง ข้อสอบเหล่านี้ไม่ยาก แต่จะกำหนดเวลาทำเพียงจำกัดที่สุด ซึ่งท่านต้องตอบให้เร็วที่สุดและถูกมากที่สุด

28. 789

ก. 987

ข. 789

ค. 879

ง. 869

(คำตอบที่ถูกคือ ง.)

คำสั่ง ข้อ 28—35 ในแต่ละข้อเป็นข้อความหรือตัวเลขซึ่งแบ่งพิมพ์เป็น 3 แถว ท่านจงตรวจทานว่า ข้อความหรือตัวเลขใดในแถวทั้งสามนั้นเขียนเหมือนกัน หรือต่างกัน แล้วตอบในกระดาษคำตอบ ดังนี้

ตอบข้อ ก. ถ้าข้อความทั้งสามแถวเขียนเหมือนกัน

ตอบข้อ ข. ถ้าข้อความแถวที่ 1 เขียนเหมือนกับแถวที่ 2

ตอบข้อ ค. ถ้าข้อความแถวที่ 1 เขียนเหมือนกับแถวที่ 3

ตอบข้อ ง. ถ้าข้อความแถวที่ 2 เขียนเหมือนกับแถวที่ 3

ตอบข้อ จ. ถ้าข้อความทั้งสามแถวเขียนไม่เหมือนกันเลย

³ สำนักงาน ก.พ., คู่มือการสอบข้อเขียน, เอกสารหมายเลข 3, หน้า 22—24.

แถวที่หนึ่ง	แถวที่สอง	แถวที่สาม
29. กองวิทยาศาสตร์ชีวภาพ	กองวิทยาศาสตร์ชีวภาพ	กองวิทยาศาสตร์ชีวภาพ
30. หน่วยทดลองเส้นใยฝ้าย และไหม	หน่วยทดลองใยฝ้าย และไหม	หน่วยทดลองฝ้าย และ ไหม
31. สำนักงานองค์การน้ำตาล ไทย	สำนักงานกลางองค์การ น้ำตาลไทย	สำนักงานองค์การน้ำตาล ไทย
32. With the compliments of P.R.D.	With the compliments of P.R.D.	With the compliment of P.R.D.
33. Mr. Edwin F. Black	Mr. Edwin E. Black	Mrs. Edwin G. Black
34. 646,278.32	647,278.32	647,278.32
35. 13,167.58	13,167.68	13,167.58

คำตอบสำหรับแต่ละข้อมีดังนี้

- | | |
|--------|--------|
| 29. ก. | 33. จ. |
| 30. จ. | 34. ง. |
| 31. ค. | 35. ค. |
| 32. ข. | |

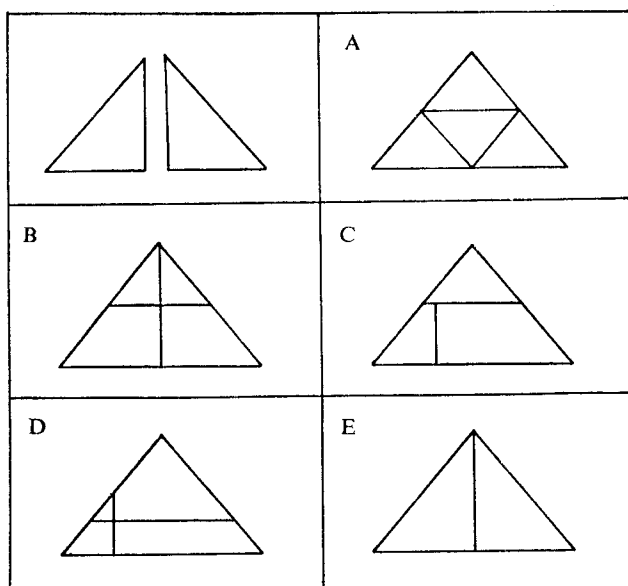
ความถนัดด้านเครื่องจักรกล

เครื่องมือวัดต่าง ๆ ที่พยายามสร้างขึ้นเพื่อวัดว่าคุณจะสามารถปฏิบัติงานเกี่ยวกับเครื่องจักรได้ วัดองค์ประกอบที่เรียกว่า ความสามารถทางมิติ (spatial ability) ความสามารถนี้เป็นความสามารถในการที่จะพิจารณาว่า ภาพหนึ่ง ๆ นั้น ประกอบด้วยส่วนต่าง ๆ อย่างไร ภาพสามมิติเกี่ยวข้องกับภาพสองมิติประการใด วัตถุต่าง ๆ จะมีลักษณะอย่างไรเมื่อมองจากด้านต่าง ๆ กัน และหากวัตถุที่ประกอบด้วยชิ้นส่วนต่าง ๆ เคลื่อนไหวแล้ว ส่วนต่าง ๆ แต่ละชิ้นจะมีอาการเคลื่อนไหวที่สัมพันธ์กันอย่างไร

แบบทดสอบที่วัดความถนัดด้านเครื่องจักรกลในระยะแรก เป็นข้อสอบเกี่ยวกับส่วนต่าง ๆ ของเครื่องจักรในโรงงาน ต่อมาก็ได้มีแบบทดสอบที่ตามความรู้เกี่ยวกับชื่อและการใช้เครื่องมือต่าง ๆ แล้วจึงได้มีผู้สร้างแบบทดสอบที่ครอบคลุมความถนัดในด้านนี้อย่างกว้างขวางออกไป โดยถามหลักการง่าย ๆ ทางเครื่องจักรและตั้งปัญหาเกี่ยวกับสภาพความเกี่ยวข้องของคาน รอก และตัวประกอบต่าง ๆ ทางกลศาสตร์ เช่น มีรูปปล่องอ้อมโค้ง ผู้ตอบจะต้องตอบว่า รถจะไถลไปทางใด หรือมีรูปเขื่อนกั้นน้ำ ผู้ตอบจะ

ต้องอธิบายว่า จุดใดรับน้ำหนักของน้ำมากที่สุด ตัวอย่างของแบบทดสอบความถนัดด้านเครื่องจักรกล ที่ใช้กันแพร่หลายมีดังต่อไปนี้

Revised Minnesota Paper Form Board Test เป็นแบบทดสอบที่แพร่หลายมากในการใช้วัดความถนัดด้านเครื่องจักรกล สร้างขึ้นในปี 1934 แบบทดสอบนี้ประกอบด้วยคำถามที่เป็นรูปภาพ 64 ข้อ ต้องทำให้เสร็จภายใน 20 นาที คำถามแต่ละข้อประกอบด้วยรูปภาพ 6 รูป ผู้เข้าสอบต้องพิจารณาเลือกรูปใดรูปหนึ่งเพียงรูปเดียว ที่มีลักษณะตรงกันกับรูปตัวอย่างที่ให้ไว้แต่ละข้อ รูปที่ให้เลือกลักษณะกลับข้างหรือสลับบนล่างตรงกันข้ามกับรูปตัวอย่างก็ได้ ดังแสดงไว้ในแผนภาพที่ 19



แผนภาพที่ 19 ตัวอย่างแบบทดสอบ MPFB

แหล่งที่มา : Guion, Robert M., *Personnel Testing*, p. 246.

แบบทดสอบนี้เรียกย่อๆ ว่า MPFB ในแผนภาพที่ 19 ซึ่งแสดงตัวอย่างคำถามจะเห็นว่ารูปบนซ้ายสุด เป็นตัวอย่างของรูปในคำถามข้อหนึ่ง ผู้เข้าสอบจะต้องเลือกรูประหว่าง A ถึง E เพียงรูปเดียว โดยเลือกรูปที่ตนคิดว่าเป็นตัวแทนของรูปตัวอย่างที่นำมาเชื่อมกันแล้วถูกต้อง แบบทดสอบนี้จึงมีลักษณะเป็นการวัดความสามารถทางมิติ เพราะผู้เข้าสอบจะต้องใช้ความคิดคำนึงของตนเองในการประกอบชิ้นส่วนต่างๆ จากผลการวิจัยปรากฏว่า แบบทดสอบฉบับนี้ มีสัมประสิทธิ์ความเชื่อมั่นและความเที่ยงตรงสูง เป็นที่นิยมใช้กันกว้างขวางในวงอาชีพ วิศวกร สถาปนิก ช่างเขียนแบบ

และช่างกล แบบทดสอบอื่น ๆ ที่พยายามวัดความถนัดด้านเครื่องจักรกลก็ดำเนินการสร้างคำถาม ตามทำนองเดียวกับแบบทดสอบนี้

Survey of Space Relations Ability ประกอบด้วยคำถามที่มีลักษณะทำนองเดียวกับ MPFB แต่แทนที่จะจัดคำถามเป็นกลุ่ม กลับจัดเรียงไว้ในแถวเดียวกัน ในการตอบคำถามแต่ละข้อ ผู้เข้าสอบจะต้องเลือกรูปภาพรูปหนึ่งรูปใดจากในจำนวน 10 รูป แทนที่จะเป็นเพียง 6 รูป คำถามทั้งหมดมี 35 ข้อ ต้องตอบภายใน 15 นาที แบบทดสอบนี้มีสัมประสิทธิ์สหสัมพันธ์กับ MPFB ระหว่าง 0.69 ถึง 0.77 ซึ่งนับว่าอยู่ในระดับค่อนข้างสูง ส่วนสัมประสิทธิ์ความเชื่อมั่นก็สูงถึง 0.80

แบบทดสอบผสมที่ใช้วัดความถนัดพิเศษกับความเที่ยงตรง

ในการวัดบุคคล เชื่อกันว่า แบบทดสอบผสม (test battery) จะมีคุณสมบัติดีกว่าแบบทดสอบหลายฉบับที่ยกตัวอย่างมาแล้ว เพราะเนื้อหาของสิ่งที่จะวัดได้มีการกำหนดตามแผนของการวัดความถนัดพิเศษทั้งหมด ทำให้สามารถวัดความถนัดพิเศษได้ทุก ๆ ด้านในระยะเวลาสั้นที่สุด แบบทดสอบผสมที่เด่นที่สุดในการใช้วัดความถนัดพิเศษมีอยู่ 2 ฉบับ คือ Differential Aptitude Tests และ General Aptitude Test Battery แบบทดสอบฉบับแรกมีค่าสหสัมพันธ์ระหว่างข้อสอบภาคต่าง ๆ โดยเฉลี่ย 0.42 นับว่าตัวข้อสอบมีความคล้ายคลึงกันในกลุ่ม (homogeneity) ของเนื้อหาเป็นอย่างดีพอสมควร และสัมประสิทธิ์นั้นนับได้ว่าสูงกว่าแบบทดสอบอื่น ๆ ที่ยอมรับใช้กันอีกหลายฉบับ แต่แบบทดสอบนี้ค่อนข้างยาว ต้องใช้เวลาทดสอบถึง 4 ชั่วโมง ฉะนั้นในที่นี้จึงจะกล่าวแต่แบบทดสอบผสมฉบับที่สองเท่านั้น

General Aptitude Test Battery เป็นแบบทดสอบที่ใช้กันแพร่หลายในการบริหารงานบุคคลของรัฐต่าง ๆ ในสหรัฐ ฯ องค์ประกอบที่วัดมี 10 ตัว คือ สถิติปัญญา ความถนัดทางภาษา ความถนัดทางตัวเลข ความสามารถทางมิติ สัญชาตญาณของรูป สัญชาตญาณในงานสารบรรณ ความแม่นยำ ความไวของกล้ามเนื้อ ความคล่องของนิ้ว และความคล่องของมือ คำถามที่ถามเป็นปัญหาเกี่ยวกับการเปรียบเทียบชื่อ การคำนวณ การพิจารณารูปแบบศัพท์ ประเด็นจากเรื่องที่เล่า แบบทดสอบนี้ประกอบด้วยข้อสอบ 15 ภาค กำหนดขั้นต่ำไว้สำหรับผู้ที่ควรจะได้ก่อนที่จะได้รับการพิจารณาให้ปฏิบัติงานตำแหน่งต่าง ๆ แต่งานที่ระบุไว้ส่วนใหญ่เป็นงานขั้นต้นมิใช่เป็นงานในระดับบังคับบัญชา งานเหล่านี้ได้แก่งานประเภทไร่ฝมือ และกึ่งฝมือ มีงานในระดับสูงกว่านี้เพียง 3 ตำแหน่ง คือ หัวหน้าหน่วยบริการจัดงานศพ (funeral director) นักบัญชี และนักตรวจบัญชี จาก

องค์ประกอบที่กล่าวมาข้างต้นจะเห็นได้ว่า มีข้อสงสัยเกี่ยวกับข้อสอบนอ่ยู่บ้าง คือ มีการวัดองค์ประกอบเรื่องสติปัญญา ทั้ง ๆ ที่ข้อสอบนี้มุ่งจะวัดความถนัดพิเศษ⁴ อย่งไรก็ดี แบบทดสอบนี้ใช้กันแพร่หลาย และได้มีการนำผลของการวัดตามแบบทดสอบฉบับนี้มาวิจัยอยู่เรื่อย ๆ

ความเที่ยงตรงของแบบทดสอบความถนัดพิเศษ หลักฐานเชิงสถิติเกี่ยวกับความเที่ยงตรงในการคาดคะเนของแบบทดสอบประเภทนี้ น้อยมาก ทั้งนี้เพราะการที่จะทราบความเที่ยงตรงของแบบทดสอบแต่ละฉบับนั้น ต้องใช้เวลาเป็นสัปดาห์ Ghiselli ได้ประมวลค่าความเที่ยงตรงของแบบทดสอบวัดความถนัดในด้านต่าง ๆ กับอาชีพแต่ละประเภทไว้ ดังแสดงในตารางที่ 28

ตารางที่ 28*

ความตรงประเด้นในการคาดคะเนของแบบทดสอบด้านต่าง ๆ

ตำแหน่งงาน	แบบทดสอบ		
	สติปัญญา	ความสามารถทางมิติ	ความคล่องของกล้ามเนื้อ
หัวหน้างานทั่วไป	53	60	48
หัวหน้าคนงาน	52	49	46
เสมียนทั่วไป	60	34	48
เสมียนจดบันทึก	59	32	53
เสมียนเครื่องจักรกลคณนา	58	56	43
เสมียนฝ่ายขาย	32	50	57
พนักงานขาย	67	64	24
ผู้ปฏิบัติฝ่ายคุ้มครองรักษา	55	40	54
คนงานฝ่ายบริการ	33	59	41
พนักงานขับยานพาหนะ	35	50	69
ช่างแก้ไขเครื่อง	50	53	48
ช่างไฟฟ้า	69	65	56

*แหล่งที่มา : Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, p. 87.

⁴ Guion, *op. cit.*, p. 270.

ความคล่องของกล้ามเนื้อ

ความคล่องของกล้ามเนื้อ (Psychomotor skills) ในแง่การปฏิบัติงาน หมายถึง ความสามารถที่จะประสานกันระหว่างจิตใจและอวัยวะต่างๆ ของร่างกาย (mind-muscle cooperation)⁵ ศูนย์วิจัยงานบุคคลของกองทัพอากาศสหรัฐ ฯ (Air Force Personnel and Training Research Center) ได้ศึกษาเรื่องนี้เป็นพิเศษ และได้แยกความสามารถนี้ออกเป็น 11 ประการ คือ⁶

1. ความไวในการควบคุม (fine control sensitivity) คือ ความสามารถควบคุมกล้ามเนื้อส่วนต่างๆ ที่ไวต่อความรู้สึกได้

2. ความประสานระหว่างแขนขา (multiple limb coordination) คือ ความสามารถเคลื่อนไหวแขนขาพร้อมๆ กัน เพื่อการปฏิบัติงานต่างๆ

3. การควบคุมการสนองตอบ (response orientation) เป็นความสามารถที่จะเลือกการสนองตอบในทางที่ถูก

4. การปรับท่า (postural discrimination) เป็นความสามารถที่จะปรับท่าของร่างกาย โดยไม่ต้องอาศัยการเห็นภาพ

5. ความคล่องของมือ (manual dexterity) เป็นความสามารถที่จะเคลื่อนไหวมือและแขนให้สอดคล้องกับวัตถุที่เกี่ยวข้องกับการปฏิบัติงาน

6. ความคล่องของนิ้ว (finger dexterity) คือ ความคล่องในการใช้นิ้วต่างๆ

7. ความมั่นคงของมือและแขน (arm-hand steadiness) คือ ความสามารถจัดทำของมือและแขนตามที่ต้องการด้วยความแม่นยำและเที่ยงตรง

8. ความเร็วในปฏิกิริยา (reaction time) คือ ความสามารถที่จะมีปฏิกิริยาโต้ตอบต่อสิ่งเร้า (stimulus) ได้อย่างรวดเร็ว

9. ความไวของแขน (speed of arm movement) คือ ความว่องไวในการเคลื่อนไหวแขน

10. ความเร็วในการควบคุม (rate control) ความสามารถในการปรับอวัยวะและท่าของร่างกายให้สอดคล้องกับวัตถุที่เคลื่อนไหวในอัตราความเร็วต่างๆ

11. ความประสานในการสนองตอบ (response integration) คือ ความสามารถที่จะใช้ประโยชน์จากข่าวสารที่ได้รับจากแหล่งต่างๆ ให้สามารถเคลื่อนไหวในลักษณะหนึ่งๆ

⁵ Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, p. 118.

⁶ *Ibid.*, p. 119.

ลักษณะต่าง ๆ ของความสามารถนี้เมื่อพิจารณาในแง่การที่จะใช้ปฏิบัติงานในวงการอุตสาหกรรมและธุรกิจ อาจแยกออกเป็นทักษะต่าง ๆ ได้ 5 ด้าน คือ

1. ความแม่นยำ (aiming) คือ ความชำนาญในการเคลื่อนไหวมือได้อย่างรวดเร็วและแม่นยำเมื่อเห็นภาพหนึ่ง ๆ ในการปฏิบัติงานลักษณะต่าง ๆ
2. ความไวของแขน (speed of arm movement) คือ ความชำนาญในการเคลื่อนไหวก่อนได้อย่างรวดเร็ว
3. ความคล่องของนิ้ว (finger dexterity) คือ ทักษะในการเคลื่อนนิ้วตัวเล็ก ๆ ด้วยนิ้วตามลักษณะที่ต้องการได้อย่างรวดเร็ว
4. ความคล่องของมือ (manual dexterity) คือ ทักษะในการเคลื่อนนิ้วตัวใหญ่ ๆ ด้วยมือและแขนตามลักษณะที่ต้องการได้อย่างรวดเร็ว
5. ความไวของข้อมือและนิ้ว (wrist-finger speed) คือ ทักษะในการเคลื่อนข้อมือในลักษณะหมุนหรือขยับได้อย่างรวดเร็ว

Purdue Pegboard เป็นแบบทดสอบที่ใช้วัสดุจริง ๆ ในโรงงาน วัสดุที่ใช้ทดสอบมีหมุด (pin) วงแหวนแบน (washer) และวงแหวนกลม (collar) ค่าคะแนนที่วัดได้แยกเป็น 4 ส่วน คือ ความสามารถของมือขวา ความสามารถของมือซ้าย ความสามารถของทั้งสองมือพร้อมกัน และความสามารถในการประกอบชิ้นส่วนทั้งหมด เครื่องมือทดสอบเป็นแผ่นไม้ที่มีช่องใส่ หมุด วงแหวนแบน และวงแหวนกลม เป็นส่วน ๆ แยกจากกัน ผู้เข้าสอบมีโอกาสใช้มือซ้ายและมือขวาข้างละ 30 วินาที หยิบหมุดมาใส่ในช่องบนแผ่นไม้และมีเวลาอีกหนึ่งนาทีที่จะใช้ทั้งสองมือพร้อม ๆ กัน หยิบหมุดและวงแหวนมาสวมใส่เข้าด้วยกันเพื่อบรรจุในช่องบนแผ่นไม้ จากผลการวิจัยปรากฏว่าแบบทดสอบนี้มีสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้ตามวิธีทดลองเครื่องมือวัดซ้ำสูงถึงระหว่าง 0.60 ถึง 0.70 ส่วนสัมประสิทธิ์ความเที่ยงตรงนั้นยังต่ำอยู่ ค่าเฉลี่ยของสัมประสิทธิ์ความเที่ยงตรงคำนวณได้ 0.33

ปัญหาการใช้เครื่องมือทดสอบความคล่องของกล้ามเนื้อ การวัดความสามารถด้านนี้ โดยการใช้เครื่องมือจริง ๆ ในการปฏิบัติงาน ทำให้เกิดปัญหาหลายอย่าง ประการแรก การกำหนดเวลาจำกัดในการทำการทดสอบครั้งหนึ่ง ๆ เช่น แบบ Purdue Pegboard อาจทำให้เป็นผลเสียต่อความเชื่อมั่นของเครื่องมือวัด เพราะแบบทดสอบที่วัดความแคล่วคล่องไวของความสามารถประเภทนี้อาจไม่สามารถวัดคุณภาพของการปฏิบัติได้ มี

แต่ Purdue Pegboard เท่านั้นที่แสดงว่ามีค่าความเชื่อมั่นสูง ประการที่สอง เครื่องมือที่ใช้ทำให้มีปัญหारे่องสถานที่ ไม่เหมือนกับข้อสอบที่เขียนตอบในกระดาษ ซึ่งไม่มีปัญหาในการเก็บเครื่องมือทดสอบ ประการที่สาม ความสักรของเครื่องมือ อาจทำให้การทำข้อสอบง่ายขึ้น เพราะช่องหรือช่วงว่างที่ทำไว้กว้างและมีความลึกขึ้น และประการสุดท้าย ค่าใช้จ่ายของเครื่องมือทดสอบสูง เช่น อาจมีราคาถึงชิ้นละ 2,000 บาท เมื่อเทียบกับข้อสอบที่เป็นคำถามในกระดาษ นับว่าแพงกว่ากันมาก

ความสามารถของประสาทตา

การวัดประสาทตาที่มีใช้เป็นการวัดสายตาทั่วไปตามวิธีการทางแพทย์ แต่เป็นการวัดในการปฏิบัติงานหนึ่ง บุคคลจะบังคับกับกลัมตาของตนได้อย่างคล่องแคล่วเพียงใด ความสามารถของประสาทตา (visual skills) ที่จำเป็นต้องมีในการปฏิบัติงานหนึ่ง ๆ อาจแยกพิจารณาได้เป็น 12 ด้าน คือ⁷

ก. ระยะใกล้ (26 ฟุต)

1. ท่าพักของตาในทางตั้ง (vertical phoria)
2. ท่าพักของตาในทางราบ (lateral phoria)
3. ความไวของทั้งสองตา (acuity, both eyes)
4. ความไวของตาขวา
5. ความไวของตาซ้าย
6. ความสามารถมองความลึก (depth perception)
7. ความสามารถเห็นสี (color vision)

ข. ระยะใกล้ (13 นิ้ว)

1. ความไวของทั้งสองตา
2. ความไวของตาขวา
3. ความไวของตาซ้าย
4. ท่าพักของตาในทางตั้ง
5. ท่าพักของตาในทางราบ

⁷ Guion, *op. cit.*, p. 281.

งานต่าง ๆ ในวงการอุตสาหกรรมและธุรกิจ อาจแบ่งเป็นประเภทได้ 6 ประเภทตามระดับของความสามารถของประสาทตา ดังต่อไปนี้ คือ⁸

1. งานเสมียน งานธุรการ และงานในสำนักงานโดยทั่วไป ต้องใช้ความสามารถทุกด้าน ยกเว้นแต่ความสามารถมองลึก แต่ความสามารถในด้านต่าง ๆ ไม่จำเป็นต้องสูงนัก

2. งานตรวจ และงานที่ต้องทำใกล้เครื่องจักร เช่น พนักงานเย็บผ้า หรือพนักงานประกอบชิ้นส่วนวิทยุ ต้องใช้ความไวของตาในระยะใกล้เป็นพิเศษ ในกรณีของผู้ปฏิบัติงานบางหน้าที่ เช่น พนักงานถักถุงเท้าด้วยเครื่องจักร ผู้ที่มีความไวของตาในระยะไกลสูง จะปฏิบัติได้สะดวกเป็นสัดส่วนกับความสามารถที่มี

3. พนักงานขับยานพาหนะ เช่น พนักงานขับรถบรรทุก และยังรวมไปถึงพนักงานควบคุมบินขึ้น และเครื่องมือที่มีลักษณะคล้าย ๆ กัน ความไวของตาในระยะไกลมีความสำคัญที่สุด

4. พนักงานควบคุมเครื่องจักร หมายถึงผู้ที่ปฏิบัติงานกับเครื่องจักรโดยอยู่ห่างจากเครื่องจักรไม่เกินช่วงแขน จำเป็นต้องใช้ความสามารถของประสาทตาทุกด้าน

5. กรรมกรหรือพนักงานแบกหาม ต้องใช้ความไวของตาในระยะไกลสูงกว่างานประเภทที่ 2 ความไวของตาในระยะใกล้จำเป็นแต่เพียงส่วนน้อย ไม่มีปัญหาในเรื่องท่าพักของตาทั้งในทางตั้งและในทางราบเลย

6. ช่างกลและช่างฝีมือต่าง ๆ รวมทั้งช่างแก้ไขทุกชนิด ต้องการความสามารถทุกด้านของประสาทตา โดยเฉพาะความไวในระยะใกล้มีความสำคัญมาก

ปัญหาการวัดความสามารถของประสาทตา ถึงแม้ความสามารถในด้านนี้มีลักษณะพิเศษไปกว่าการวัดความสามารถด้านอื่น ๆ ในเรื่องงานบุคคล ลักษณะของเครื่องมือวัดก็จำเป็นต้องมีความเที่ยงตรงเช่นเดียวกับแบบทดสอบต่าง ๆ แผ่นกระดาษที่มีตัวอักษรซึ่งจักษุแพทย์ใช้ในการวัดสายตานั้นขนาดความเที่ยงตรง โดยเฉพาะในการวัดเพื่องานบุคคล วิธีการของจักษุแพทย์ช่วยให้ทราบแต่เพียงว่าบุคคลสามารถอ่านตัวอักษรได้ในระยะ 20 ฟุตหรือไม่ ระยะที่กำหนดเป็นระยะตายตัวในการวัดมิได้มีการทดลองใช้แสง

⁸ *Ibid.*, pp. 284–285.

ในปริมาณและคุณภาพต่าง ๆ กัน ทำให้ไม่สามารถทราบความสามารถของประสาทตาในแง่งานบุคคลได้

ในปัจจุบันได้มีเครื่องมือวัดความสามารถของประสาทตาเพื่อใช้ในงานบุคคล โดยเฉพาะ 3 แบบ คือ Telebinocular, Sight Screener และ Ortho-Rater เครื่องมือทั้งสามชนิดมีลักษณะเป็นกล่อง 2 ตา ซึ่งผู้เข้าสอบจะต้องส่องมองวัตถุที่กำหนดไว้ กล้องที่ใช้ของวัตถุที่มองเห็น เครื่องมือวัดเหล่านี้จะวัดความสามารถของประสาทตาออกมาเป็น 12 ด้าน ซึ่งได้กล่าวมาแล้วแต่ตอนต้นของหัวข้อนี้ จากการวิจัยได้ผลว่า การให้ผู้เข้าสอบมองวัตถุในกล้อง โดยกำหนดระยะให้ใกล้เคียงตามจะกำหนดนั้น เป็นวิธีการวัดที่มีความเชื่อมั่นสูงกว่าการวัดโดยให้ผู้เข้าสอบมองวัตถุที่อยู่ห่างจากตัวจริง ๆ ตามระยะใกล้เคียงที่กำหนดนั้นเสียอีก

ทฤษฎีการวัดบุคลิกภาพ

แม้นักวิชาการบางคน เช่น Guilford จะถือว่าในการพิจารณาบุคลิกภาพ จะพิจารณาโดยแยกออกจากสติปัญญา หรือความถนัดพิเศษไม่ได้ก็ตาม ในการวัดบุคลิกก็มีหลักฐานชี้ให้เห็นว่า สามารถวัดสติปัญญาหรือความถนัดพิเศษไว้ได้เป็นส่วนหนึ่ง ต่างไปจากการวัดบุคลิกภาพ การวัดสติปัญญาหรือความถนัดพิเศษ เป็นการวัดเพื่อให้ทราบว่า บุคคลหนึ่ง ๆ เมื่ออยู่ในสภาพการณ์และมีแรงกระตุ้นที่จะทำงานให้ดีที่สุด จะทำได้ อย่างไร เพียงใด กล่าวสั้น ๆ คือ ประสงค์จะวัดความสามารถของบุคคล ผู้ถูกวัดทำข้อสอบ หรือตอบคำถามด้วยความเข้าใจว่า ตนมีหน้าที่แก้ปัญหาวางอย่างตามที่กำหนดไว้ตามเครื่องมื่อวัด แต่ในสภาพความเป็นจริงของการทำงาน ทุกคนไม่จำเป็นต้องอยู่ในสภาพการณ์ที่จะได้รับแรงกระตุ้นให้ทำงานอย่างดีที่สุด ฉะนั้น จึงต้องมีการวัดบุคลิกภาพเพื่อให้ทราบว่า บุคคลหนึ่ง ๆ มีปฏิริยาความสนใจ หรือความโน้มเอียงที่จะทำให้เอาใจใส่ในการปฏิบัติงานประเภทหนึ่ง ๆ อย่างไรบ้าง ในการวัดบุคลิกภาพนี้อาจกล่าวได้ว่า เป็นการวัดเพื่อให้ทราบว่าบุคคลหนึ่ง ๆ จะเป็นอย่างไรในช่วงเวลาหนึ่ง ๆ และทำไมจึงเป็นเช่นนั้น ข้อเท็จจริงที่ทราบในเรื่องนี้ช่วยในการวิเคราะห์ว่า การที่บุคคลเป็นเช่นนั้นจะช่วยให้ปฏิบัติงานได้สำเร็จดีหรือไม่

การวัดบุคลิกภาพมีลักษณะพิเศษผิดแผกไปจากการวัดความสามารถด้านต่าง ๆ ของบุคคลที่ได้กล่าวมาแล้วในบทก่อน ๆ ในการวัดความสามารถ ผู้วัดและผู้ถูกวัดมีความเข้าใจตรงกันว่า ผู้ถูกวัดจะต้องตอบปัญหาหรือคำถามให้ถูกต้องมากที่สุด เพื่อให้ได้คะแนนมากที่สุด แต่ในการวัดบุคลิกภาพ ผู้วัดและผู้ถูกวัดอาจมีความตั้งใจขัดกัน ผู้วัดพยายามตั้งคำถามเพื่อให้ผู้ถูกวัดอธิบายความรู้สึกที่แท้จริง แต่ผู้ถูกวัดอาจพยายาม

ที่จะตอบคำถามหรือปัญหาที่กำหนด โดยซ่อนความรู้สึกของตนไว้ให้มากที่สุดเท่าที่จะมากได้ หรืออาจพยายามแสดงความรู้สึกในด้านใดด้านหนึ่งให้มากที่สุดเกินกว่าความจริง ในการทำให้คำตอบของผู้ถูกวัดอาจเป็นในลักษณะเปะปะ หรือฟุ่มเฟือยก็ได้ กล่าวโดยสั้น ๆ ก็คือ ในการวัดบุคลิกภาพ นักวัดบุคคลต้องคำนึงเสมอว่าเครื่องมือวัดของตนสามารถแก้ปัญหาการที่ผู้ถูกวัดขาดการเร่งเร้าในการแสดงภาพที่แท้จริงของตนได้หรือไม่เพียงใด

การวัดบุคลิกภาพไม่ได้ก้าวหน้าเท่าเทียมกับการวัดสติปัญญา หรือความถนัดพิเศษ ความจริงนักวัดบุคคล เช่น Galton, Spearman และ Thurstone ได้พยายามอย่างยิ่งที่จะสร้างเครื่องมือวัดบุคคลในด้านนี้ แต่การที่การวัดบุคลิกภาพก้าวหน้าไม่ทัดเทียมกับการวัดสติปัญญาด้านต่าง ๆ ของบุคคลก็เพราะไม่สามารถสร้างเครื่องมือวัดที่มีความเที่ยงตรงถึงระดับที่น่าพอใจได้ ในทางจิตวิทยา ได้มีการวิเคราะห์ลักษณะนิสัยไว้แล้วหลายอย่าง เช่น ความกังวล ความซาเหิน ความขย้อตนเอง ความเห็นแก่ตัว เป็นต้น แต่ก็ยังไม่มีเครื่องมือวัดอันใดที่อ้างได้ว่า สามารถวัดลักษณะนิสัยเหล่านั้นได้อย่างเที่ยงตรง แต่ถึงกระนั้นการศึกษาเรื่องการวัดบุคลิกภาพก็ได้ก้าวหน้ามาจนถึงจะกล่าวต่อไป

ความหมายของบุคลิกภาพ คำว่า “บุคลิกภาพ” ได้มีผู้นิยามไว้ต่าง ๆ กัน หลังจากได้พิจารณาคำนิยามไม่น้อยกว่า 50 แบบแล้ว Guilford ได้นิยามบุคลิกภาพไว้อย่างรัดกุมว่า “บุคลิกภาพ คือ ลักษณะนิสัย (traits) ที่รวมกันเป็นแบบฉบับเฉพาะตัวของแต่ละบุคคล”¹ และได้อธิบายความหมายของนิยามนี้ต่อไปว่า บุคลิกภาพเป็นสิ่งที่สร้างความแตกต่างระหว่างบุคคล การที่จะทราบได้ว่าผู้ใดมีบุคลิกภาพอย่างใดก็ต่อเมื่อนำบุคคลต่าง ๆ มาเปรียบเทียบกัน แต่การที่จะนำลักษณะทั้งหมดของบุคคลแต่ละคนมาเทียบกันจริง ๆ เป็นสิ่งที่พื้นวิสัย ในการเปรียบเทียบนี้อาจทำได้แต่เพียงนำเอาคุณสมบัติบางอย่างของแต่ละบุคคลมาพิจารณาทีละอย่างโดยกำหนดปริมาณของคุณสมบัตินั้นให้แน่นอน และระยะเวลาที่บุคคลมีคุณสมบัตินั้น ๆ ให้ชัดเจนลงไป เช่น เปรียบเทียบว่า นาย ก. ต่างกับ นาย ข. ในแง่คุณสมบัติประการที่ 1, 2, 3 และอื่น ๆ หรือ นาย ก. ต่างกับ นาย ง. ในแง่คุณสมบัติประการที่ 4, 5, 6 และอื่น ๆ หรือ นาย ฮ. ต่างกับบุคคลเฉลี่ยโดยทั่ว ๆ ไป ในแง่คุณสมบัติประการที่ 7, 8, 9 และอื่น ๆ คุณสมบัตินี้แต่ละประการที่นำมาเปรียบเทียบเพื่อให้ทราบลักษณะของบุคลิกภาพนั้น จำกัดเฉพาะลักษณะ

¹ Guilford, G.P., *Personality*, pp. 2, 4.

นิสัย เท่านั้น ลักษณะนิสัยนี้เป็นลักษณะประจำตัวที่ทำให้บุคคลแต่ละคนต่างกัน ลักษณะนิสัยที่นำมาเทียบอาจเป็นประเภทที่มีลักษณะกว้าง ๆ มาก เช่น ลักษณะความเชื่อมั่นในตนเอง หรืออาจเป็นประเภทที่มีลักษณะแคบ ๆ ก็ได้ เช่น ความยืดหยุ่นของกล้ามเนื้อในระดับเสียงต่าง ๆ

ตามคำนิยามนี้ สถิติปัญญา ความถนัดพิเศษ ทักษะทางกาย ตลอดจนลักษณะความต้องการเฉพาะของแต่ละบุคคล (needs) ล้วนเป็นเรื่องที่เกี่ยวข้องกับบุคลิกภาพทั้งสิ้น แต่ในแง่วิชาการวัดบุคคลที่ได้กล่าวมาแล้วในบทที่ 9 ได้ชี้ให้เห็นแล้วว่า สถิติปัญญา ความถนัดพิเศษและทักษะทางกาย เป็นสิ่งที่สามารถวัดได้โดยเฉพาะเป็นเรื่อง ๆ ไป จริงอยู่ อาจมีผู้ค้านว่า ลักษณะทั้งสามประการนี้มีส่วนทำให้บุคคลมีบุคลิกภาพอย่างหนึ่ง ๆ ฉะนั้น ในการพิจารณาบุคลิกภาพจะต้องพิจารณาลักษณะทั้งสามประการรวมพร้อมกันไปด้วย แต่ในการวัดผลนั้นเมื่อจะทำการวัดบุคลิกภาพ นักทฤษฎีได้พยายามวัดแนวโน้มของพฤติกรรมมากกว่าที่จะมุ่งไปแยกแยะระดับของสถิติปัญญา หรือความถนัดพิเศษ หรือทักษะทางกาย ฉะนั้น ในบทที่ 1 จึงได้สรุปพฤติกรรมที่จะใช้เป็นแนวในการวัดบุคลิกภาพไว้ว่า มี 5 ประการ คือ อุปนิสัย (character) ความสามารถในการปรับตัว (adjustment) อารมณ์ (temperament) ความสนใจ (interest) และทัศนคติ (attitude)

การตัดสินใจว่าบุคคลมีอุปนิสัยเช่นใด ขึ้นอยู่กับค่านิยมของสังคม (social value) บุคคลที่มีอุปนิสัยดีจึงเป็นคนที่มัลักษณะอย่างหนึ่งอย่างใด ดังต่อไปนี้ คือ ซื่อสัตย์ ใจกว้าง ขยันขันแข็ง ร่วมมือกับผู้อื่น และอื่น ๆ ส่วนผู้ที่สามารถปรับตัว คือ ผู้ที่ครองชีวิตในสังคมโดยไม่ประสบปัญหา หรือกล่าวอีกนัยหนึ่ง คือผู้ที่ยอมรับสภาพของตัวเอง และสังคมอย่างมีความสุขสบาย คำว่า อารมณ์ นั้น ในตำราภาษาอังกฤษหลายเล่มจะใช้สลับกันไปกับคำว่า บุคลิกภาพ เมื่อมีการวัดอารมณ์ก็ถือว่าได้วัดบุคลิกภาพแล้ว แต่แท้ที่จริงอารมณ์เป็นเพียงลักษณะด้านหนึ่งของบุคลิกภาพเท่านั้น อารมณ์เป็นเพียงลักษณะความรู้สึกประจำตน อารมณ์มีได้หลายแบบ เช่น ร่าเริงหรือเศร้าซึม แข็งแรงหรือเฉื่อยชา ตื่นเต้นหรือเฉื่อยสงบ เป็นต้น ความสนใจเป็นแนวโน้มที่จะทำให้พยายามแสวงหา หรือพยายามมีส่วนในกิจกรรมอย่างหนึ่ง ๆ ลักษณะสุดท้าย คือ ทัศนคตินั้น ก็เป็นแนวโน้มในการยอมรับหรือไม่ยอมรับ บุคคลบางกลุ่มบางประเภท หรือแนวความคิดบางอย่าง หรือสถาบันทางสังคมบางลักษณะ

วิธีการในการวัดบุคลิกภาพ

จากคำอธิบายความหมายของบุคลิกภาพ จะเห็นว่า บุคลิกภาพมีลักษณะพิเศษไปกว่าคุณสมบัติด้านความสามารถของบุคคล ในการวัดสติปัญญาที่ดี หรือในการวัดความสัมฤทธิ์ผลที่ดี นักวิชาการอาจสร้างแบบทดสอบใช้เป็นเครื่องมือวัดได้ แต่บุคลิกภาพไม่ใช่เป็นสิ่งที่สามารถใช้แบบทดสอบได้โดยตรง จริงอยู่ ได้มีนักวิชาการบางคนพยายามสร้างแบบทดสอบเพื่อวัดบุคลิกภาพ แต่แบบทดสอบเหล่านั้นก็ไม่ได้เป็นที่ยอมรับกันว่าสามารถวัดบุคลิกภาพได้ทุกแง่ทุกมุม จึงได้มีผู้พยายามคิดค้นแนววิธีต่าง ๆ หลายอย่างที่วัดลักษณะด้านนี้ของบุคคล แนววิธีต่าง ๆ นี้ อาจแยกได้เป็น 4 แบบ คือ

1. วัดพฤติกรรมจริง ๆ ของแต่ละบุคคล คือ พิจารณาว่าบุคคลนั้นทำอะไร ประพฤติตนอย่างไรต่อสภาพแวดล้อมและบุคคลทั่วไป
2. วัดความเห็นของบุคคลอื่น ๆ คือ ค้นหาว่าบุคคลอื่น ๆ กล่าวถึงบุคคลนั้นอย่างไร
3. พิจารณาสິงที่บุคคลนั้นกล่าวถึงตนเอง คือ วิเคราะห์ว่าบุคคลนั้นมองเห็นตนเองอย่างไร
4. สมมติสภาพการณ์ให้บุคคลนั้นพิจารณาแล้ววัด พฤติกรรมของบุคคลนั้นในสภาพการณ์นั้น ๆ

การวัดพฤติกรรมจริง ๆ ของแต่ละบุคคลนั้น ทำได้ทั้งวิธีทดสอบ และวิธีสังเกต รายละเอียดของวิธีทั้งสองนี้ กล่าวไว้แล้วในบทที่ 1 อย่างไรก็ตามวิธีทดสอบบุคลิกภาพนั้นมีข้อสำคัญว่า ควรจะต้องกำหนดให้มีลักษณะที่ซ่อนเงื่อนเพื่อให้ผู้ถูกวัดสามารถแสดงสิ่งที่บิดเบือนไปจากบุคลิกภาพจริง ๆ ของตนได้² การซ่อนเงื่อนนี้เป็นของจำเป็นเพราะในการวัดบุคลิกภาพนี้ไม่ต้องการทราบพฤติกรรมที่ดีที่สุดของแต่ละบุคคล แต่ต้องการทราบพฤติกรรมทั่วไปของแต่ละคนมากกว่า

ในด้าน การวัดความเห็นของบุคคลอื่น เพื่อที่จะให้ทราบบุคลิกภาพของคนหนึ่งๆนั้น เป็นการวัดในแง่ว่าบุคคลอื่นมองเห็นบุคคลนั้นอย่างไร เพื่อนร่วมงานถือว่าบุคคลนั้นเป็นคนที่ประพฤติเป็นกันเองกับคนอื่นหรือเปล่า บริษัทลูกค้านั้นถือว่าพนักงานเดินตลาดคนหนึ่ง ๆ ปฏิบัติตนให้เป็นที่น่าเชื่อถือหรือเปล่า ในการวัดความเห็นของบุคคลอื่นเพื่อที่จะให้ทราบว่าบุคคลนั้น ๆ มีบุคลิกภาพเช่นใด ทำได้โดยใช้วิธีการให้คะแนน

² Venon, P.E., *The Measurement of Abilities*, pp. 193-194.

โดยผู้อื่น (rating procedure) บริษัท ห้างร้านบางแห่งได้ถือปฏิบัติในการที่จะคัดเลือกบุคคลที่มาสมัครงานว่า จะต้องส่งแบบฟอร์มให้คะแนนไปให้บุคคลที่รู้จักผู้สมัครงานตอบ เพื่อที่จะให้ทราบว่าคนที่สมัครงานนั้นมีบุคลิกภาพเช่นใด เป็นการประกอบการพิจารณา แบบฟอร์มให้คะแนนนี้ทำได้หลายแบบ แบบข้างล่างนี้เป็นแบบหนึ่ง

	ดีมาก	พอใช้ได้	ไม่ดี
ความพยายาม	--	--	--
ความสามารถปรับตัว	--	--	--
ความร่วมมือกับผู้อื่น	--	--	--
ความกระตือรือร้น	--	--	--

ในการพิจารณาสิ่งทีบุคคลหนึ่ง ๆ กล่าวถึงตนเองเพื่อที่จะให้ทราบว่าบุคคลนั้น ๆ มีบุคลิกภาพอย่างไร เป็นการพิจารณาบุคลิกภาพของแต่ละคนโดยถือหลักว่า การที่จะเข้าใจบุคลิกภาพของแต่ละคนนั้นจะต้องพิจารณาจากภายในตัวบุคคลนั้น ๆ วิธีการสองวิธีแรกเป็นการพิจารณาบุคลิกภาพของแต่ละบุคคล โดยพิจารณาจากสภาพที่มองเห็นว่าเป็นจริง กล่าวคือ พิจารณาว่าบุคคลหนึ่ง ๆ มีบุคลิกภาพอย่างไร โดยถือเอาสิ่งที่ผู้อื่นมองเห็นผู้นั้นเป็นเกณฑ์ แต่การที่เปิดโอกาสให้เจ้าตัวกล่าวถึงตนเองเพื่อให้ทราบบุคลิกภาพของผู้นั้นย่อมเป็นทางที่จะวัดบุคลิกภาพของผู้นั้นจากสิ่งที่มีอยู่ภายใน การที่จะวัดด้วยวิธีนี้ได้ก็อาจทำได้โดยการสัมภาษณ์ ให้ตอบแบบสอบถาม และให้กรอบัญชีรายการ (inventory) วิธีการใช้บัญชีรายการเพื่อวัดบุคลิกภาพของแต่ละคนเป็นวิธีการที่แพร่หลายที่สุด เพราะสามารถทำให้เป็นมาตรฐานได้ คะแนนของคำตอบที่มีอยู่ในบัญชีรายการจะเป็นเครื่องชี้ให้เห็นภาพในเรื่องบุคลิกภาพของผู้นั้น

ส่วนการวัดบุคลิกภาพโดยการสมมติสภาพการณ์เพื่อให้บุคคลหนึ่ง ๆ พิจารณาแล้ววัดพฤติกรรมของผู้นั้น ในสภาพการณ์นั้น ๆ เป็นการเปิดโอกาสให้บุคคลนั้นได้แสดงตนเองของตนเองอย่างเต็มที่ ทั้งนี้โดยถือหลักว่าการตอบคำถามต่าง ๆ หรือการอธิบายตนเองยังอาจจะทำให้ไม่สามารถทราบลักษณะทั้งหมดที่เกี่ยวกับบุคลิกภาพของผู้นั้นได้ การสมมติสภาพการณ์มีหลักว่า สิ่งทีบุคคลหนึ่ง ๆ มองเห็นนั้นขึ้นอยู่กับบุคลิกภาพของผู้นั้น สิ่งทีคนหนึ่ง ๆ มองเห็นในสิ่งต่าง ๆ เป็นการสะท้อนลักษณะภายในของบุคคลนั้น ฉะนั้น เมื่อบุคคลหนึ่ง ๆ มีโอกาสพิจารณาสภาพการณ์หนึ่ง ๆ แล้วแสดงปฏิกิริยาได้ตอบได้โดยไม่ต้องมีกรอบจำกัด ย่อมทำให้เห็นภาพทั้งหมดของบุคคลนั้น วิธีการนี้ปฏิบัติ

ได้โดยสมมุติรูปหมึก หรือรูปภาพ หรือรูปปั้นต่าง ๆ ให้บุคคลที่ถูกวัดพิจารณาแล้ว ให้แสดงพฤติกรรมออกมา พฤติกรรมของบุคคลที่ถูกวัดที่มีต่อรูปเหล่านั้นเป็นการสะท้อนบุคลิกภาพของผู้นั้น อย่างไรก็ตาม แนวคิดในการวัดด้วยวิธีนี้แม้จะเป็นที่ยอมรับกันอย่างแพร่หลาย แต่ก็ยังมีบุคคลบางส่วนที่ยังเชื่อว่า การวัดบุคลิกภาพด้วยวิธีนี้ไม่ทำให้ทราบบุคลิกภาพอย่างแท้จริงของแต่ละคน³

แบบทดสอบ แบบทดสอบในที่นี้หมายถึงแบบคำถามทดสอบเท่านั้น ในการสร้างเครื่องมือวัดแบบนี้ ปรากฏว่าได้มีผู้พยายามสร้างขึ้นเพื่อวัดอารมณ์ ลักษณะบางอย่างของนิสัย อุปนิสัย และความสนใจ แต่ยังไม่มียแบบทดสอบที่วัดทัศนคติ หรือความสามารถในการปรับตัว แบบทดสอบที่สร้างขึ้นนี้มีปัญหาในเรื่องลักษณะการตอบพลิกแพลง (stylistic response sets) กล่าวคือ ผู้ตอบมักจะปฏิบัติอย่างใดอย่างหนึ่งหรือทั้ง 3 อย่าง ดังต่อไปนี้ คือ พยายามที่จะหลอกผู้วัดโดยพยายามตอบให้ไม่ตรงกับที่ใจคิด ตั้งใจที่จะแสดงภาพที่ตนคิดว่าต่างกับภาพที่ตนเป็นจริง ๆ หรือประการสุดท้ายขาดความเข้าใจทำให้คำตอบบิดเบือนไป การปฏิบัติของผู้ถูกทดสอบเช่นที่กล่าวมานี้ย่อมทำให้ไม่ได้ทราบบุคลิกภาพที่แท้จริงของผู้นั้น จะเห็นได้ง่าย ๆ ถ้าหากว่าลองพิจารณาคำถามต่อไปนี้ “ท่านชอบใจลอยหรือไม่?” “ท่านขี้เกียจสนทนากับผู้อื่นหรือไม่?” ผู้ที่มาสมัครงานเมื่อถูกถามเช่นนั้น ย่อมต้องพยายามตอบเพื่อแสดงว่าตนเป็นผู้ที่ขยันขันแข็งชอบเข้ากับผู้อื่น ก็จะพยายามที่จะตอบให้เป็นประโยชน์ต่อตนเอง Guion ได้อ้างผลงานวิจัยล่าสุดของ Wagner ในปี 1960 ซึ่งพิสูจน์ว่า การทดสอบบุคลิกภาพนั้นเท่าที่ปรากฏจากการวิเคราะห์แบบทดสอบที่มีอยู่ ไม่สามารถที่จะแยกให้ชัดเจนลงได้ว่างานในตำแหน่งหนึ่ง ๆ ต้องการบุคลิกภาพในด้านใดเพียงใด โดยเฉพาะงานในตำแหน่งผู้บริหาร การที่เป็นเช่นนั้นก็เพราะ แบบทดสอบบุคลิกภาพที่มีอยู่ยังขาดมาตรฐานการวัดความสำเร็จที่จะนำมาเทียบหาความตรงประเด็นในแง่การคาดคะเนได้

แบบทดสอบที่ได้ผลดีออกมาแล้ว และอาจนำมากล่าวในที่นี้ได้ คือ Test of Social Insight และ Social Intelligence Test⁴ แบบทดสอบแบบแรกมีคำถาม 60 คำถาม ซึ่งถามผู้ถูกวัดว่า จะทำอย่างไรบ้างในเหตุการณ์ต่าง ๆ เช่น เมื่อตกผิดคน เมื่อถูกขอให้บริจาคเงินของขวัญให้ผู้บังคับบัญชาซึ่งตนไม่ชอบ หรือเมื่อไปกินอาหารในร้านแล้วไม่มีเงินจ่าย แบบทดสอบแบบที่สองประกอบด้วยคำถาม 5 ตอน ซึ่งถามทั้งในแง่การปฏิบัติตนเมื่อตกอยู่ในสภาพการณ์อย่างหนึ่ง ๆ ความสามารถในการจำชื่อและหน้า

³ Guion, Robert M., *Personnel Testing*, pp. 340–344.

⁴ *Ibid.*, p. 331.

คนที่รู้จัก และอารมณ์ขันต่าง ๆ ในปี 1962 Cattell ได้พยายามสร้างแบบทดสอบชื่อ Objective-Analytic Personality Test Batteries⁵ แบบทดสอบนี้เป็นกลุ่มของแบบทดสอบซึ่งพยายามวัดองค์ประกอบต่าง ๆ ของบุคลิกภาพ ซึ่ง Cattell ได้พิสูจน์ไว้และได้กล่าวอ้างถึงมาแล้วในที่นี้ อย่างไรก็ตาม ความสำเร็จและความตรงประเด็นของแบบทดสอบนี้ก็ยังไม่เป็นที่ยอมรับทั่วไปอย่างเต็มที่ จะต้องรอรยะเวลาหนึ่งซึ่งพิสูจน์เสียก่อน

ในการวัดบุคลิกภาพ ได้เชื่อกันว่าหากวัดความสนใจแยกต่างหากออกมาโดยเฉพาะจะทำให้ได้ภาพของบุคคล ซึ่งจะประกอบการพิจารณาความเหมาะสมของผู้นั้นในการปฏิบัติงานได้ชัดเจนยิ่งขึ้น จึงได้มีผู้สร้างแบบทดสอบเพื่อวัดความสนใจในขั้นนี้ แต่แบบทดสอบที่ใช้วัดความสนใจนี้ มิใช่แบบทดสอบที่มีลักษณะเหมือนกับแบบทดสอบทั่วไป แบบทดสอบที่วัดความสนใจมีลักษณะเป็นแบบทดสอบอำพราง (disguised interest tests) กล่าวคือ ไม่ได้มีการถามตรง ๆ ว่าสนใจอะไรบ้าง หรืออย่างไร แต่ไปถามเนื้อหาข้อเท็จจริงบางอย่างที่บุคคลที่สนใจในเรื่องหนึ่ง ๆ ควรทราบ เช่น ถามว่าผู้ถูกวัดรู้คำศัพท์เกี่ยวกับอาชีพต่าง ๆ มากน้อยเพียงใด หรือถามว่า รู้จักส่วนประกอบของเครื่องมือที่ใช้ในการปฏิบัติงานหนึ่ง ๆ เพียงใด การถามเช่นนี้ย่อมเป็นที่น่าสงสัยว่ามีความตรงประเด็นในแง่ความสัมพันธ์กับความสำเร็จในการปฏิบัติงานนั้น ๆ เพียงใด

การสังเกต ได้มีผู้ค้นคว้าเทคนิคในการสังเกตบุคคลไว้มากมาย และได้นำไปใช้ได้ผลตามสมควรในงานวิชาการหลายอย่างแล้ว เช่น งานวิจัยและงานแนะแนว เป็นต้น แม้ในการวัดบุคลิกภาพก็เชื่อว่า การสังเกตสามารถช่วยให้วัดบุคลิกภาพของแต่ละคนได้ แต่เนื่องจากการสังเกตนี้จะต้องมีขอบวนพิศการ ซึ่งเตรียมเป็นพิเศษ บุคคลที่จะถูกวัดบุคลิกภาพด้วยการสังเกตจึงมีแต่เพียงบุคคลบางประเภท เช่น เด็ก นักเรียน และคนไข้ทางจิต เป็นต้น นอกจากนี้การที่จะนำวิธีการสังเกตมาวัดบุคลิกภาพของบุคคลเพื่อประโยชน์ในการบริหารนั้นยังมีปัญหาอีกมากมาย ซึ่งทำให้ไม่เหมาะที่จะนำวิธีการสังเกตนี้มาใช้⁶ เช่น ต้องสิ้นเปลืองค่าใช้จ่ายมากเกินไป ผลที่วัดได้มีความลำเอียงของผู้สังเกตประกอบด้วย การแยกแยะประเด็นย่อยของบุคลิกภาพที่สังเกตได้ทำได้ยาก เป็นต้น

⁵ *Ibid.*, p. 345.

⁶ Thorndike Robert L., and Hagen, Elizabeth, *Measurement and Evaluation in Psychology and Education* pp. 321–324.

การให้คะแนนโดยผู้อื่น (*rating procedure*) วิธีนี้เป็นวิธีที่ใช้อย่างกว้างขวางในการประเมินผลการปฏิบัติงาน ซึ่งอาจจะเรียกใหม่ได้ว่าเป็นการให้คะแนนตามระบบคุณธรรม (*merit rating system*) เนื่องจากวิธีวัดแบบนี้มีกล่าวไว้ในวิชาที่เกี่ยวกับการบริหารงานบุคคลเป็นส่วนใหญ่ ในที่นี้จึงจะไม่อธิบายรายละเอียดของการวัดผลตามวิธีนี้ แต่จะกล่าวถึงหัวข้อต่าง ๆ ที่ใช้ในวิธีนี้โดยย่อ คือ⁷

1. การให้คะแนนทั่วไป (*rating scales*) ทำได้โดยกำหนดลักษณะหรือคุณสมบัติต่าง ๆ ของผู้ปฏิบัติงาน แล้วให้ผู้บังคับบัญชาหรือผู้เกี่ยวข้องเป็นผู้ให้คะแนน ทั้งนี้อาจจะทำแบบฟอร์มได้ดั่งข้างล่างนี้โดยพิจารณากำหนดพฤติกรรมแต่ละอย่างให้ละเอียดลงไป เช่น

การให้ความร่วมมือในการปฏิบัติงานพิเศษ

กระตือรือร้นที่จะแสดง ความเห็น พยายามเสีย สละทุนทรัพย์และแรงงาน และยอมปฏิบัติงานเกิน เวลาราชการ	ทำในสิ่งที่ผู้บังคับบัญชา ขอร้องให้ความร่วมมือ ในด้านต่าง ๆ แต่ไม่ริเริ่ม	พยายามทำน้อยที่สุด ให้ความร่วมมือเท่าที่ คิดว่าจำเป็นจะต้องทำ
---	---	---

2. การจัดอัตราร้อยละ (*forced distribution system or percentage of group*) วิธีนี้ใช้กับหน่วยงานที่มีคนมาก ผู้บังคับบัญชาอาจจะต้องให้คะแนนผู้ร่วมงานโดยจัดว่าผู้ร่วมงานคนหนึ่ง ๆ สามารถหรือมีลักษณะในการปฏิบัติงานอยู่ในระดับสิบเปอร์เซ็นต์ ฯลฯ ต่ำสุด หรือสิบเปอร์เซ็นต์สูงสุด หรือถัดจากสูงสุดลงมาอีกยี่สิบเปอร์เซ็นต์ ฯลฯ

3. การจัดที่ (*ranking or rank-order system*) ผู้ที่จะออกความเห็นเกี่ยวกับบุคลิกภาพของบุคคลหนึ่ง ๆ ในการทำงานจะต้องจัดลำดับบุคคลต่าง ๆ ในแง่ของลักษณะที่เป็นประโยชน์ต่อการทำงาน เช่น กำหนดว่าในเรื่องความเป็นผู้นำนั้นแต่ละคนมีลักษณะเด่นต่างกันอย่างใด หรือในเรื่องความตรงเวลามีลำดับที่ต่างกันอย่างไร

⁷ U.S. Civil Service Commission, "Selecting Supervisors", *Personnel Methods Series, No. 2*, pp. 18-21.

Kellogg, M.S., *What to do About Performance Appraisal*, pp. 68-82.

Dooher, M. Joseph, and Marquis, Vivienne, eds., *Rating Employees and Supervisory Performance*, pp. 11-43.

4. การตอบรายการตรวจสอบ (check list rating systems) ผู้วัดจะต้องเลือกจากรายการตรวจสอบที่ระบุคุณลักษณะในการทำงาน เช่น ในการคัดเลือกนายทหารอากาศที่จะไปเป็นอาจารย์สอนวิชาช่าง ผู้บังคับบัญชาที่คัดเลือกนายทหารเหล่านั้นจะต้องกาสิ่งที่ตนเลือกตอบหนึ่งข้อจากสี่ข้อ คือ

- ก. บุคคลผู้เป็นคนอดทนไม่โกรธผู้ที่เรียนช้า
- ข. บุคคลผู้บรรยายด้วยความมั่นใจ
- ค. บุคคลผู้ทำให้ทหารที่อยู่ในชั้นสนใจและตั้งใจ
- ง. บุคคลผู้ทำให้ทหารที่เข้ารับการอบรมเข้าใจประเด็น และความมุ่งหมายของการอบรมในทุกหัวข้อ

5. การคัดเลือกเหตุการณ์ที่สำคัญ (critical incident technique) ทำได้โดยให้ผู้บังคับบัญชาพิจารณาว่าในการปฏิบัติงานหนึ่ง ๆ บุคคลแต่ละคนมีพฤติกรรมที่ดีเด่น และไม่พึงปรารถนาอย่างไรบ้าง พฤติกรรมที่บันทึกไว้เหล่านั้นจะได้นำมาจัดเรียงลำดับเพื่อชี้ให้เห็นว่าในการปฏิบัติงานหนึ่ง ๆ ให้ได้ผลดี หรือปฏิบัติไม่ได้ดีนั้นมียังไร รายการที่จัดเรียงไว้นั้นอาจจะทำได้เป็นหลายแบบ เช่น จัดเรียงว่า

- ก. วิธีการพิจารณาตัดสินใจ
- ข. ลักษณะการที่ฟังผู้อื่น
- ค. การปฏิบัติตามกฎข้อบังคับ
- ง. ความตรงต่อเวลาในการรายงานผลงาน
- จ. การริเริ่ม
- ฉ. ความรับผิดชอบ
- ช. ความสามารถเข้ากับผู้อื่น

การให้คะแนนโดยผู้อื่นนั้นก็คือ การวัดบุคลิกภาพของบุคคลหนึ่ง ๆ โดยขอให้ผู้อื่นสังเกตผู้ถูกวัดนั่นเอง ถึงแม้ว่าวิธีนี้เป็นวิธีที่ใช้กันแพร่หลายในการบริหารงานบุคคล แต่ถ้าพิจารณาในแง่การวัดผลแล้วเป็นที่น่าสงสัยว่าวิธีนี้จะวัดได้ผลหรือไม่ เพราะสาเหตุหลายประการ คือ

1. ผู้ที่ให้คะแนนนั้นไม่ได้ใช้ความพยายามหรือความรอบคอบอย่างเต็มที่ในการให้คะแนน
2. ผู้ให้คะแนนมักจะถือตนว่ามีความสำคัญเกี่ยวข้องกับผู้ถูกให้คะแนน ไม่มากนัก

3. ประเด็นที่ต้องให้คะแนนนี้อาจตีความหมายได้หลายแง่หลายมุม
4. คะแนนที่ให้นั้นเป็นผลจากการพิจารณากิริยาอาการภายนอกของผู้ที่ถูกรัด
5. ผู้ให้คะแนนไม่มีโอกาสที่จะสังเกตผู้ถูกให้คะแนนตลอดเวลา
6. มาตรฐานการให้คะแนนของแต่ละคนไม่เท่ากัน
7. ผู้ให้คะแนนแต่ละคนมีลักษณะพิเศษประจำตน ทำให้ถือประเด็นที่สำคัญในการให้คะแนนไม่เหมือนกัน

การที่มีเหตุถึง 7 ประการที่จะทำให้วิธีการให้คะแนนนี้วัดไม่ได้จริง ๆ ทำให้ผลที่วัดได้นั้นมีส่วนเสียหลายอย่าง คือ

1. คะแนนที่ให้มักจะเป็นส่วนดีแก่ผู้ถูกรัด (generosity error)
2. คะแนนที่ได้บิดเบือนเพราะอคติของผู้ให้ (halo error)
3. ขาดความเชื่อถือได้เพราะมาตรฐานการวัดหรือการให้คะแนนไม่เหมือนกัน

การสัมภาษณ์และการส่งแบบสอบถาม เนื่องจากบุคลิกภาพเป็นเรื่องที่เจ้าตัวเท่านั้นจะมีโอกาสสังเกตว่าตนมีพฤติกรรมและปฏิกิริยาต่อสิ่งแวดล้อมเช่นไรมากกว่าผู้อื่น จึงได้ถือกันว่าการวัดบุคลิกภาพด้วยเครื่องมือต่าง ๆ ดังกล่าวมาแล้วยังไม่ได้ผลดี แม้จะถือว่าแบบทดสอบเป็นเครื่องมือที่ดีที่สุดในการวัดสติปัญญา ความถนัดทางธรรมชาติและความสำเร็จผลก็ตาม เพราะไม่มีใครที่จะอยู่ในฐานะที่จะมองสภาพภายในของบุคคลหนึ่ง ๆ ได้ดีกว่าเจ้าตัวเอง จึงได้มีผู้พยายามวัดบุคลิกภาพด้วยการสัมภาษณ์ และการส่งแบบสอบถาม แต่เนื่องจากว่าการสัมภาษณ์นั้นจะทำให้เป็นมาตรฐานได้ยาก เพราะวิธีการปฏิบัตินั้นขึ้นอยู่กับผู้สัมภาษณ์และการแปลความหมายของผู้สัมภาษณ์มากกว่าตัวผู้ให้สัมภาษณ์เอง ฉะนั้น ถึงแม้ว่าการสัมภาษณ์ยังเป็นที่ยอมรับดำเนินการในงานวิชาการบางอย่าง เช่น งานวิจัยทั่วไปก็ตาม แต่ก็เป็นที่เชื่อกันว่าเป็นวิธีที่ไม่สามารถวัดบุคลิกภาพของแต่ละคนได้อย่างแท้จริง ส่วนในการทดสอบแบบสอบถามก็ได้มีการค้นคว้าเทคนิคการออกแบบสอบถามและการบริหารงานแบบสอบถามไว้อย่างกว้างขวางจนเป็นที่ยอมรับใช้แพร่หลายในงานวิจัยสังคมศาสตร์ทั่วไป แต่ก็ยังเป็นปัญหาในการที่จะนำมาใช้เพื่อการวัดงานบุคคล⁸ ฉะนั้นจึงได้มีการทำแบบสอบถามให้มีลักษณะเป็นมาตรฐาน เพื่อให้ผู้ถูกรัดสามารถแสดงบุคลิกภาพของตนเองออกมา แบบสอบถามประเภทนี้เรียกได้ว่า บัญชีรายการ

⁸ Lawshe and Balma, *op. cit.*, pp. 108–109, 233–234.

บัญชีรายการ (*inventory*) เป็นเครื่องมือที่แพร่หลายที่สุดในการวัดบุคลิกภาพในด้านต่าง ๆ ได้แก่ อารมณ์ ความสนใจ ทักษะคิด และลักษณะบางอย่างของนิสัยและอุปนิสัย บัญชีรายการที่มีชื่อที่สุดได้ถือหลักอันประกอบของบุคลิกภาพซึ่ง Guilford และ Cattell ได้แยกไว้ วิธีการสร้างรายการในบัญชีทำโดยตั้งคำถามให้ผู้ถูกวัดเลือก (*forced-choice*) วิธีนี้เชื่อกันว่า ผู้ถูกวัดจำเป็นต้องเลือกสิ่งที่แสดงความเป็นตนของตนเองในระยะเวลาจำกัด ทำให้ไม่สามารถพลิกแพลงคำตอบหลอกลวงผู้วัดได้ บัญชีรายการที่วัดบุคลิกภาพที่มีชื่อที่สุด คือ Minnesota Multiphasic Personality และ Guilford-Zimmerman Temperament Survey⁹ บัญชีรายการแบบแรก ประกอบด้วยรายการที่ถามบุคลิกภาพของผู้ถูกวัดเป็นกลุ่มและผู้ถูกวัดเดี่ยว คำถามที่บังคับให้ผู้ถูกวัดเลือกนั้นจะแสดงอารมณ์ของผู้ถูกวัดในระดับต่าง ๆ กัน เช่นเป็นคนชอบเป็นทุกข์ เป็นคนชอบกังวล ส่วนในบัญชีรายการแบบที่สองนั้น คือ องค์ประกอบของบุคลิกภาพ ซึ่ง Guilford ได้แยกไว้เป็นหลักในการตั้งคำถาม

ในการวัดบุคลิกภาพด้วยบัญชีรายการก็ถือหลักเช่นเดียวกับแบบทดสอบ (แบบคำถามทดสอบ) ที่กล่าวมาแล้ว คือ ผู้วัดเชื่อว่านอกจากจะวัดบุคลิกภาพรวม ๆ แล้วยังควรจะวัดความสนใจแยกออกไปต่างหากอีกด้วย แต่ในการสร้างบัญชีรายการนี้ยังได้มีความคิดแตกแขนงเพิ่มเติมไปกว่า การแยกวัดทัศนคติก็จะเป็นประโยชน์ด้วยเหมือนกัน ฉะนั้น จึงได้มีผู้สร้างบัญชีรายการที่วัดประเด็นเฉพาะทั้งสองอย่างของบุคลิกภาพไว้ด้วย ในเรื่องความสนใจนั้นได้มีผู้สร้างบัญชีรายการไว้หลายแบบ แบบที่มีชื่อที่สุดคือ Strong Vocational Interest Blank¹⁰ ซึ่งมีทั้งฉบับที่วัดความสนใจของเพศหญิงและเพศชายโดยเฉพาะ ผู้สร้างบัญชีรายการนี้ได้ใช้เวลาวิจัยถึง 22 ปี เพื่อพิสูจน์ความตรงประเด็นของเครื่องมือ กล่าวคือ ในการสร้างบัญชีรายการนี้ เมื่อผู้วัดได้สร้างรายการที่จะถามความสนใจของผู้ที่มีความสำเร็จในการประกอบอาชีพต่าง ๆ แล้ว ก็ได้มีการติดตามทุก ๆ 5 ปี เพื่อที่จะให้ทราบว่าความสนใจของผู้ประกอบอาชีพนั้น ๆ ยังมีค่าสหสัมพันธ์เป็นอย่างไร นอกจากนั้นในปีที่ 18 ยังมีการติดตามอีกว่า เมื่อได้ใช้บัญชีรายการนี้ทดลองกับนักศึกษามหาวิทยาลัยสแตนฟอร์ดแล้ว ปรากฏว่านักศึกษาเหล่านั้นมีความสนใจในขณะที่ทดสอบสอดคล้องกับความเป็นจริงเมื่อ 18 ปีที่ผ่านมาไปแล้วเพียงใด ปรากฏว่าค่าสหสัมพันธ์ต่าง ๆ เหล่านี้ได้ผลเป็นที่น่าพอใจ คำถามที่ปรากฏในรายการแบ่งออกได้เป็น 8 ภาค เป็นคำถามเกี่ยวกับการปฏิบัติงานวิชาที่เรียน การละเล่น

⁹ Guion, *op. cit.*, pp. 249, 151.

¹⁰ *Ibid.*, pp. 307-311.

กิจกรรมทั่วไป และลักษณะพิเศษต่าง ๆ ผู้ถูกวัดจะต้องเลือกรายการต่าง ๆ ว่าตนไม่ชอบหรือไม่สนใจอย่างใด ผู้สร้างบัญชีรายการได้คำนวณคะแนนมาตรฐานของความสนใจของผู้ที่สำเร็จในอาชีพต่าง ๆ ไว้ เพื่อเปรียบเทียบกับระดับค่าความสนใจของผู้ถูกวัดแต่ละคนไว้ อย่างไรก็ตามบัญชีรายการนี้ถึงแม้จะมีชื่อเสียงเด่นที่สุด และมีรายงานว่าหน่วยงานบางแห่งได้ใช้ประกอบการพิจารณาตัดสินใจ จะปฏิบัติการเกี่ยวกับการบรรจุแต่งตั้ง เลื่อนตำแหน่งผู้ปฏิบัติงานก็ตาม ก็ยังไม่เป็นที่ยอมรับกันอย่างสิ้นเชิงว่า เป็นเครื่องมือวัดที่ถูกต้องเที่ยงตรงที่สุด เพราะยังไม่ได้พิสูจน์ความตรงประเด็นกับความสำเร็จของงาน

ในเรื่องทัศนคตินั้น ไม่มีรายการใดที่มีชื่อเสียงเด่นที่สุดเหมือนกับบัญชีรายการของ Strong บัญชีรายการเท่าที่ได้มีผู้คิดขึ้นแล้ว และสมควรนำมากล่าวก็คือ บัญชีรายการของ Thurstone และ Likert บัญชีรายการของ Thurstone ประกอบด้วยรายการ 20 รายการ ผู้ถูกวัดจะต้องระบุว่าตนเห็นด้วยและไม่เห็นด้วย รายการแต่ละข้อมีคะแนนตั้งแต่ 0 ไปจนถึง 11.0 เช่น¹¹

ลัทธิคอมมิวนิสต์เป็นวิธีแก้ปัญหามหาเศรษฐกิจปัจจุบัน (คะแนน 9.1)

ตำรวจยิงคอมมิวนิสต์ตายไม่ผิด (คะแนน 0.3)

คนที่แสวงหาทรัพย์สินเงินทองเป็นของตนเองได้มากเท่าไร ควรจะมีสิทธิ์ใช้ทรัพย์สินเงินทองของตนเองอย่างเต็มที่ (คะแนน 3.3)

ส่วนบัญชีรายการของ Likert นั้นสร้างโดยกำหนดระดับของทัศนคติแล้วเลือกประโยคที่จะแสดงทัศนคติในระดับต่าง ๆ มาให้ผู้ถูกวัดเลือกเพื่อที่จะวิเคราะห์ให้ได้ว่า ผู้นั้นเป็นคนชอบจับผิดหรือเป็นคนที่มีมองโลกในแง่ดี เป็นคนที่มีปมด้อยหรือไม่ เป็นคนที่หัวเก่าหรือไม่ เช่น มีประโยคให้ผู้ถูกวัดเลือกว่า

การที่จะกล่าวสิ่งที่ถูกในเวลาที่เหมาะสมนั้นยาก

ทุกคนควรจะเสียสละทุกสิ่งทุกอย่างเพื่อครอบครัวของตนเอง

รัฐสภาออกกฎหมายควบคุมเสรีภาพทางการค้ามากเกินไป

คนเราทำงาน 4 ปี ได้รับความรู้มากกว่าไปโรงเรียนเสียอีก

อย่างไรก็ดี อาจขำในทีนี้อีกครั้งหนึ่งได้ว่า บัญชีรายการเหล่านี้ยังเป็นที่น่าสงสัยในเรื่องความตรงประเด็น บัญชีรายการที่วัดความสนใจนั้น มีหลักฐานพิสูจน์ความเชื่อถือ

¹¹ Tiffin, Joseph, and McCormick, Ernest J., *Industrial Psychology*, pp. 327-334.

ได้ แต่บัญชีรายการที่วัดทัศนคตินั้นยังขาดหลักฐานที่จะพิสูจน์ทั้งในเรื่องความเชื่อถือได้ และความตรงประเด็น

วิธีที่ให้ผู้ถูกวัดแสดงตนเอง (*Projective technique*) เครื่องมือวัดที่แพร่หลายที่สุดตามวิธีนี้ก็คือ Rorschach Test¹² เครื่องมือวัดนี้ประกอบด้วยภาพที่เลอะหมึก 10 ภาพแต่ละภาพมีสีต่างๆ กัน เช่น ดำ เทา ผู้ถูกวัดจะต้องพิจารณาภาพเหล่านี้ แล้วบอกผู้วัดว่าเห็นอะไรบ้าง เมื่อผู้ถูกวัดบอกครบทั้ง 10 ภาพแล้ว ผู้วัดก็ต้องถามเหตุผลที่ผู้ถูกวัดเช่นนั้น การให้คะแนนไม่ให้คะแนนเป็นตัวเลข ผู้วัดจะแยกแยะภาพที่ผู้ถูกวัดมองเห็นออกมาเป็นลักษณะต่างๆ เช่น ผ้า อาหาร แล้วแปลความหมายออกเป็นบุคลิกภาพของผู้ถูกวัด เครื่องมือวัดนี้ยังไม่มีหลักฐานการพิสูจน์สถิติว่าจะให้ความตรงประเด็นและความเชื่อถือได้เพียงใด ขบวนการวัดด้วยเครื่องมือนี้ถือว่าผู้ถูกวัดซึ่งได้รับการฝึกอบรมมาเป็นอย่างดีแล้วจะแปลความหมายของบุคลิกภาพของผู้ถูกวัดได้ถูกต้อง เครื่องมือวัดด้วยวิธีนี้ได้มีผู้คิดขึ้นอีกหลายแบบโดยถือหลักการทำนองเดียวกัน ต่างกันตรงที่ภาพซึ่งนำมาให้ผู้ถูกวัดอธิบายความหมาย เช่น อาจจะเป็นภาพของบุคคล 3 คน นั่งรับประทานอาหารร่วมกัน หรืออาจจะเป็นภาพของพนักงานโรงงานในลักษณะต่างๆ

ทฤษฎีองค์ประกอบและทฤษฎีลำดับลดหลั่นของบุคลิกภาพ

นักวิชาการได้พยายามจัดประเภทของบุคลิกภาพมาช้านานแล้ว Hippocrates (ประมาณ 460 ถึง 377 ปีก่อนคริสตศักราช) แบ่งบุคลิกภาพออกเป็น 4 ประเภท คือ ร่าเริง (sanguine) ขี้โมโห (choleric) ใจเย็น (phlegmatic) และเศร้า (melancholic) ส่วน Freud (1856—1939) แบ่งบุคลิกภาพออกเป็น 3 ประเภท คือ ประเภทปาก (oral) ประเภททวารหนัก (anal) และประเภทปลายอวัยวะเพศ (phallic) ทั้งนี้โดยถือตามทฤษฎีที่ว่า บุคลิกภาพของแต่ละคนเริ่มพัฒนามาตั้งแต่สมัยเป็นทารก โดยเกิดจากประสบการณ์ที่ได้จากการรับประทานอาหาร การขับถ่าย การได้รับการเร้งเร้าทางเพศ ในขณะเดียวกัน Jung (1875—1961) แบ่งบุคลิกภาพเป็น 2 ประเภท คือ ประเภทไม่สนใจผู้อื่น (introvert) และประเภทสนใจปรากฏการณ์รอบๆ ตัวตนเอง (extrovert) การแบ่งประเภทเช่นนี้เป็นที่สนใจไม่เพียงพอที่จะอธิบายบุคลิกภาพของบุคคลทุก ๆ คน เช่น ถ้าแบ่งประเภทตามแบบ Jung ก็ยังต้องอาจแยกประเภททั้งสองออกเป็นประเภทย่อยๆ

¹² Guion, *op. cit.*, pp. 340—341.

อีกมากมาย นักวิชาการในปัจจุบันบางคน เช่น Allport และ Odbert¹³ ได้พยายามค้นคว้าลักษณะต่าง ๆ ของบุคลิกภาพและได้พบว่า มีคำที่อาจใช้อธิบายบุคลิกภาพได้ถึง 18,000 คำ และแม้จะมีมากเช่นนี้ก็ยังไม่พอที่จะใช้อธิบายลักษณะทั้งหมดของบุคลิกภาพได้อย่างแท้จริง

จากสภาพการณ์ที่ขัดแย้งกันข้างบนนี้ จึงทำให้มีการศึกษาทฤษฎีเกณฑ์ในการวิเคราะห์บุคลิกภาพมาเรื่อย ๆ ทฤษฎีวิเคราะห์บุคลิกภาพที่ยอมรับกันในปัจจุบัน คือ ทฤษฎีองค์ประกอบของบุคลิกภาพ (personality factor theory) และทฤษฎีลำดับลดหลั่นของบุคลิกภาพ (hierarchy of traits theory) ของ Guilford และ Cattell ทฤษฎีแรกได้มีการวิจัยถึงระดับที่ชี้ให้เห็นว่า บุคลิกภาพอาจแยกเป็นส่วน ๆ ได้ประการใด ส่วนทฤษฎีที่สอง เพียงเสนอรูปแบบสมมุติ (model) ในเชิงวิชาการเท่านั้น

ทฤษฎีองค์ประกอบของบุคลิกภาพ เครื่องมือนวัตกรรม Guilford-Zimmerman Temperament Survey¹⁴ นับเป็นเครื่องมือวัดบุคลิกภาพที่แพร่หลาย และเป็นที่ยกย่องกันอย่างกว้างขวางมากที่สุด ในการสร้างเครื่องมือวัดชุดนี้ Guilford¹⁵ ได้ตั้งสมมติฐานว่า การวิเคราะห์บุคลิกภาพตามทฤษฎีของ Jung จะต้องพิจารณาลักษณะนิสัยอื่น ๆ ซึ่งแทรกอยู่ระหว่างความไม่สนใจผู้อื่น (introversion) และความสนใจปรากฏการณ์รอบ ๆ ตน (extroversion) ข้อเท็จจริงในเรื่องนี้พิสูจน์ได้โดยนำวิธีการวิเคราะห์องค์ประกอบมาคำนวณผลที่ได้จากการใช้แบบทดสอบที่สร้างขึ้นตามทฤษฎี ก่อนที่จะสร้างเครื่องมือวัดที่กล่าวนี้ Guilford ได้สร้างบัญชีรายการขึ้นทดลอง 3 ชุด คือ Inventory of Factors STDCR, Inventory of Factors GAMIN และ Personnel Inventory

ใน Inventory Factors STDCR สามารถวัดองค์ประกอบได้ดังนี้

การแยกอยู่โดดเดี่ยว	(S: social introversion, seclusiveness)
ความช่างคิด	(T: thinking introversion, reflectiveness)
การมองแต่แง่ร้าย	(D: depression, pessimism)
ความไม่มั่นคงทางอารมณ์	(C: cycloid disposition, emotional instability)
การปล่อยตามสบาย	(R: rhathymia, carefreeness)

¹³ Guilford, *op. cit.*, p. 92.

¹⁴ Guion, *op. cit.*, p. 318.

¹⁵ Guilford, *op. cit.*, pp. 183-4.

องค์ประกอบ S, T, R เป็นลักษณะที่แตกแขนงออกไปของความไม่สนใจผู้อื่น และความสนใจปรากฏการณ์รอบๆ ตัว องค์ประกอบ D, C เป็นลักษณะผสมของการที่ขาดการปรับตัวทางอารมณ์

ใน Inventory Factors GAMIN มีองค์ประกอบดังต่อไปนี้

- ความเข้มแข็ง (G : general drive for activity, energy)
- ความเหนือผู้อื่น (A : ascendancy, social boldness)
- ความเป็นชาย (M : masculinity of emotions and interests)
- ความเชื่อมั่น (I : confidence)
- ความใจเย็น (N : composure, calmness)

ใน Personnel Inventory มีองค์ประกอบดังต่อไปนี้

- ความเป็นปรนัย (O : objectivity)
- ความเป็นมิตร (Ag : agreeableness)
- ความร่วมมือ (Co : cooperativeness)

เมื่อได้ทดลองใช้บัญชีรายการทั้งสามชุด วัดองค์ประกอบได้ 13 ตัว เช่นนี้แล้ว Guilford ก็พยายามที่จะแก้ปัญหาของการที่องค์ประกอบบางตัวมีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันสูงเป็นพิเศษ โดยยุบรวมองค์ประกอบที่มีลักษณะเช่นนี้เข้าด้วยกันเสีย จนเหลือเพียงองค์ประกอบ 10 ตัว ซึ่งอาจถือว่าเป็นลักษณะที่อาจวิเคราะห์ได้จากบุคลิกภาพของบุคคลต่างๆ องค์ประกอบที่จัดใหม่ 10 ตัวนี้ ได้แก่

1. ความขยันแข็ง (G : general activity) เป็นแนวโน้มที่จะกระตือรือร้นในการเคลื่อนไหว หรือทำงาน ซึ่งตรงกันข้ามกับความเชื่องช้าเื่องหงอย
2. ความควบคุมตน (R : restraint) ได้แก่ การเป็นคนเอาจริงเอาจังไม่ปล่อยตัวตามสบาย หรือกระทำไปตามอารมณ์ชั่วแค้น
3. ความเหนือผู้อื่น (A : ascendancy) เป็นลักษณะที่จะพูด ออกความเห็น หรือแสดงตนโดยไม่รีรอ หรือชะงัก หรือยอมเป็นเบี้ยล่างผู้อื่น รวมทั้งลักษณะที่จะแสดงตนเป็นผู้นำ
4. การชอบสังสรรค์ (S : sociability) เป็นลักษณะที่ชอบติดต่อกับผู้อื่น ไม่เป็นคนขี้อาย หรือชอบเก็บตัวอยู่คนเดียว

5. ความมั่นคงทางอารมณ์ (E: emotional stability) เป็นแนวโน้มที่จะแสดงอารมณ์อย่างคงเส้นคงวา ในลักษณะที่ร่าเริง และเห็นสิ่งต่างๆ และบุคคลต่างๆ รอบตัวในด้านดี ไม่เป็นคนที่มีอารมณ์ผันแปรง่าย ๆ หรือเป็นคนเศร้าซึม

6. ความเป็นปรนัย (O: objectivity) เป็นลักษณะของคนที่มีใจหนักแน่นมองโลกในแง่ของความเป็นจริง ไม่หลอกตัวเอง พยายามมองสิ่งต่างๆ โดยไม่พิจารณาแต่ในแง่ของตนเอง ไม่เป็นคนเห็นแก่ตัว หรือขี้สงสัยผู้อื่น

7. ความเป็นมิตร (F: friendliness) เป็นลักษณะของการที่ชอบยกย่องผู้อื่น ยอมรับหรืออดทนต่ออารมณ์ศัตรูหรือการถูกกดขี่ข่มเหง ไม่เป็นคนที่คอยคิดแต่จะแก้แค้นหรืออาฆาต หรือดูหมิ่นผู้อื่น

8. การชอบคิด (T: thoughtfulness) เป็นลักษณะของการช่างคิด สนใจที่จะพิจารณาสิ่งต่างๆ แทนที่จะถือการปฏิบัติแต่อย่างเดียว

9. ความสัมพันธ์กับบุคคลอื่น (personal relations) เป็นลักษณะของการที่ชอบให้ความร่วมมือแก่ผู้อื่น สามารถทนบุคคลอื่น ๆ ได้ และมีศรัทธาในสถาบันสังคมในลักษณะต่างๆ ไม่เป็นคนที่ชอบจับผิดผู้อื่น

10. ความเป็นผู้ชาย (M: masculinity) เป็นแนวโน้มที่จะแสดงความสนใจในสิ่งต่างๆ ตามลักษณะของผู้ชาย หรือกล่าวอีกอย่างหนึ่งได้ว่าไม่เป็นคนเจ้าอารมณ์ ขี้ขลาดหวาดกลัว หรือขี้สงสารแบบผู้หญิง

Cattell¹⁶ เป็นอีกผู้หนึ่งที่ได้นำหลักการวิเคราะห์องค์ประกอบ มาศึกษาบุคลิกภาพจากการวิจัย ปรากฏผลว่า บุคลิกภาพมีองค์ประกอบ 16 ตัว แต่ก็มีนักวิชาการบางคนได้กำหนดว่า องค์ประกอบของ Cattell ยังไม่อาจแสดงส่วนสัดที่แท้จริงของบุคลิกภาพอย่างไรก็ดี เพื่อความสมบูรณ์ของเรื่อง จะได้นำรายชื่อองค์ประกอบมากล่าวไว้ในที่นี้ด้วย คือ

1. ไว้วางใจ (aloof) ตรงข้ามกับการแสดงท่าทีให้ความอบอุ่นต่อผู้อื่น
2. ทื่อ (dull) ตรงข้ามกับความคล่อง
3. หวั่นไหว (emotional) ตรงข้ามกับความเป็นผู้ใหญ่ หรือ วุฒิภาวะ

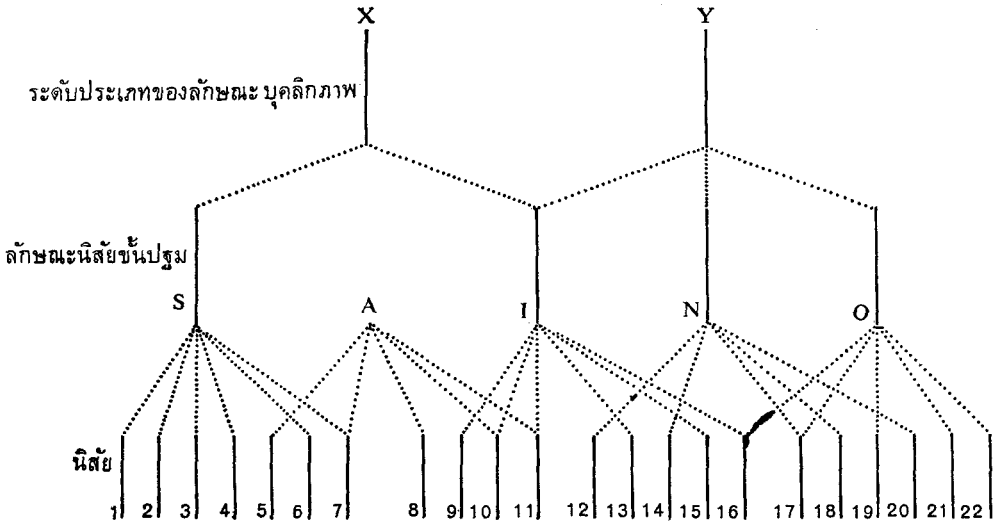
¹⁶ Guion, *op. cit.*, p. 320.

4. ขรึม (glum—silent) ตรงข้ามกับกระตือรือร้น
5. ยอม (submissive) ตรงข้ามกับการมีลักษณะชอบเหนือผู้อื่น
6. เรือเบอย (casual) ตรงข้ามกับการรู้สึกตระหนักรู้
7. ขลาด (timid) ตรงข้ามกับการเป็นคนชอบผจญภัย
8. ใจหนักแน่น (tough) ตรงข้ามกับเป็นคนอ่อนไหว
9. เชื่อผู้อื่น (trustful) ตรงข้ามกับขิงสงสัย
10. ชอบระเบียบ (conventional) ตรงข้ามกับการชอบคิดอะไรแปลก ๆ
11. ง่าย (simple) ตรงข้ามกับคนที่มีลักษณะชอบยกย่องคนอื่น
12. มั่นใจ (confident) ตรงข้ามกับคนที่รู้สึกว่าคุณาของตนเองไม่แน่นอน
13. หัวเก่า (conservative) ตรงข้ามกับคนที่ชอบทดลอง
14. ขาดสติ (lax) ตรงข้ามกับคนที่รู้จักบังคับตนเอง
15. ชอบพึ่งผู้อื่น (dependent) ตรงข้ามกับคนที่รู้สึกมั่นใจในตัวเอง
16. ใจสบาย (stable) ตรงข้ามกับคนที่อารมณ์แปรปรวน

ทฤษฎีลำดับลดหลั่นของบุคลิกภาพ Guilford¹⁷ เสนอหลักการวิเคราะห์บุคลิกภาพว่า อาจแยกพิจารณาบุคลิกภาพได้เป็น 3 ระดับ โดยเรียงจากระดับต่ำสุดไปหาระดับสูงสุด ดังต่อไปนี้ ระดับนิสัย (hexis level) ระดับลักษณะนิสัยขั้นปฐม (primary—trait level) และระดับประเภทของลักษณะบุคลิกภาพ (type level) การวิเคราะห์บุคลิกภาพตามหลักนี้เรียกว่า เป็นการกำหนดลำดับลดหลั่นของลักษณะนิสัยในบุคลิกภาพ (hierarchy of traits) รูปแบบตามทฤษฎีลำดับลดหลั่น (hierarchical model) มีแสดงในแผนภาพที่ 20

บุคคลแต่ละคนมีนิสัยประจำตัว การที่แต่ละคนสร้างนิสัยเป็นของตนเองได้ก็ด้วยจากอิทธิพล 2 ทาง ทางแรก คือ พันธุกรรม ทางที่สอง คือ การเรียนรู้ หรือฝึกฝนอบรมในระหว่างเจริญวัยเติบโตขึ้น การวิเคราะห์นิสัยของบุคคลทำได้โดยการสังเกตพฤติกรรม ในสภาพการณ์หนึ่ง ๆ แต่ละคนจะปฏิบัติกรอย่างหนึ่ง ๆ เมื่อได้สังเกตการปฏิบัติกรของแต่ละคนในสภาพการณ์เช่นเดียวกันแล้ว ก็อาจกำหนดได้ว่า บุคคลนั้น ๆ มีนิสัยเช่นใด ในแผนภาพได้ใส่เครื่องหมายเลขแทนนิสัยที่จะวิเคราะห์ได้ไว้เป็นจำนวน 22 ตัว Guilford ได้ลองสมมติตัวอย่างนิสัยไว้เพื่อให้สะดวกแก่การเข้าใจง่าย ดังต่อไปนี้

¹⁷ Guilford, *op. cit.*, pp. 99—102.



แผนภาพที่ 20 ลำดับลดหลั่นของบุคลิกภาพ

แหล่งที่มา : Guilford, J.P., *Personality*, p. 100.

1. ชอบไปงานสังสรรค์
2. ชอบทำงานกับผู้อื่น ไม่ชอบทำงานคนเดียว
3. ชอบคุยกับคนแปลกหน้า
4. ชอบไปในสถานที่ที่มีผู้คนมาก ๆ
5. ไม่ยอมให้ใครมาข่ม
6. ชอบพูดในที่สาธารณะ
7. ชอบทำงานในหน่วยงาน
8. พร้อมที่จะกล่าวโต้ตอบแก่ไม่ให้ตนผิด
9. พร้อมที่จะทำงานยาก ๆ
10. พร้อมที่จะอภิปรายความเห็นของตน
11. เมื่อได้รับการดูหมิ่น จะรีบดำเนินการอย่างหนึ่งอย่างใดทันที

ลักษณะนิสัยขั้นปฐม คือ ลักษณะของบุคลิกภาพที่พิสูจน์ได้ว่ามีมติของตนเอง หรือกล่าวอย่างง่าย ๆ ก็คือ ลักษณะด้านหนึ่ง ๆ โดยเฉพาะของบุคลิกภาพ ลักษณะนิสัย

ขั้นปฐม แสดงไว้ในแผนภาพที่ 20 ด้วย ตัวอักษร S, A, I, N, O ลักษณะนิสัยขั้นปฐมจะเป็นตัวกำหนดว่าบุคคลจะมีนิสัยเช่นใด เช่นเดียวกับที่นิสัยเป็นเครื่องชี้ว่า บุคคลจะปฏิบัติเช่นใดในสภาพการณ์หนึ่ง ๆ ตัวอักษรย่อดังกล่าวใช้แทนลักษณะนิสัยขั้นปฐมที่สมมุติ ดังต่อไปนี้

- การชอบสังสรรค์ (S: sociability) กลุ่มนิสัย 1, 2, 3, 4, 6, 7
- ความเหนือผู้อื่น (A: ascendance) กลุ่มนิสัย 3, 5, 7, 8, 10, 11
- ความเชื่อมั่น (I: confidence) กลุ่มนิสัย 9, 10, 11, ฯลฯ
- ความใจเย็น (N: composure) กลุ่มนิสัย 12, 14, ฯลฯ
- ความเป็นปรนัย (O: objectivity) กลุ่มนิสัย 16, 17, ฯลฯ

ในแผนภาพที่ 20 ได้แสดงลักษณะบุคลิกภาพเป็น 2 ประเภท คือ ประเภท X, Y ลักษณะบุคลิกภาพทั้ง 2 ประเภท มีส่วนผสมของลักษณะนิสัยขั้นปฐมร่วมกัน คือ ตัว I ซึ่งแทนความเชื่อมั่น ลักษณะผสมของบุคลิกภาพนี้อาจออกมาในรูปลักษณะต่างๆ เช่น ความสามารถในการปรับตัวกับสังคม

ทฤษฎีลำดับลดหลั่นของบุคลิกภาพสัมพันธ์เกี่ยวข้องกับทฤษฎีองค์ประกอบของบุคลิกภาพ การวิเคราะห์บุคลิกภาพตามทฤษฎีองค์ประกอบจะทำให้ทราบมิติของลักษณะของบุคลิกภาพ เมื่อได้ทราบองค์ประกอบแล้ว จะต้องนำเอาองค์ประกอบที่วิจัยได้มาพิจารณาว่าเป็นบุคลิกภาพในระดับใด องค์ประกอบของบุคลิกภาพที่วิเคราะห์ไม่จำเป็นต้องเป็นลักษณะนิสัยขั้นปฐม หรือลักษณะของบุคลิกภาพระดับประเภทสัมพันธ์ที่สัมพันธ์กันระหว่างองค์ประกอบ จะชี้ให้เห็นว่า องค์ประกอบนั้นมีลักษณะที่ร่วมกันอยู่มากหรือน้อยในระดับใด การที่จะตัดสินว่า องค์ประกอบตัวใดควรจัดให้อยู่ในระดับใด อาจถือหลัก 2 ประการ ประการแรก อิทธิพลขององค์ประกอบตัวนั้นกับองค์ประกอบตัวอื่น ๆ (determination between levels) และประการที่สอง ความสามารถในการคาดคะเนลักษณะขององค์ประกอบตัวอื่น ๆ จากองค์ประกอบตัวนั้น ๆ (prediction between levels) ยกตัวอย่าง เช่น องค์ประกอบที่แสดงว่าบุคคลมีอุดมการณ์ทางศีลธรรมสูง จะต้องครอบคลุมลักษณะนิสัยที่ซื่อสัตย์ รู้จักควบคุมตนเอง ลักษณะเช่นนี้เป็นการแสดงอิทธิพลของอุดมการณ์ทางศีลธรรม ในขณะที่เดียวกัน ผู้วิจัยก็อาจคาดคะเนต่อไปได้ว่า ผู้ที่มีอุดมการณ์สูง จะไม่เป็นคนชอบทุจริต เช่นนี้เป็นต้น

เทคนิคการวัดทัศนคติ

ในบรรดาลักษณะต่าง ๆ ของบุคลิกภาพ ทัศนคติเป็นลักษณะด้านหนึ่งที่มีความสำคัญต่อการปฏิบัติงานเป็นอย่างมาก นักบริหารที่ขยันพยายามสร้างแรงจูงใจ (incentive) เพื่อกระตุ้นให้ทุกคนร่วมมือกันคิดและปฏิบัติงานให้ลุล่วงไปอย่างมีประสิทธิภาพ แรงจูงใจนี้อาจเป็นในรูปของวัสดุสิ่งของมีค่าหรือมีประโยชน์ เช่น เงิน หรืออาจเป็นในรูปของสิ่งแวดล้อม เช่น สภาพการทำงาน หรืออาจเป็นในรูปพลังผลักดันภายในใจของผู้ปฏิบัติงานเอง ทัศนคติมีส่วนอย่างมากในการตะล่อมแรงจูงใจประเภทสุดท้ายให้เป็นประโยชน์ต่อฝ่ายบริหาร ในบรรดาผู้ปฏิบัติงานเอง อาจมีทัศนคติที่ดีหรือไม่ดีต่อฝ่ายบริหารก็ได้ ทัศนคติที่ไม่ดีต่อฝ่ายบริหาร เช่น ทุกคนไม่ได้รับความเป็นธรรมจากฝ่ายบริหาร หรือ การวางระเบียบในเรื่องนั้น ๆ เป็นการที่ผู้บังคับบัญชาเอาเปรียบเจ้าหน้าที่ผู้ปฏิบัติงาน ล้วนแต่เป็นอุปสรรคต่อการบริหารงานทั้งสิ้น ทัศนคติเหล่านี้อาจจะขาดเหตุผล หรือขัดแย้งต่อความเป็นจริง แต่ตราบดีเท่าที่ผู้ปฏิบัติงานทั้งหลายยังมีทัศนคติที่ไม่ดีต่อฝ่ายบริหารหรือผู้บังคับบัญชา ตราบนั้นก็จะมีผู้ปฏิบัติงานคนหนึ่งคนใดกล้ามีทัศนคติที่ขัดต่อทัศนคติของกลุ่มผู้ปฏิบัติงานส่วนใหญ่ การวัดทัศนคติจึงเป็นเรื่องที่นักวิชาการวัดผลบุคคลได้พยายามศึกษาค้นคว้าเป็นพิเศษ แต่ก็เป็นที่น่าเสียดายว่า ผลของการศึกษาค้นคว้าในด้านนี้ยังไม่เพียงพอที่จะทำได้เครื่องมือวัดที่ความเชื่อมั่น และความเที่ยงตรงสูงพอที่จะใช้ในการบริหารงานบุคคลได้อย่างเป็นที่พอใจเหมือนเครื่องมือวัดในด้านความสามารถของบุคคล ถึงกระนั้นก็ตาม การศึกษาค้นคว้าในเรื่องการวัดทัศนคติได้ทำให้ทราบเทคนิคหลายอย่างที่มีลักษณะพิเศษ ไปกว่าการวัดบุคลิกภาพด้านอื่น จึงได้นำมาอธิบายแยกเป็นอีกบทหนึ่งต่างหากในขั้นนี้

ความหมายของทศนคติ

ทศนคติเป็นตัวแปรเชิงสมมติฐาน (hypothetical variable) เพราะทศนคติวัดไม่ได้จากการกระทำ หรือพฤติกรรมอันหนึ่งอันใดของบุคคล แต่ทศนคติเป็นนามธรรมที่ใช้เรียกกลุ่มของการกระทำหรือพฤติกรรมหลายอย่าง นักจิตวิทยาได้นิยามทศนคติไว้ต่าง ๆ กัน เช่น Allport นิยามทศนคติว่าเป็นสภาพทางจิต และประสาทที่ทำให้บุคคลพร้อมที่จะสนองต่อวัตถุและสถานการณ์ต่างๆ Fuson นิยามว่า ทศนคติคือ ความน่าจะเป็นของการที่พฤติกรรมซึ่งนิยามแล้วจะเกิดขึ้นในสถานการณ์ซึ่งนิยามแล้ว แต่คำนิยามของ Guilford¹ นับว่ามีความกะทัดรัด และเข้าใจได้ง่ายกว่าคำนิยามของผู้อื่น ตามความเห็นของ Guilford ทศนคติคือ การที่บุคคลโน้มเอียงในทางที่ชอบ เห็นด้วย สนับสนุน หรือไม่ชอบ ไม่เห็นด้วย ไม่สนับสนุนการกระทำ หรือมีสิ่งที่มีอยู่ประเภทหนึ่งประเภทใดในสังคม การกระทำในสังคมได้แก่ขนบธรรมเนียมและวิธีการปฏิบัติต่างๆ เช่น ระบอบการปกครอง วิธีการในทางศาล หรือการกีฬาประจำชาติ เป็นต้น สิ่งที่มีอยู่ในสังคม มี 2 แบบใหญ่ ๆ คือ สถาบันทางสังคมต่างๆ และกลุ่มสังคมทั้งหลาย สถาบันทางสังคมมีหลายอย่าง เช่น วัด โรงเรียน รัฐบาล เป็นต้น กลุ่มสังคมได้แก่ กลุ่มเชื้อชาติ กลุ่มสัญชาติ และกลุ่มบุคคลที่รวมกันเป็นชั้นวรรณะ หรืออาชีพต่างๆ

บุคคลแต่ละคนเกิดมีทศนคติต่อการกระทำหรือสิ่งที่มีอยู่ในสังคม โดยอาศัยความเชื่อประจำตนประกอบกับความรู้สึกรู้ตัว ในการพิจารณาทศนคติของบุคคล ได้มีนักวิชาการหลายรายถือเอาความคิดเห็นของแต่ละคนเป็นดัชนีวัดระดับทศนคติของบุคคลนั้น ๆ ความคิดเห็นเป็นการที่บอกว่าบุคคลมีความเชื่อหรือมีความรู้สึกรู้ตัวต่อสิ่งหนึ่ง ๆ อย่างไร ความคิดเห็นนี้อาจมีสัมพันธสัมพันธ์กับทศนคติในบางกรณี แต่ไม่จำเป็นต้องเสมอไป เช่น ผู้ที่เคร่งศาสนาบางคนอาจประกาศตนในที่สาธารณะว่า การพนันเป็นอภัยมูข แต่โดยความเป็นจริงก็ยังรักสนุกที่จะเล่นไพ่หรือซออ๊วสลากกินแบ่งรัฐบาล La Pierre² ได้ทำการวิจัยในปี 1934 เรื่อง ทศนคติของชาวอเมริกันต่อชาวจีนในสหรัฐ ฯ ผู้วิจัยได้เดินทางไปกับสามีภรรยาชาวจีนคู่หนึ่ง ในระหว่างทางได้พักแรมตามโรงแรม 66 แห่ง และรับประทานอาหารตามร้านต่างๆ อีก 184 แห่ง มีสถานที่พักแรมเพียงแห่งเดียวที่ไม่ยอมต้อนรับคณะผู้เดินทาง เมื่อสิ้นสุดการเดินทางแล้ว ผู้วิจัยได้ทอดแบบสอบถามไปยังเจ้าของกิจการเหล่านั้นทุกแห่ง เพื่อขอทราบว่ามีสถานที่เหล่านั้นให้บริการแก่ชาว

¹ Guilford, J.P., *Personality*, p. 223.

² *Ibid.*, pp. 223-224.

จีนหรือไม่ ปรากฏผลว่ามีเจ้าของโรงแรมและเจ้าของร้านอาหารในอัตราร้อยละ 93 และ 92 ตามลำดับ ที่ระบุว่าไม่ให้บริการแก่ชาวจีน ในขณะที่เดียวกันผู้วิจัยได้ทดลองกับกลุ่มที่ควบคุมสภาพการณ์ (control group)³ อีกด้วย ปรากฏผลการแสดงความรังเกียจชาวจีนในอัตรากว้างไกลเหมือนกัน จากผลการวิจัยนี้แสดงให้เห็นว่า การแสดงความคิดเห็นกับทัศนคติจริง ๆ นั้นอาจขัดกันก็ได้ การที่เป็นเช่นนั้นก็เพราะพฤติกรรมบางอย่างที่บุคคลต้องแสดงให้สังคมเห็น อาจเป็นพฤติกรรมที่บุคคลต้องทำในลักษณะที่สอดคล้องกับบุคคลอื่นในสังคม ทั้ง ๆ ที่บุคคลเหล่านั้นมีทัศนคติต่าง ๆ กันไป หรือมีเจตนาที่อาจเป็นเพราะบุคคลเหล่านั้นมิได้มีความยึดมั่นอย่างแท้จริงในความคิดเห็นที่แสดงออก เมื่อถึงเวลาแสดงพฤติกรรมก็อาจแสดงพฤติกรรมไปในอีกทางหนึ่ง ซึ่งไม่จำเป็นต้องเป็นในลักษณะทำนองเดียวกับความคิดเห็นที่มีอยู่ อาจเป็นไปได้ว่าบุคคลจะต้องมีความยึดมั่นอย่างเหนียวแน่นจึงจะแสดงพฤติกรรมที่เปิดเผยอย่างสอดคล้องกับความคิดเห็นที่แสดงออกไป

ถึงแม้จะเป็นที่เห็นพ้องกันในหมู่นักวิชาการแล้วว่าทัศนคติคืออะไร แต่เมื่อแยกวิเคราะห์ทัศนคติออกเป็นประเภท (type) หรือลักษณะนิสัยขั้นปฐม (primary traits) แล้วก็ยังประสบปัญหาในการกำหนดเนื้อหาของละเอียดของคุณสมบัติที่จะต้องมิในทัศนคติประเภทหนึ่ง ๆ ฉะนั้น ในที่นี้จะได้นำทัศนคติในลักษณะต่าง ๆ ตามที่ได้รายงานไว้ใน การวิจัยของต่างประเทศมาประมวลพอเป็นตัวอย่างไว้ในที่นี้⁴

เสรีนิยม และอนุรักษนิยม (liberalism and conservatism) เสรีนิยมเป็นแนวโน้มที่จะชอบหรือสนับสนุนการเปลี่ยนแปลงทางสังคม เศรษฐกิจ การเมือง และศิลปกรรม บุคคลที่มีทัศนคติแบบเสรีนิยม เชื่อในการปฏิรูป สนับสนุนการคุมกำเนิด การหย่าร้าง ชอบสิ่งใหม่ ๆ บุคคลที่มีวัยสูงกว่า 40 ปี มักจะมีทัศนคติแบบอนุรักษนิยมชอบเห็นด้วยกับประชาชนหนุ่มมาก พวกเขาเสรีนิยมเป็นบุคคลที่ทราบความเคลื่อนไหวในปัจจุบันเป็นอย่างดี ส่วนพวกอนุรักษนิยมสนใจเรื่องราวเกี่ยวกับ การเงิน การคลัง และนโยบายของรัฐบาล พวกเขาเสรีนิยมยกย่องนักวิทยาศาสตร์ นักค้นคว้า นักเขียน กวี และรัฐบุรุษ นักศึกษามหาวิทยาลัยที่มีทัศนคติแบบเสรีนิยมชอบเรียนนิเวศวิทยา วารสารศาสตร์ สังคมสงเคราะห์

³ กลุ่มควบคุมสภาพการณ์ คือ กลุ่มทดลองที่ผู้วิจัยควบคุมสภาพการณ์ให้มีลักษณะคล้ายกับสภาพการณ์ปกติ โดยไม่ต้องนำตัวแปรไปทดลอง ในที่นี้หมายถึงกลุ่มโรงแรมและร้านอาหารที่ผู้วิจัยไม่ได้ลองนำสามิภรรยาชาวจีนไปพักแรม และรับประทานอาหารจริง ๆ

⁴ Thurstone, L.L., *The Measurement of Values*, pp. 219-221.
Laski, Harold J., *A Grammar of Politics*, pp. 220-231.

นิติศาสตร์ และเกษตรศาสตร์ ในขณะที่พวกอนุรักษนิยมยกย่องนักการทหาร นักกีฬา นักการคลัง นักอุตสาหกรรม และดาราบันเทิง นักศึกษามหาวิทยาลัยที่มีทัศนคติแบบอนุรักษนิยม ชอบเรียนวิชาการธนาคาร ดุริยางคศิลป์ ทัศนแพทย์ และงานอื่นๆ ที่ใช้ปฏิบัติในหน่วยราชการ นักศึกษาที่มีความสามารถสร้างสรรค์มักจะมีทัศนคติแบบเสรีนิยมมากกว่าอนุรักษนิยม

ศาสนนิยม (religionism) ทัศนคตินี้หมายถึงความเชื่อในคริสต์ศาสนา กล่าวคือมีศรัทธาในพระเจ้า ไม่เชื่อในการปฏิรูปของสัตว์และสิ่งที่มีชีวิต ไม่สนับสนุนการคุมกำเนิด จากลักษณะดังกล่าวจะเห็นว่า ทัศนคติแบบศาสนนิยมมีความเกี่ยวข้องกับทัศนคติแบบอนุรักษนิยมอยู่มาก ในบรรดาผู้ที่ทัศนคติแบบศาสนนิยมส่วนใหญ่เป็นสตรี ชุมชนที่มีขนาดยังเล็กจะยังมีบุคคลที่มีทัศนคติแบบศาสนนิยมมากขึ้นตามส่วน นักศึกษามหาวิทยาลัยที่เคร่งศาสนาจะเลือกเรียนวิชาศาสนศาสตร์ แพทยศาสตร์ ดุริยางคศิลป์ และครุศาสตร์ ส่วนพวกที่เคร่งน้อยกว่าจะเลือกเรียนวิชาการบิน วารสารศาสตร์ นิติศาสตร์ และพาณิชย์ศาสตร์

ชาตินิยม (nationalism) บุคคลที่มีทัศนคติแบบชาตินิยมสูง เป็นผู้ที่เชื่อในความสำคัญของการรักชาติ และเห็นคุณค่าของการทำสงคราม บุคคลประเภทนี้เชื่อว่ารัฐต้องมีอำนาจในการควบคุมสิ่งต่างๆ และเห็นว่า ลัทธิคอมมิวนิสต์เป็นภัยต่อประเทศ ผู้มีชาตินิยมสูงสนับสนุนให้รัฐมีกองทัพเรือขนาดใหญ่ และเห็นควรให้รัฐใช้อำนาจเด็ดขาดในการควบคุมการค้ากับต่างประเทศ บุคคลเหล่านี้เห็นว่าการเดินขบวนนัดหยุดงาน การปล่อยให้ชาวต่างชาติเข้าเมืองโดยเสรี และการใช้ของที่ผลิตจากต่างประเทศ เป็นความผิด และเป็นเรื่องเสียหายต่อประเทศ

หลักการของเครื่องวัดช่วงเท่ากัน

Thurstone เป็นบุคคลแรกที่ยืนยันว่า ทัศนคติอาจวัดได้โดยการวิเคราะห์ความคิดเห็นที่เจ้าตัวถือว่าเป็นความคิดเห็นของตน ในปี 1928 Thurstone⁵ ได้เสนอวิธีการวัดทัศนคติซึ่งมีชื่อว่า วิธีวัดช่วงเท่ากัน (method of equal-appearing intervals) Thurstone ให้เหตุผลในการที่ใช้ความคิดเห็นเป็นแหล่งวัดทัศนคติว่า เพราะมีสัมประสิทธิ์สหสัมพันธ์ในเชิงปฏิสัมพันธ์ระหว่างสิ่งที่บุคคลกล่าวเกี่ยวกับเรื่องหนึ่งๆ และการที่บุคคลจะปฏิบัติเกี่ยวกับเรื่องนั้น จริงอยู่ สัมประสิทธิ์สหสัมพันธ์ในเรื่องนี้ย่อมไม่มีค่าสูง

⁵ Thurstone, *op. cit.*, pp. 215-233.

การทำจะสร้างความเชื่อมั่นในการวัดทำได้โดยใช้ความคิดเห็นหลาย ๆ ครั้งประกอบกัน และการสร้างสัมประสิทธิ์สหสัมพันธ์ระหว่างความคิดเห็นกับการปฏิบัติ ก็ทำได้โดยนำเหตุการณ์เกี่ยวกับการปฏิบัติหลาย ๆ ตอนมาเทียบกัน ในการสร้างเครื่องวัดช่วงเท่ากัน (equal-appearing interval scales) กระทำโดยอาศัยสมมติฐาน 4 ประการดังต่อไปนี้

1. ทศนคติของบุคคลในแต่ละเรื่อง เป็นช่วงของความชอบที่แยกเป็นส่วน ๆ ไม่ได้
2. ความคิดเห็นที่แสดงออกเป็นดัชนีของระดับทศนคติ
3. ความคิดเห็นในแต่ละเรื่องของแต่ละบุคคล ซึ่งได้ว่าบุคคลมีทศนคติในระดับใดในช่วงของความชอบ ฉะนั้น ความคิดเห็นนั้นจึงต้องกำหนดค่าได้ในระดับใดในช่วงของความชอบ
4. ระดับของทศนคติในช่วงของความชอบ ได้แก่ ระดับในเกณฑ์เฉลี่ยของความความคิดเห็นที่แสดงออก ความคิดเห็นแต่ละข้อของบุคคลคนเดียวกันย่อมมีค่าในช่วงของความชอบใกล้เคียงกัน

ขั้นตอนในการกำหนดค่า ขั้นแรกผู้สร้างเครื่องวัดจะต้องประมวลความคิดเห็นทั้งหมดในเรื่องเกี่ยวกับทศนคติในด้านที่จะวัด เพื่อให้คลุมช่วงของความชอบในเรื่องนั้น ๆ ทั้งช่วง เช่น สมมติว่า ต้องการสร้างเครื่องวัดทศนคติของบุคคลที่มีต่อศาสนา ผู้สร้างเครื่องวัดจะต้องถามความคิดเห็นของบุคคลฝ่ายต่าง ๆ ที่มีในเรื่องนี้ และประมวลความคิดเห็นที่ต่าง ๆ กันนั้นให้มากที่สุด ถ้ายังขาดความคิดเห็นที่สุดท้ายสุดขว ผู้สร้างก็อาจต้องเพิ่มเติมความคิดเห็นในลักษณะที่ไม่สนใจเรื่องนี้่อีกในกรณีที่ยังขาดอยู่ ความเห็นในแต่ละเรื่องอาจมีจำนวนถึง 200 ข้อ เพื่อให้คลุมช่วงความชอบทั้งช่วง

ขั้นต่อไปต้องมีการคัดเลือกข้อความความคิดเห็นที่รวบรวมได้ ข้อความคิดเห็นที่จะคงไว้ต้องสั้นตรงประเด็น และมีความชัดเจนที่ทำให้ผู้อ่าน ผู้ฟัง สามารถแสดงความคิดเห็นด้วยหรือคัดค้านได้ ข้อความคิดเห็นที่เลือกไว้จะต้องสัมพันธ์กับจุดหนึ่งจุดใดภายในช่วงความชอบ ข้อความคิดเห็นที่ควรตัดทิ้งได้แก่ ข้อความคิดเห็นที่มีความหมายสองง่าม ข้อความที่กำกวม ข้อความที่บุคคลที่เห็นตรงข้ามก็อาจจะยอมรับได้ ข้อความที่เป็นความจริงอยู่แล้ว และข้อความที่บุคคลทุกคนทุกฝ่ายเห็นด้วย หรือที่จะไม่มีใครเห็นด้วยเลย ยกตัวอย่าง เช่น ศาสนาทำให้คนเป็นคนดี แต่ก็ทำให้คนเฉื่อยชาไม่ขวนขวายต่อสู้ในชีวิต ข้อความเช่นนี้มีความเห็นถึง 2 ประการอยู่ด้วยกัน หรือคนเคร่งศาสนาเป็นคนไม่ชอบผู้หญิง ข้อความประโยคหลังนี้ไม่ชัดเจนว่า ถ้าเป็นผู้หญิงแล้วจะรู้สึกเป็นอย่างไร หรือ มนุษย์นับถือศาสนามาหลายร้อยปีแล้ว ข้อความนี้เป็นเรื่องที่

บุคคลทั่วไปจะต้องเห็นว่า เป็นความจริง ไม่เป็นข้อความที่จะต้องให้ความเห็นชอบหรือคัดค้านกับความเชื่อแต่อย่างใด

การดำเนินงานขั้นที่สาม คือ การให้ค่าเป็นปริมาณแก่ความคิดเห็นแต่ละข้อ ผู้สร้างเครื่องวัดต้องแยกเขียนหรือพิมพ์ความคิดเห็นแต่ละข้อในบัตรแต่ละใบ เมื่อจัดบัตรเหล่านี้เสร็จแล้วก็ขอให้บุคคลอื่นหลาย ๆ คนช่วยกันตัดสินว่า ควรจะจัดเรียงลำดับตามลักษณะของความเชื่อหรือการสนับสนุนเรื่องนั้น ๆ จากนั้นไปหามาได้อย่างไร ในการจัดลำดับนี้ ผู้สร้างเครื่องวัดอาจจะบอกความคิดเห็นในข้อที่มีลักษณะเป็นกลาง ๆ ไว้ด้วยก็ได้ ถ้ามี เช่น ข้อความที่ว่า การมีศาสนาอยู่ในโลกนี้หรือไม่ ไม่เป็นสิ่งสำคัญสำหรับข้าพเจ้า ข้อความที่มีลักษณะเป็นกลางนี้จะช่วยให้คณะผู้ตัดสินสามารถจัดประเภทข้อความที่ชอบมากและน้อยไม่เท่ากันออกจากกันได้ ตามปกติ ก็อาจจัดความคิดเห็นในข้อต่าง ๆ นี้ได้เป็นจำนวนประมาณ 11 ประเภท โดยให้แต่ละประเภทมีความต่างกันในระดับความชอบในเรื่องนั้น ๆ เท่ากัน การจัดประเภทเช่นนี้ เป็นการจัดตามวิธีจิตวิทยากายภาพ (psychophysical method) คณะผู้ตัดสินประกอบด้วยบุคคลเป็นจำนวนประมาณ 25 คน กันบ้างว่าเพียงพอแก่การดำเนินการ ได้มีการวิจัยในที่ต่าง ๆ หลายแห่ง ซึ่งพิสูจน์แล้วว่า⁶ ความคิดเห็นส่วนตัวของผู้ตัดสินมิได้ทำให้การจัดประเภทนั้นบิดเบือนไป เช่น ผลงานวิจัยในเรื่องความรักชาติ ของ R. Pintner และ G. Forlano ในปี 1937 และเรื่องความเกลียดชังชาวยิว ของ H.J. Eysenck และ S. Crown ในปี 1949 เป็นต้น แต่ผลการตัดสินนี้จะเปลี่ยนไปตามยุคสมัย จากผลการวิจัยของ P.R. Farnsworth ในเรื่องทัศนคติต่อการสงครามในปี 1943 ปรากฏว่า ความคิดเห็นที่เคยจัดประเภทไว้ในทศวรรษ 1930 ว่ามีลักษณะรุนแรงใส่สงคราม กลับถูกจัดให้อยู่ในประเภทที่มีลักษณะใฝ่สันติ ในปี 1940-1941

ผลการจัดประเภทของคณะผู้ตัดสิน ทำให้สามารถคำนวณคุณสมบัติทางสถิติของความคิดเห็นแต่ละข้อได้ ค่ามัธยยะของพิสัยตรงกลาง (semi-interquartile range) ทำให้ทราบลักษณะการกระจายของความคิดเห็นแต่ละข้อที่มีน้ำหนักความชอบในระดับต่าง ๆ แต่ละระดับ ความคิดเห็นในข้อที่มีค่านี้สูงจะต้องได้รับการตัดทิ้งเสีย เพราะแสดงว่าเป็นข้อที่มีความหมายไม่ชัดเจน นอกจากนี้ยังจะต้องคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างข้อ (item intercorrelations) และสัมประสิทธิ์สหสัมพันธ์ร่วมของทุกข้อ (item-total correlations) อีกด้วย เพื่อให้ทราบว่า ความคิดเห็นแต่ละข้อ มีความตรงประเด็น

⁶ Guilferd, J.P., *Psychometric Methods*, p. 458.

กับทัศนคติในลักษณะที่ต้องการวัดหรือไม่ จำนวนข้อที่เลือกไว้จะมีประมาณ 25 ข้อ โดยจัดให้แต่ละข้อมีความห่างกันในระดับความชอบเป็นปริมาณเสมอกัน หรือถ้าเป็นไปได้ก็อาจเลือกความคิดเห็นไว้เป็น 2 ชุด ๆ ละ 25 ข้อ เพื่อจะได้ใช้สลับกัน เมื่อเลือกไว้เรียบร้อยแล้ว ก็นำข้อต่างนี้มาเรียงโดยการสุ่ม (random order) ไม่เรียงตามลำดับในช่วงของความชอบ การที่เรียงข้อเช่นนี้ก็เพื่อไม่ให้ผู้ถูกวัดเกิดความลำเอียงในการตอบ

Thurstone ได้สร้างเครื่องมือวัดทัศนคติแบบเครื่องวัดช่วงเท่ากันไว้หลายเรื่อง เช่น ทัศนคติต่อสงคราม ทัศนคติต่อศาสนาคริสต์ ทัศนคติในเรื่องพระผู้เป็นเจ้า ทัศนคติในการลงโทษอาชญากร ทัศนคติต่อการคุมกำเนิด ทัศนคติต่อลัทธิคอมมิวนิสต์ เป็นต้น ตัวอย่างข้างล่างนี้เป็นคำถามในเรื่องทัศนคติต่อกีฬาฟุตบอลในสถาบันอุดมศึกษา ซึ่งกำหนดให้มีช่วงวัด 9 ช่วง

- 8.1 ฟุตบอลเป็นกิจกรรมสำคัญของมหาวิทยาลัย
- 7.1 ทีมฟุตบอลที่เก่งจะทำให้วิทยาลัยมีชื่อเสียงทั่วประเทศ
- 6.0 ฟุตบอลทำให้บรรดานักศึกษาของวิทยาลัยรวมกันเป็นอันหนึ่งอันเดียวกัน
- 5.0 ฟุตบอลเป็นทางให้ระบายความคิดเห็นที่รุนแรงออก
- 3.5 ฟุตบอลมีความสำคัญรองลงมาจากการเล่น
- 2.0 นักฟุตบอลได้รับประโยชน์จากการที่การเล่นน้อยมากจนไม่คุ้มกับเวลาที่เสียไป
- 1.2 ฟุตบอลเป็นวิธีการหาเงินในวิทยาลัยเท่านั้นเอง
- 0.5 ข้าพเจ้าเชื่อว่าสมควรเลิกกีฬาฟุตบอลในวิทยาลัยเสีย

การให้คะแนนผู้ถูกวัด เมื่อได้เรียบเรียงความคิดเห็นเกี่ยวกับทัศนคติในเรื่องที่จะวัดเรียบร้อยแล้ว และประสงค์จะวัดทัศนคติของผู้ใดในเรื่องนั้น ก็นำรายการความคิดเห็นที่เลือกไว้มาให้ผู้ถูกวัดเลือก ผู้ถูกวัดจะต้องเลือกข้อที่สะท้อนหรือมีลักษณะใกล้เคียงกับความเห็นของตน ระดับของทัศนคติของแต่ละคนได้แก่อัตราเฉลี่ยของค่าความคิดเห็นที่ผู้ใดได้เลือกไว้จากรายการความคิดเห็นทั้งหมดในเครื่องวัดช่วงเท่ากัน อัตราเฉลี่ยนี้อาจคำนวณได้เป็นค่าเฉลี่ย (mean) หรือค่ามัธยะ (median) ก็ได้ แต่ I. Lorge⁷ ได้ทำการวิจัยและค้นพบว่า ค่ามัธยะมีความเชื่อถือได้มากกว่าค่าเฉลี่ย ในกรณีนี้ คะแนนทัศนคติที่เป็นอัตราเฉลี่ยของค่าความคิดเห็นนี้อาจมีทางบกพร่องได้ในกรณีที่ไม่จำกัดจำนวนข้อความคิดเห็นที่ผู้ถูกวัดจะเลือกได้ จริงอยู่ ผู้ถูกวัดย่อมเลือกเฉพาะข้อที่มีลักษณะละม้ายคล้ายคลึงกับความเห็นของตนเท่านั้น แต่จำนวนข้อที่เลือกโดยไม่

⁷ Ibid., p. 459.

จำกัดจะทำให้พิสัยของคะแนนที่ได้บิดเบือนไป ผู้ที่เลือกไว้หลายข้อจะได้คะแนนเฉลี่ยที่ใกล้เคียงกับค่าที่กำหนดไว้เป็นกลางๆ โดยเฉพาะในกรณีของบุคคลที่มีทัศนคติรุนแรงสุดขั้วสุดขั้ว จำนวนข้อที่เลือกมากจะทำให้ได้อัตราเฉลี่ยของค่าความคิดเห็นที่ไม่ตรงกับความเป็นจริง ฉะนั้นในการใช้เครื่องวัดช่วงเท่ากันนี้ จึงมักกำหนดให้ผู้ถูกวัดเลือกแต่เพียง 3 ข้อ หรือ 5 ข้อ เท่านั้น การกำหนดจำนวนจำกัดนี้ จะทำให้ผู้ถูกวัดตั้งใจเลือกเฉพาะความคิดเห็นที่ใกล้เคียงกับของตนจริงๆ เท่านั้น

หลักการของเครื่องวัดค่ารวม

การวัดทัศนคติด้วยเครื่องวัดค่ารวม (summative scales or summated rated scales) อาศัยรูปแบบค่ารวม (summative model) เป็นหลัก เนื่องจากหลักการนี้ R.A. Likert เป็นผู้ยืนยันว่าดีที่สุดในการใช้วัดทัศนคติ จึงเรียกกันว่า Likert scales เช่นเดียวกับหลักการของเครื่องวัดช่วงเท่ากัน ซึ่งเรียกกันว่า Thurstone scales เพราะ L.L. Thurstone เป็นผู้อ้างว่าดีที่สุดในการใช้วัดทัศนคติ รูปแบบค่ารวมมีสมมติฐานใหญ่ 3 ประการ ประการแรก การสนองตอบต่อคำถามหรือรายการแต่ละข้อในเครื่องวัดจะมีลักษณะคงที่ (monotonic trace line) แต่ทั้งนี้ ไม่ได้หมายความว่าลักษณะคงที่ของการสนองตอบในทุกๆ ข้อจะต้องเป็นเส้นทับกัน ประการที่สอง ผลรวมของลักษณะคงที่ของการสนองตอบ (sums or averages) ในข้อทั้งหมดของแต่ละบุคคลจะมีลักษณะเป็นเส้นตรงหรือเกือบเป็นเส้นตรง (linear) เพราะถึงแม้ว่าลักษณะคงที่ในทุกๆ ข้อจะไม่เป็นเส้นทับกัน แต่เมื่อนำค่าคงที่นั้นมารวมกันแล้วจะทำให้ส่วนที่เอนจากเส้นตรงหักลบกันไป และประการสุดท้าย ผลรวมของลักษณะคงที่ของการสนองตอบในข้อหนึ่งๆ จะมีองค์ประกอบร่วมกันอยู่หนึ่งตัวนั้น คือ ผลรวมนี้ แทนค่าลักษณะนิสัยที่จะวัดได้อย่างหนึ่งๆ เพียงอย่างเดียว จากหลักการของรูปแบบค่ารวมที่ถือว่า การสนองตอบต่อคำถามหรือรายการแต่ละข้อมีความสัมพันธ์ในเชิงลักษณะคงที่ (monotonically related) กับลักษณะนิสัยแต่ละอัน และผลรวมของการสนองตอบสัมพันธ์กับลักษณะนิสัยนั้นๆ ในเชิงเส้นตรงเช่นนี้ Likert จึงสามารถวัดทัศนคติในเรื่องหนึ่งๆ โดยอาศัยหลักการรวมคะแนนทั้งหมดที่บุคคลได้จากการสนองตอบคำถามทุกๆ ข้อเข้าด้วยกัน

การสร้างเครื่องวัดค่ารวม วิธีการในการสร้างเครื่องวัดนี้มีลักษณะคล้ายกับการสร้างแบบทดสอบทั่วไป เครื่องวัดค่ารวมประกอบด้วยคำถามแบบเลือกหลาย ๆ ข้อ (multiple-choice type) กล่าวคือ ในคำถามแต่ละข้อ ผู้ถูกวัดอาจเลือกคำตอบข้อหนึ่งจาก 3 ข้อ หรือ 5 ข้อ หรือ 7 ข้อ คำตอบที่กำหนดให้เลือกอาจเป็น การเห็นด้วย การไม่เห็นด้วย และการไม่มีความเห็น หรืออาจเป็นการเห็นด้วยทั้งหมด การเห็นด้วยบ้าง การไม่มี

ความเห็น การเห็นด้วยบางส่วน และการไม่เห็นด้วยทั้งหมด คำถามแต่ละข้อเป็นคำกล่าวในประเด็นที่เกี่ยวกับทัศนคติที่จะวัด ตามธรรมชาติเครื่องวัดแต่ละชุดควรมีคำถามนี้ประมาณ 40 ข้อ คำกล่าวที่ตั้งเป็นคำถามนี้ ควรเป็นความคิดเห็นที่ไม่รุนแรง และเป็นความคิดเห็นทั้งประเภทที่เห็นชอบและไม่เห็นสอดคล้องด้วยกับเรื่องที่เกี่ยวข้องกับทัศนคติที่จะวัด ความคิดเห็นที่เป็นกลางไม่ควรตั้งเป็นคำถาม เพราะทำให้ไม่สามารถวัดทัศนคติในเรื่องนั้น ๆ ได้ ความคิดเห็นที่รุนแรงก็เช่นเดียวกันไม่สามารถทำให้ทราบความแปรปรวนในการวัด ความคิดเห็นทั้งประเภทที่เห็นชอบ และไม่เห็นสอดคล้องอย่างปกติจะทำให้ได้ค่าแปรปรวนเป็นทางให้ทราบทัศนคติในเรื่องนั้น ความคิดเห็นที่ตั้งไว้เป็นคำถาม ควรจะมีประเภทที่เห็นชอบเป็นปริมาณก้ำกึ่งกับประเภทที่ไม่เห็นสอดคล้อง การเปิดโอกาสให้ผู้ถูกวัดเลือกแสดงความเห็นหรือไม่เห็นด้วยกับความคิดเห็นทั้งประเภทที่เห็นชอบ และไม่เห็นสอดคล้องกับเรื่องที่เกี่ยวข้องกับทัศนคติเรื่องหนึ่ง ๆ นี้ ทำให้ผู้ถูกวัดสามารถแสดงน้ำหนักความเชื่อของตนในเรื่องนั้น ๆ ได้อย่างเต็มที่ แต่อย่างไรก็ดี ค่าความแปรปรวนของความเชื่อที่ได้อาจเบี่ยงเบนไปจากความจริง เพราะมีค่าความแปรปรวนของการตอบมาเกี่ยวข้องด้วย กล่าวคือ บุคคลแต่ละคนมีแนวโน้มในการตอบคำถามในลักษณะต่างๆ กัน บางคนชอบตอบอย่างรุนแรง หรือบางคนชอบตอบเป็นกลาง ๆ ค่าความแปรปรวนของการตอบอาจทำให้ไม่ทราบความแปรปรวนของทัศนคติอย่างแท้จริง Cronbach และ Guilford⁸ มีความเห็นพ้องกันว่า ค่าความแปรปรวนของการตอบคำถามชนิดนี้เกิดขึ้นจากสาเหตุใหญ่ ๆ 6 ประการ คือ ประการแรก ความรู้สึกในเรื่องการเล็ง บางคนพยายามตอบเป็นกลาง ๆ ไว้ เพื่อป้องกันการเสียหาย ประการที่สอง ความเข้าใจไม่ตรงกันในความหมายของภาษา บุคคลหนึ่งอาจเข้าใจความหมายของการเห็นด้วยบ้าง ไม่ตรงก็บอื่บุคคลหนึ่ง ประการที่สาม แรงจูงใจในการตอบ (impulsion) การให้เลือกตอบทำให้บุคคลสักแต่ว่าหากคำตอบลงไป ประการที่สี่ การยอมรับ (acquiescence) บุคคลที่มีลักษณะชอบเห็นด้วยกับบุคคลอื่น จะเลือกคำตอบประเภทที่เห็นสอดคล้องกับเรื่องที่ถาม ได้ถูกต้องมากกว่าประเภทที่ไม่เห็นสอดคล้องกับเรื่องที่ถาม ประการที่ห้า ปัญหาในเรื่องเวลาที่ให้ตอบ เวลาที่กำหนดไว้ ทำให้การตอบอาจไม่ถี่ถ้วน และประการสุดท้าย ความรู้สึกที่จะซ่อนเร้น (falsification) ผู้ถูกวัดย่อมพยายามที่จะแสดงแต่ลักษณะที่ดีของตน และพยายามปิดบังส่วนเสียของบุคลิกลักษณะ สภาพการณ์เช่นนี้ เป็นทางให้ผู้ถูกวัดเลือกคำตอบที่อาจจะไม่แสดงลักษณะที่แท้จริงของตน

⁸ Kerlinger, Fred N., *Foundations of Behavioral Research*, pp. 484—485.

การสร้างเครื่องวัดค่าร่วมเพื่อวัดทัศนคติ อาจทำตามตัวอย่างข้างล่าง ซึ่งยกมาเป็นตัวอย่างสมมติในการวัดทัศนคติของประชาชนต่อการดำเนินงานของรัฐสภา

ผู้ถูกวัดจะต้องให้ความเห็นเกี่ยวกับข้อความต่อไปนี้ โดยระบุว่าเห็นด้วยเป็นที่สุด เห็นด้วยอย่างแท้จริง เห็นด้วย ไม่เห็นด้วย ไม่เห็นด้วยอย่างแท้จริง หรือไม่เห็นด้วยเป็นที่สุด

1. การดำเนินงานของรัฐสภาช่วยส่งเสริมระบอบประชาธิปไตยของชาติ โดยการทำให้ประชาชนมีสิทธิมีเสียงในการปกครองประเทศ

2. กฎหมายที่รัฐสภาได้ตราออกมาถูกต้อง เพราะช่วยส่งเสริมให้การบริหารประเทศเป็นไปตามคัลลองที่ถูกต้อง

การให้คะแนนผู้ถูกวัด คะแนนที่ใช้ในการเลือกตอบว่า เห็นด้วยหรือไม่เห็นด้วย ขึ้นกับจำนวนคำตอบที่ให้เลือก เช่น ถ้าเป็นคำถามที่ต้องเลือกคำตอบในระหว่าง 2 คำตอบ (dichotomous items) ก็อาจให้คะแนนผู้ที่ตอบว่าเห็นด้วย 1 คะแนน และไม่เห็นด้วยเป็นคะแนนศูนย์ แต่ถ้าเป็นคำถามที่ต้องเลือกระหว่าง 6 คำตอบ ดังตัวอย่างข้างบน ก็อาจให้คะแนนเป็น 7 ชั้น คือ เห็นด้วยเป็นที่สุด 7 คะแนน เห็นด้วยอย่างแท้จริง 6 คะแนน เห็นด้วย 5 คะแนน ไม่เห็นด้วย 3 คะแนน ไม่เห็นด้วยอย่างแท้จริง 2 คะแนน ไม่เห็นด้วยเป็นที่สุด 1 คะแนน และให้คะแนน 4 คะแนน แก่ผู้ที่ไม่เลือกตอบข้อใดเลย ในกรณีที่คำถามซึ่งนำมาตั้งเป็นคำถามเป็นประเภทไม่เห็นสอดคล้องกับเรื่องที่เกี่ยวข้องกับทัศนคติที่วัด (negative statements) ก็จะต้องให้คะแนนกลับกัน เช่น ถ้าเป็นคำถามที่ต้องเลือกคำตอบในระหว่าง 2 คำตอบ ผู้ที่เห็นด้วยได้คะแนนเป็นศูนย์ ผู้ที่ไม่เห็นด้วยได้ 1 คะแนน แต่ถ้าเป็นคำถามที่ต้องเลือกระหว่างคำตอบหลายคำตอบ (multipoint items) ก็ให้เอาคะแนนที่กำหนดไว้เดิมไปหักออกจากคะแนนสูงสุดที่เพิ่มชั้นอีก 1 คะแนนแล้ว สำหรับตัวอย่างที่กล่าวมาแล้ว อาจให้คะแนนดังนี้ เห็นด้วยเป็นที่สุด 1 คะแนน เห็นด้วยอย่างแท้จริง 2 คะแนน เห็นด้วย 3 คะแนน ไม่ได้ตอบ 4 คะแนน ไม่เห็นด้วย 5 คะแนน ไม่เห็นด้วยอย่างแท้จริง 6 คะแนน และไม่เห็นด้วยเป็นที่สุด 7 คะแนน เป็นต้น

จากคำอธิบายทั้งหมดนี้จะเห็นได้ว่า แม้เครื่องวัดค่ารวมจะนำเอาความคิดเห็นมาตั้งเป็นหัวข้อในการวัดทัศนคติเช่นเดียวกับเครื่องวัดช่วงเท่ากัน แต่เครื่องวัดประเภทแรกก็มีลักษณะต่างไปจากประเภทหลัง โดยที่มิได้นำความคิดเห็นแต่ละข้อมาจัดช่วงระยะ

ระหว่างกัน จากผลการวิจัยหลายแห่ง⁹ ปรากฏว่าเครื่องวัดค่ารวมสามารถวัดทัศนคติได้แน่นอนกว่าเครื่องวัดช่วงเท่ากัน เพราะในการวัดด้วยเครื่องวัดค่ารวมผู้ถูกวัดจะต้องแสดงความคิดเห็นกับข้อความทุกข้อความที่ปรากฏในเครื่องวัด แต่ในการวัดด้วยเครื่องวัดช่วงเท่ากัน ผู้ถูกวัดอาจไม่แสดงปฏิกิริยาต่อความคิดเห็นบางข้อได้ ลักษณะการเช่นนี้ทำให้ค่าคะแนนที่วัดด้วยเครื่องวัดช่วงเท่ากันมีโอกาสบิดเบือนมากกว่าค่าคะแนนที่วัดด้วยเครื่องวัดค่ารวม

แนวคิวในการวัดทัศนคติ

ในบทที่ 7 เรื่องการวิเคราะห์องค์ประกอบ ได้อธิบายเทคนิคต่าง ๆ ที่ใช้ในการวิเคราะห์ตัวแปร 3 ตัว คือ สภาพการณ์ (circumstances) ตัวบุคคล (persons) และสิ่งประกอบ (attributes) และได้ชี้ให้เห็นว่า แนวคิว (Q technique) เป็นเทคนิคที่ใช้วิเคราะห์ลักษณะทางจิตวิทยาของบุคคล โดยเน้นตัวบุคคลโดยเฉพาะ (ipsative) ปรากฏว่า ในปี 1953 W. Stephenson ได้เสนอรายละเอียดของเทคนิคนี้จนเป็นที่ยอมรับใช้กันอย่างแพร่หลายในการวิเคราะห์ทัศนคติ สาระสำคัญของหลักการนี้ก็คือ การวัดทัศนคติควรกระทำโดยการเปรียบเทียบการสนองตอบในลักษณะต่าง ๆ ของบุคคลซึ่งเป็นเจ้าตัวเอง ไม่ใช่เปรียบเทียบการสนองตอบระหว่างบุคคลต่าง ๆ หรือกล่าวอีกนัยหนึ่งก็คือ ให้วัดทัศนคติโดยการนำผลของการวัดด้วยวิธีต่าง ๆ มาเปรียบเทียบกันเป็นคน ๆ ไป มิใช่ใช้การวัดแต่เพียงวิธีหนึ่งวิธีใดแล้วเปรียบเทียบผลการวัดนั้นระหว่างบุคคลต่าง ๆ บุคคลซึ่งเป็นเจ้าตัวเองที่กำหนดอาจเป็นบุคคลเพียงคนเดียวหรือกลุ่มของบุคคลก็ได้ แต่โดยปกติมักใช้กลุ่มของบุคคลเพื่อจะให้ครอบคลุมพิสัยได้กว้างขวางขึ้น

การที่จะเข้าใจแนวคิวได้ชัดเจนควรจะลองเปรียบเทียบกับการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ตามปกติ หรือมีชื่อเรียกโดยเฉพาะเพื่อให้แตกต่างกับเทคนิคการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ที่พิเศษอื่น ๆ ว่า แนวอาร์ (R technique or R methodology) ในการวิเคราะห์ตามแนวอาร์ ผู้วิจัยใช้เครื่องมือวัดวัดบุคคลหลาย ๆ คน เพื่อรวบรวมข้อมูลมาคำนวณหาสัมประสิทธิ์สหสัมพันธ์ร่วมระหว่างการใช้เครื่องมือวัดแต่ละชุด ผลที่วิเคราะห์จะแสดงค่าของตัวแปรต่าง ๆ แต่ในแนวนั้นผู้วิจัยจะต้องพยายามหาสัมประสิทธิ์สหสัมพันธ์ระหว่างบุคคลแต่ละคน หรือกลุ่มของบุคคลแต่ละกลุ่ม โดยใช้เครื่องมือวัดหลาย ๆ ชุด ผลที่วิเคราะห์ได้จะแสดงชนิดของสิ่งประกอบของแต่ละคนหรือแต่ละกลุ่ม ลักษณะการวิเคราะห์ตามแนวคิว อาจพิจารณาเห็นได้จากตัวอย่างทางคณิตศาสตร์อย่างง่าย ๆ ดังต่อไปนี้

⁹ Guilford, *Personality*, p. 229.

สมมติว่ามีนักการเมืองอยู่ 6 คน ผู้วิจัยอาจตั้งคำถามถามราษฎร 4 ราย ซึ่งประกอบด้วยสุภาพบุรุษ 2 ราย และสุภาพสตรี 2 ราย ว่า ถ้าจะเลือกนักการเมืองเหล่านี้เป็นผู้แทนเพื่อไปบริหารประเทศ จะเลือกผู้ใดตามลำดับก่อนหลังอย่างไร ข้อมูลที่ได้นำมาจัดเป็นตารางได้ดังนี้

	ผู้ชายคนที่ หนึ่ง	ผู้ชายคนที่ สอง	ผู้หญิงคนที่ หนึ่ง	ผู้หญิงคนที่ สอง
นักการเมืองคนที่หนึ่ง	1	2	6	4
นักการเมืองคนที่สอง	2	1	5	6
นักการเมืองคนที่สาม	3	3	4	5
นักการเมืองคนที่สี่	5	4	3	1
นักการเมืองคนที่ห้า	4	5	2	2
นักการเมืองคนที่หก	6	6	1	3

จากข้อมูลสมมติ ปรากฏว่า ผู้ชายและผู้หญิงต่างมีทัศนคติต่อนักการเมืองแต่ละคนไม่เหมือนกัน เช่น ผู้ชายคนที่หนึ่งชอบนักการเมืองคนที่หนึ่งมากที่สุด และชอบนักการเมืองคนที่หกน้อยที่สุด ส่วนผู้ชายคนที่สองชอบนักการเมืองคนที่สองมากที่สุด และชอบนักการเมืองคนที่หกน้อยที่สุด เป็นต้น ลำดับที่ที่ผู้ชายและผู้หญิงแต่ละคนจัดเลือกนักการเมืองต่าง ๆ กันเช่นนั้น ทำให้สามารถคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างลำดับที่ (rank-order coefficients of correlation) ได้โดยสูตร

$$R = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$$

ผลจากการคำนวณจะได้สัมประสิทธิ์สหสัมพันธ์ระหว่าง ผู้ชายและผู้หญิงแต่ละคน ดังแสดงไว้ในตารางข้างล่าง

	ผู้ชายคนที่ หนึ่ง	ผู้ชายคนที่ สอง	ผู้หญิงคนที่ หนึ่ง	ผู้หญิงคนที่ สอง
ผู้ชายคนที่หนึ่ง		0.89	-0.17	-0.19
ผู้ชายคนที่สอง	0.89		-0.17	-0.18
ผู้หญิงคนที่หนึ่ง	-0.17	-0.17		0.87
ผู้หญิงคนที่สอง	-0.19	-0.18	0.87	

จากสัมประสิทธิ์สหสัมพันธ์ระหว่างลำดับที่ที่คำนวณได้ข้างบนนี้ ปรากฏว่า ลำดับที่ของผู้ชายคนที่หนึ่งมีค่าสหสัมพันธ์กับลำดับที่ของผู้ชายคนที่สองในระดับสูง แต่กลับมีค่าสหสัมพันธ์กับลำดับที่ของผู้หญิงทั้งสองคนในเชิงนิเสธ และในทำนองเดียวกัน ลำดับที่ของผู้หญิงทั้งสองคนมีสหสัมพันธ์ต่อกัน แต่ไม่มีสหสัมพันธ์กับผู้ชายทั้งสองคน จากผลของการวิเคราะห์ข้อมูลสมมตินี้ ทำให้ทราบว่า มีบุคคลสองกลุ่มในแง่ความชอบนักการเมือง 6 คน คือ กลุ่มผู้ชาย 2 คน และกลุ่มผู้หญิง 2 คน แต่ละกลุ่มประกอบด้วยบุคคลที่มีความคิดเห็นเกี่ยวกับตัวนักการเมือง ในทำนองเดียวกัน

การจำแนกแบบคิว การที่จะวัดค่าทัศนคติตามเทคนิคคิวได้ ต้องอาศัยการจำแนกแบบคิว (Q sort) เป็นหลัก การจำแนกแบบคิวนี้คือ การกำหนดให้ผู้ถูกวัดเลือกบัตรซึ่งมีข้อความหรือรูปภาพที่เป็นตัวแทนของประเด็นที่เกี่ยวข้องกับทัศนคติในด้านที่จะวัด เช่นบัตรที่มีข้อความต่อไปนี้ในบัตรแต่ละใบเป็นข้อ

บัตรใบที่หนึ่ง : การบริหารเป็นการประสานงาน ผู้ปฏิบัติงานต้องได้รับการสนับสนุนจากผู้บังคับบัญชา ให้มีกำลังใจร่วมกันในการทำงาน

บัตรใบที่สอง : การบรรจุ แต่งตั้ง และเลื่อนชั้นผู้ปฏิบัติงาน ต้องอาศัยการพิจารณาความรู้ ความสามารถ ประสิทธิภาพ และผลงานโดยวิธีอัตโนมัติ

บัตรใบที่สาม : นักบริหารต้องมีลักษณะเป็นผู้นำ ผู้บังคับบัญชาที่ไม่มีลักษณะเป็นผู้นำ ถือว่าเป็นผู้บังคับบัญชาที่ไร้การไม่ได้

บัตรใบที่สี่ : การที่ผู้ปฏิบัติงานแต่ละคนจะปฏิบัติงานได้ดีเพียงใดขึ้นอยู่กับความตั้งใจ อุทิศของของตนเอง

บัตรใบที่ห้า : การกำหนดหน้าที่ความรับผิดชอบไว้เป็นลายลักษณ์อักษรชัดเจนเท่านั้นที่จะช่วยให้การปฏิบัติงานของหน่วยงานลุล่วงไปด้วยดี

บัตรใบที่หก : การทำราชการส่วนใหญ่ในปัจจุบันไม่ค่อยตั้งใจปฏิบัติงาน ก็เพราะได้เงินเดือนต่ำ

หรือถ้าเป็นบัตรที่แสดง รูปภาพ บัตรแต่ละใบก็อาจมีรูปภาพของบุคคลหรือวัตถุในเรื่องที่เกี่ยวข้องกับการวัดทัศนคติ เช่น ถ้าจะวัดทัศนคติต่อรูปปั้นในสมัยต่างๆ ก็อาจมีบัตรแต่ละใบที่มีภาพถาวรรูปปั้น ตั้งแต่สมัยโบราณมาจนถึงสมัยปัจจุบัน ดังนี้เป็นต้น

จำนวนบัตรที่มีให้เลือกนั้น ขึ้นอยู่กับความเหมาะสมทางสถิติ โดยทั่วไปไม่ควรต่ำกว่า 60 ใบ และไม่จำเป็นต้องเกิน 100 ใบ¹⁰

ผู้ถูกวัดมีหน้าที่เลือกกว่า บัตรแต่ละใบซึ่งบรรจุข้อความแต่ละข้อหรือรูปภาพแต่ละรูปนั้น ถูกใจตนมากน้อยเพียงใด ในการเลือกนี้อาจระบุความชอบมากน้อยที่มีต่อบัตรแต่ละใบ เป็นลำดับลดหลั่น 11 ชั้น เพื่อให้มีจำแนกบัตรในลักษณะที่ใกล้เคียงกับการแจกแจงปกติ (normal distribution) เช่น สมมติว่ามีบัตรให้เลือกบัตร 100 ใบก็อาจจัดลำดับได้ดังตัวอย่างข้างล่างนี้

ชอบมากที่สุด	$\frac{2 \ 4 \ 8 \ 12 \ 14 \ 20 \ 14 \ 12 \ 8 \ 4 \ 2}{10 \ 9 \ 8 \ 7 \ 6 \ 5 \ 4 \ 3 \ 2 \ 1 \ 0}$	ชอบน้อยที่สุด
--------------	---	---------------

ก่อนการเลือกบัตร ผู้วัดจะมีการสับบัตรเพื่อให้บัตรทุกใบมีโอกาสได้รับเลือกเท่า ๆ กัน ผู้ถูกวัดจะต้องพิจารณาบัตรทั้งหมดทีละใบ และแปะบัตรทั้งหมดบนโต๊ะใหญ่ ๆ ในการเตรียมการจัดลำดับผู้ถูกวัดมักจะได้รับการขอให้แยกบัตรเป็นประเภทใหญ่ ๆ เสียก่อน คือ ประเภทที่รู้สึกชอบแน่นอน ประเภทที่รู้สึกแน่ใจว่าไม่ชอบและประเภทที่ไม่แน่ใจว่าจะชอบหรือไม่ชอบ การแยกประเภทใหญ่ขึ้นเป็นการช่วยให้จัดลำดับได้สะดวกขึ้น ในการจัดลำดับ ผู้ถูกวัดจะต้องจัดลำดับประเภทที่ชอบมากที่สุดและที่ชอบน้อยที่สุดเสียก่อน แล้วจึงจัดลำดับอื่น ๆ เรียงลงไป จากรูปตัวอย่างข้างต้นจะเห็นว่าค่าคะแนนที่แสดงความชอบมากที่สุด คือ 10 และค่าคะแนนที่ชอบน้อยที่สุด คือ ศูนย์ ค่าคะแนนที่ชอบน้อยลงมาเป็นลำดับลดหลั่นก็มีระหว่าง 9 ถึง 1 ค่าคะแนนเหล่านี้เขียนไว้ใต้บรรทัด ส่วนตัวเลขที่อยู่เหนือบรรทัด คือ จำนวนบัตรที่เลือกไว้ว่า ชอบมาก ชอบน้อยตามความประสงค์ของผู้ถูกวัด ในกรณีของตัวอย่างที่สมมตินี้ ปรากฏว่าผู้ถูกวัดเลือกบัตรที่ชอบมากที่สุดไว้ 2 ใบ และที่ชอบรองลงมาเป็นจำนวน 4, 8, 12, ... ตามลำดับ บัตรที่อยู่ในช่องคะแนน 5 คือ บัตรที่มีลักษณะเป็นกลาง ผู้ถูกวัดไม่แน่ใจว่าจะชอบมากหรือชอบน้อย

ตัวอย่างการจำแนกแบบคิว ผู้วัดจะต้องเอาค่าคะแนนที่เป็นผลลัพธ์จากการเลือกบัตรมาคำนวณสัมประสิทธิ์สหสัมพันธ์ของลำดับที่ในการเลือกระหว่างบุคคลต่าง ๆ เพื่อให้ความสะดวกแก่การวิเคราะห์ตัวเลข จะได้สมมติตัวอย่างง่าย ๆ เป็นการศึกษาเรื่องนี้ สมมติว่า ประสงค์จะศึกษาทัศนคติของบุคคลที่จะเลือกบัตรเพียง 4 คน ผู้เลือกต้องจัดลำดับความชอบเป็น 5 ลำดับ คือ ชอบมากที่สุด 4 คะแนน ชอบรองลงมา 3 คะแนน

¹⁰ Kerlinger, *op. cit.*, p. 583.

ไม่แน่ใจ 2 คะแนน ชอบบ้าง 1 คะแนน ชอบน้อยที่สุด 0 คะแนน สมมติว่าผู้เลือกแต่ละคนได้เลือกบัตรเรียบร้อยแล้วตามวิธีการที่กล่าว และผลปรากฏว่า บัตรแต่ละใบได้ค่าคะแนนตามการเลือกของแต่ละคนดังต่อไปนี้

บัตรที่	คนที่หนึ่ง	คนที่สอง	คนที่สาม	คนที่สี่
1	3	3	1	1
2	4	4	3	2
3	2	2	2	2
4	3	3	2	2
5	2	1	2	2
6	1	2	3	3
7	2	2	2	4
8	0	0	4	3
9	1	1	0	0
10	2	2	1	1

ตัวเลขในสดมภ์ คนที่หนึ่งถึงคนที่สี่ เป็นค่าที่ได้จากการเลือกให้บัตรแต่ละใบอยู่ในลำดับที่ชอบมากชอบน้อยตามความพอใจของแต่ละคน จากการคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างลำดับที่แต่ละคนเลือกตามสูตรที่กล่าวไว้ก่อนหน้านั้น ก็จะได้ตารางสัมพัทธ์ข้างล่างนี้

	คนที่หนึ่ง	คนที่สอง	คนที่สาม	คนที่สี่
คนที่หนึ่ง		0.92	-0.17	-0.17
คนที่สอง	0.92		-0.08	-0.08
คนที่สาม	-0.17	-0.08		0.75
คนที่สี่	-0.17	-0.08	0.75	

การแปลความหมายของตัวเลขที่คำนวณได้ที่อาศัยการพิจารณา 2 ประการ คือ จับกลุ่มผู้ที่มีค่าสหสัมพันธ์ระหว่างกันสูงเข้าด้วยกัน และวิเคราะห์ว่า บัตรที่แต่ละกลุ่ม

เลือกไว้ว่าชอบมากที่สุดนั้น มีข้อความหรือรูปภาพในลักษณะใด จากตารางสัมพันธ
ของตัวอย่างสมมติข้างบน จะเห็นว่า คนที่หนึ่งและคนที่สองเป็นบุคคลกลุ่มเดียวกัน
เพราะมีสัมประสิทธิ์สหสัมพันธ์ของการเลือกจัดลำดับที่สูงถึง 0.92 ส่วนคนที่สามและ
คนที่สี่ ถึงแม้จะมีสัมประสิทธิ์สหสัมพันธ์ของการเลือกจัดลำดับ ที่ไม่สูงเท่า ก็อาจจัด
เป็นกลุ่มเดียวกันได้ เพราะมีลักษณะตรงกันข้ามกับกลุ่มแรก กลุ่มแต่ละกลุ่มนี้อาจตั้ง
ชื่อเรียกเพื่อให้ต่างจากกันได้เลยว่า เป็น กลุ่ม ก. และ กลุ่ม ข. เมื่อต้องการจะทราบ
ว่า กลุ่ม ก. และกลุ่ม ข. มีทัศนคติเป็นอย่างไร และต่างกันประการใด ก็กลับไป
พิจารณาบัตรที่แต่ละกลุ่มเลือกไว้ว่าชอบมากที่สุด สมมติว่า กลุ่ม ก. เลือกบัตรใบที่
หนึ่งถึงที่สาม ที่ยกเป็นตัวอย่างในหัวข้อการจำแนกแบบคือว่า เป็นบัตรที่ชอบมากที่สุด
แต่กลุ่ม ข. เลือกบัตรใบที่สี่ถึงที่หกว่า ชอบมากที่สุด เมื่อพิจารณาเนื้อหาของบัตร
ทั้งหมดแล้วก็จะเห็นได้ว่า กลุ่ม ก. ชอบให้การบังคับบัญชาในทางราชการเป็นไปตาม
หลักการบริหารแผนใหม่ คือ ประสงค์ที่จะให้ผู้บังคับบัญชาเป็นผู้นำ เป็นนักประสาน
งานและการบริหารงานบุคคลเป็นไปอย่างมีมาตรฐานทางวิทยาศาสตร์ ส่วนกลุ่ม ข.
ถือว่าเงินค่าจ้างค่าทดแทนและระเบียบต่างๆ เป็นเรื่องสำคัญในการบริหารงาน ลักษณะ
ที่ต่างกันเช่นนี้ อาจทำให้กล่าวได้ว่า กลุ่ม ก. มีทัศนคติแบบก้าวหน้า (progressive)
ส่วนกลุ่ม ข. มีทัศนคติแบบโบราณ

การกำหนดข้อความในบัตรที่จะให้เลือก ข้อความหรือรูปภาพที่จะบรรจุในบัตรแต่
ละใบสร้างขึ้นได้โดยวิธีการใหญ่ ๆ 2 อย่าง คือ วิธีสุ่ม (random) และวิธีกำหนดโครง
สร้าง (structured) การวัดทัศนคติด้วยเทคนิคนี้ ส่วนใหญ่สร้างข้อความด้วยวิธีสุ่มเป็น
หลัก วิธีการสร้างข้อความด้วยวิธีสุ่มนี้ไม่ใช่เป็นการสุ่มเลือกข้อความด้วยวิธีการสุ่ม
ตัวอย่างที่ใช้ในการวิจัยปกติ การสุ่มตัวอย่างในการวิจัยตามปกตินั้น การที่พยายามคัด
เลือกตัวแทนจริงๆ ของประชากร (universe) ทั้งหมดที่มีกำหนดไว้ตามขอบเขตของ
ปัญหาที่จะทำการศึกษา แต่การสร้างข้อความหรือเลือกรูปภาพด้วยวิธีสุ่มตามเทคนิคนี้
เป็นการประมวลข้อความหรือรูปภาพต่างๆ โดยไม่จำเป็นต้องคำนึงถึงความสัมพันธ์
ระหว่างข้อความหรือรูปภาพเหล่านั้นกับตัวแปรหรือองค์ประกอบตัวหนึ่งตัวใด โดย
หลักการกว้างๆ แล้ว ข้อความหรือรูปภาพต่างๆ ที่ไม่ทำให้เสียความคล้ายคลึงภายใน
กลุ่ม (homogeneity) ก็อาจจะได้รับการคัดเลือกให้มาบรรจุไว้ในบัตรได้ ข้อความที่

คัดเลือกมาอาจมาจากแหล่งต่าง ๆ กันได้หลายทาง เช่น บัญชีรายการบุคลิกภาพ (personality inventories) คำพูดของคนไข้เมื่อเวลาไปขอคำปรึกษาจากจิตแพทย์ หนังสือหรือเอกสารต่าง ๆ ทั่วไปที่มีข้อความที่เกี่ยวข้อง นอกจากนี้ผู้วัดก็อาจคิดข้อความอะไรขึ้นมาเองก็ได้ เนื่องจากข้อความหรือรูปภาพเหล่านี้มีลักษณะความคล้ายคลึงภายในกลุ่ม จึงได้ถือว่า จะสามารถวัดตัวแปรตัวใดตัวหนึ่งอย่างกว้าง ๆ เช่น ลักษณะประสาทหลอน (neuroticism) การปรับตัว (adjustment) และสัญชาตญาณ (perception of self) เป็นต้น ยกตัวอย่างว่าเมื่อผู้วัดต้องการวัดสัญชาตญาณ ก็อาจประมวลคำกล่าวต่าง ๆ ที่เกี่ยวข้องกับการพิจารณาตนเองมาไว้หลาย ๆ ข้อ เช่น ข้าพเจ้าชอบมนุษย์ทั้งหลาย ข้าพเจ้าไม่ก้าวหน้าในชีวิต ข้าพเจ้าตัดสินใจในเรื่องต่าง ๆ ไม่ค่อยได้ เมื่อผู้ถูกวัดได้เลือกบัตรที่บรรจุข้อความเหล่านี้ และผู้วัดได้วิเคราะห์ตารางสัมพันธ์เรียบร้อยแล้ว ก็อาจสามารถลงความเห็นในเรื่องทัศนคติของผู้ถูกวัดได้ เช่นในกรณีของบุคคลประสาทหลอน (neurotic person) หากมีสัมพันธ์สหสัมพันธ์ระหว่างการเลือกบัตรในระยะก่อนการเข้ารักษาที่ระยะหลังการเข้ารักษา ก็หมายความว่า การเข้ารักษาบุคคลผู้นั้นได้ผลเพราะแสดงว่ามีการเปลี่ยนแปลงทัศนคติ บุคคลที่ไม่ได้รับประโยชน์จากการเข้ารักษา และมีได้รับการปฏิบัติการอันใดจากจิตแพทย์ ย่อมได้ค่าสหสัมพันธ์ระหว่างการเลือกบัตร 2 ครั้ง ในระดับสูง ถ้าหากวิธีการเลือกบัตรทั้ง 2 ครั้งเหมือนกัน และวิธีการที่ใช้ในเทคนิคควินั้น มีความเชื่อมั่นได้

การสร้างข้อความหรือรูปภาพด้วยวิธีกำหนดโครงสร้าง เป็นการกระทำที่ผู้วัดกำหนดประเภทของสิ่งเร้า (stimuli) ไว้แน่นอนตายตัว กล่าวอีกอย่างหนึ่งก็คือ ผู้วัดกำหนดตัวแปรไว้อย่างเฉพาะเจาะจง โดยการสร้างสมมติฐานหรือทฤษฎีไว้อย่างชัดเจน หรืออาศัยทฤษฎีทางจิตมิติ (psychometry) ที่มีการค้นคว้ามาก่อนแล้วเป็นหลัก การสร้างข้อความด้วยวิธีนี้ตรงกันข้ามกับการสร้างด้วยวิธีสุ่มโดยผู้วิจัยกำหนดเค้าโครงของลักษณะข้อความหรือรูปภาพทั้งหมดที่จะบรรจุในบัตร ไม่ใช่เพียงแต่เลือกประมวลมาจากแหล่งต่าง ๆ โดยให้มีความคล้ายคลึงภายในกลุ่มเท่านั้น เมื่อกำหนดสมมติฐานหรือทฤษฎีไว้อย่างตรงประเด็นกับลักษณะของทัศนคติที่จะวัดแล้ว ก็จะต้องคิดข้อความที่สัมพันธ์กับรายละเอียดของสมมติฐานหรือทฤษฎีนั้น ๆ เช่น สมมติว่า จะต้องการวัดกำลังใจร่วมในการปฏิบัติงาน (morale)¹¹ และผู้วัดตั้งสมมติฐานว่า สิ่งที่จะท่อนกำลังใจร่วมในการปฏิบัติงานมี 3 ประการ คือ โอกาสแสดงออกของบุคคล (individual self-expression)

¹¹ คำนี้ มีบางแห่งใช้ว่า ขวัญ ซึ่งมีความหมายไม่ตรงกับภาษาอังกฤษทีเดียว เพราะ คำว่า ขวัญ อาจใช้กับความที่ว่า ขวัญหนีดีฝ่อก็ได้.

ความภูมิใจในฝีมือการทำงานแต่ละบุคคล (individuals' pride in his own workmanship) และการยอมรับวัตถุประสงคฺ์และค่านิยมของกลุ่ม ข้อความที่จะบรรจุในบัตรก็จะต้องเป็นข้อความที่สอดคล้องกับรายละเอียดในหัวข้อทั้งสามนี้ ก่อนที่จะบรรจุลงบัตรข้อความเหล่านี้ควรได้ผ่านการพิจารณาจากผู้อำนวยการบริหารหลาย ๆ ฝ่าย เพื่อให้แน่ใจว่าเป็นข้อความที่สอดคล้องกับรายการต่าง ๆ ในเรื่องกำลังใจร่วมในการปฏิบัติงาน อัตราส่วนข้อความที่สะท้อนรายละเอียดในหัวข้อแต่ละหัวข้อก็ต้องกำหนดให้เป็นสัดส่วน เพื่อให้สมดุลย์กัน สมมติฐานที่กำหนดนี้เป็นเครื่องชี้ว่าจะต้องมีข้อความอย่างไรบ้าง ในกรณีนี้ก็อาจมีข้อความที่สะท้อนความรู้สึกที่แต่ละบุคคลพอใจในการที่มีโอกาสแสดงออก ประมาณ 10-15 รายการ ความรำคาญในแนวปฏิบัติและวิธีของผู้บริหารที่มีต่อผู้ปฏิบัติงานแต่ละคน 10-15 รายการ ความรู้สึกที่ว่าตนเป็นเพียงบุคคลคนหนึ่ง ในหน่วยงานเท่านั้น 10-15 รายการ ความเห็นว่าวัตถุประสงคฺ์และค่านิยมถูกต้องแล้ว หรืออาจมีการเปลี่ยนแปลงปรับปรุงบ้างบางส่วน 10-15 รายการ และความรู้สึกว่าวัตถุประสงคฺ์และค่านิยมของหน่วยงานขัดกับหลักการบริหาร หรือสภาพของสังคมและการครองชีวิต อีก 10-15 รายการ บุคคลที่มีกำลังใจร่วมในการปฏิบัติงานสูง ก็จะเลือกแต่บัตรที่มีข้อความในลักษณะสะท้อนความรู้สึกที่ว่าตนมีความสามารถ มีผลงาน และเห็นว่า พิชัยการ และขบวนการในการบริหารเท่าที่เป็นอยู่ ทำให้ตนมีกำลังใจในการปฏิบัติงาน ส่วนผู้ที่ขาดกำลังใจร่วมในการปฏิบัติงาน ก็จะเลือกบัตรที่มีข้อความอีกประเภทหนึ่งที่สะท้อนความรู้สึกว่า ในหน่วยงานนั้น ๆ มีแต่การกดขี่ไม่ให้ความเป็นธรรมแก่ผู้ปฏิบัติงาน และผู้ปฏิบัติงานแต่ละคนต่างปฏิบัติงานไปวันหนึ่ง ๆ เท่านั้นเอง โดยมีได้มีความรักหน่วยงานของตนเลย

การสร้างข้อความหรือรูปภาพด้วยวิธีกำหนดโครงสร้างนี้ เป็นการกำหนดขอบเขตว่าจะให้เนื้อหาที่วัดครอบคลุมไปเพียงใด อย่างไรก็ดีในการวัดทัศนคติหลาย ๆ ลักษณะไม่สามารถกำหนดขอบเขตให้แน่ชัดลงไปได้ ในกรณีที่ต้องการทราบว่า ทัศนคติของบุคคลที่เปลี่ยนแปลงไปในระยะหนึ่ง ผู้วัดสามารถกำหนดขอบเขตให้เฉพาะเจาะจงลงไปได้ยาก ในการวัดส่วนใหญ่จึงหันไปใช้วิธีคัดเลือกข้อความหรือรูปภาพด้วยการสุ่ม เพื่อให้สามารถครอบคลุมทัศนคติในเรื่องนั้น ๆ ได้อย่างกว้าง ๆ

เทคนิคการวัดความสัมฤทธิ์ผล

การวัดความสัมฤทธิ์ผล (achievement) คือ การวัดว่า ผู้ที่จะปฏิบัติงานมีความรู้ ความชำนาญ ในวิชาการสาขาหนึ่งๆ หรือเรื่องหนึ่งๆ หรือมีทักษะและความเข้าใจ ในเนื้อหาของงานในตำแหน่งหน้าที่ที่ต้องรับผิดชอบประการใดบ้าง สิ่งที่จะวัดในการวัด ความสัมฤทธิ์ผลส่วนใหญ่จึงเป็นสิ่งที่ได้ศึกษาอบรมและเรียนรู้มา เครื่องมือวัดความ สัมฤทธิ์ผลอาจแยกเป็นประเภทใหญ่ๆ ได้ 2 ประเภท คือ ประเภทแรก ได้แก่ ข้อสอบทั่วไปที่ออกเพื่อวัดผลการศึกษาอบรม ตามหลักสูตรของผู้ปฏิบัติงานหรือของ สถาบันการศึกษาต่างๆ ข้อสอบเหล่านี้ผู้ให้การศึกษาอบรมหรือครูอาจารย์ผู้สอนเป็นผู้ กำหนดเอง ให้มีความยากง่ายในระดับต่างๆ กัน ตามประเภทและชั้นของการศึกษา อบรม ข้อสอบในประเภทแรกมีทั้งในรูปอัตนัยและปรนัย ประเภทที่สอง ได้แก่ แบบ ทดสอบมาตรฐาน ซึ่งมีใช้แพร่หลายในสหรัฐอเมริกา แบบทดสอบประเภทหลังนี้เป็น แบบที่หน่วยงานต่างๆ สร้างขึ้นเป็นพิเศษ โดยอาศัยเนื้อหาของความรู้ในวิชาหนึ่งๆ หรือขอบเขตของงานที่จำเป็นจะต้องทราบและเข้าใจในการปฏิบัติหน้าที่ในตำแหน่ง หนึ่งๆ เป็นหลัก ข้อสอบประเภทนี้มีหลักการคำนวณค่าความเชื่อมั่น ความเที่ยง ตรง และเกณฑ์การวัด (norm) ไว้เรียบร้อยแล้ว แบบทดสอบมาตรฐาน ยังอาจแยกได้ ออกเป็น 2 แบบๆ หนึ่งวัดความรู้ทั่วไป เช่น ความรู้ในเรื่องภาษา คณิตศาสตร์ หรือ สังคมศาสตร์ อีกแบบหนึ่ง วัดความรู้และความชำนาญเฉพาะอย่าง เช่น ความรู้ใน หน้าที่นักบัญชี ความรู้ในตำแหน่งงานช่างไฟฟ้า หรือความสามารถในด้านภาษาอังกฤษ เป็นต้น

การวางแผนในการสร้างแบบทดสอบ

แบบทดสอบความสัมฤทธิ์ผลเป็นเครื่องมือวัดที่ต้องการความเที่ยงตรงในแง่เนื้อหา (content validity) แบบทดสอบนี้จะต้องครอบคลุมประเด็นและสาระสำคัญของวิชา หรือรายละเอียดของงานในตำแหน่งหน้าที่ที่จะต้องปฏิบัติ การที่จะให้แบบทดสอบ มีค่าความตรงประเด็นในแง่เนื้อหาในระดับสูงได้ ควรจะต้องมีการวางแผนอย่างรอบ คอบ มิใช่รอจนกว่าได้สร้างแบบทดสอบเสร็จเรียบร้อยแล้วจึงมีการวิเคราะห์สถิติ ในขั้นต้นจะต้องมีบุคคลซึ่งมีความชำนาญทางวิชาการในวิชาหรืองานที่จะทดสอบ เพื่อ พิจารณาตกลงกันว่าวิชาหรืองานนั้น ๆ มีประเด็นหรือสาระสำคัญประการใดบ้าง แผนการ สร้างแบบทดสอบในระยะแรกประกอบด้วยเค้าโครงของเนื้อหาที่จะทดสอบ เค้าโครง ของเนื้อหาจะเป็นหลักในการพิจารณาให้แบบทดสอบมีความตรงประเด็นในแง่เนื้อหา โดยที่สามารถทำให้เห็นชัดเจนลงไปได้ว่า จะตั้งคำถามในเรื่องใด ประเด็นใด ใน ลักษณะใด และให้มีปริมาณมากน้อยเพียงใด ยกตัวอย่างเช่น สมมุติว่า ในการคัดเลือก พนักงานตรวจสำนวนตรี ได้กำหนดเค้าโครงของเนื้อหาที่จะต้องทราบ เข้าใจและ มี ความสามารถปฏิบัติงานนี้ไว้ว่า จะต้องมัทักษะในด้านกรอ่านภาษาเขียน เป็นอันดับ แรก และการอ่านภาษาเขียนนี้ จะต้องครอบคลุมความสามารถในประเด็นต่าง ๆ ดังต่อ ไปนี้

ก. ทักษะและความสามารถในการเข้าใจความหมาย ได้แก่ความสามารถ 4 ประการ คือ เข้าใจปัญหาและวัตถุประสงค์ของเรื่อง สามารถพ่วงความสนใจให้แก่ความหมาย ของข้อความ มีสัญชาตญาณ (perception) ต่อรูปแบบของข้อเขียนในลักษณะต่าง ๆ และ เข้าใจความคิดของผู้เขียนและจัดเรียงลำดับความสำคัญของประเด็นต่าง ๆ ในข้อความ หนึ่ง ๆ ได้

ข. คุณพินิจในการวิพากษ์วิจารณ์ ได้แก่ ความสามารถ 5 ประการ คือ วิเคราะห์ ได้ว่า ข้อเขียนที่อ่านมีความสำคัญอย่างไร ตัดสินได้ว่าความคิดในข้อความนั้นมีความ เป็นจริงและตรงประเด็นเพียงใด ตัดสินได้ว่า ข้อความนั้นสมบูรณ์หรือไม่ ลงความ เห็นได้ว่าเหตุผลของผู้เขียนครบถ้วนและรับฟังได้หรือไม่ และสามารถจับประเด็นที่มี ลักษณะเป็นการโฆษณาชวนเชื่อได้

เค้าโครงของความรู้ความสามารถที่จะต้องมามากไปกว่านี้ หรือลดน้อยกว่าขึ้นอยู่กับคุณพินิจทางวิชาการในวิชาหรืองานนั้น ๆ เมื่อได้กำหนดเค้าโครงเช่นที่ยกมาเป็นตัว อย่างนี้แล้ว ในการวางแผนก็จะต้องพิจารณาต่อไปว่าจะตั้งคำถามในลักษณะใด ในบท

ที่ 1 ได้กล่าวถึงแบบต่าง ๆ ของคำถามและลักษณะของการตั้งคำถามที่ดีไว้แล้ว จึงจะไม่กล่าวซ้ำในที่นี้ อีก ในการกำหนดลักษณะคำถามที่จะตั้งนั้นจะต้องวิเคราะห์ด้วยว่าจะตั้งคำถามเป็นปริมาณเท่าใด โดยแยกเป็นภาคเป็นตอน ๆ ละเท่าใด เป็นอัตราเท่าใดซึ่งกันและกัน ตลอดจนวิธีการที่จะคุมการใช้แบบทดสอบการให้ค่าคะแนนแก่ผู้ตอบแบบทดสอบ และการเทียบเกณฑ์การวัดสำหรับแบบทดสอบนี้โดยเฉพาะ

เมื่อได้กำหนดแผนการสร้างแบบทดสอบตามแนวข้างต้นนี้เรียบร้อยแล้ว จะต้องให้คณะกรรมการพิจารณาแผนการสร้างขึ้นให้รอบคอบเสียก่อน คณะกรรมการนั้นควรประกอบด้วย นักวิชาการในหัวข้อวิชาหรือเรื่องของงานที่จะวัด เจ้าหน้าที่ฝ่ายบริหารของหน่วยงานที่มีส่วนได้ส่วนเสียในการวัด และผู้ชำนาญการในด้านจิตวิทยาและการประเมินผลบุคคล เมื่อได้ประเมินว่า แผนที่สร้างขึ้นมีความรัดกุมในประเด็นและหลักการต่าง ๆ แล้ว จึงจะถือเป็นแผนที่ใช้เป็นหลักในการสร้างแบบทดสอบต่อไปได้ โดยปกติการสร้างเครื่องมือวัดความสัมฤทธิ์ผลประเภทแรก ซึ่งได้แก่ข้อสอบทั่ว ๆ ไปก็มีได้มีการวางแผนตามขั้นตอนดังกล่าว เพราะเป็นการสิ้นเปลืองเวลาของบุคคลหลายฝ่ายเป็นการล่าช้ามากเกินไป ในการสร้างแบบทดสอบมาตรฐานเท่านั้น จึงมีการดำเนินการตามหลักเกณฑ์ทั้งหมดนี้อย่างถี่ถ้วน แต่ถึงอย่างไรก็ดี ในการสร้างข้อสอบทั่ว ๆ ไปเพื่อวัดผลการฝึกอบรมผู้ปฏิบัติงานตามหลักสูตรหนึ่งหลักสูตรใด ก็ควรจะได้อนุโลมนำขั้นตอนในการวางแผนการสร้างแบบทดสอบเท่าที่กล่าวมานี้ไปใช้อย่างกว้าง ๆ เพื่อให้ได้ข้อสอบที่มีความตรงประเด็นกับการฝึกอบรมให้มากที่สุดเท่าที่จะมากได้

การทดลองแบบทดสอบ การทดลอง (pretest) แบบทดสอบที่สร้างขึ้นใหม่ มีความจำเป็นอย่างยิ่ง การทดลองอาจทำได้เป็น 3 ตอน¹ ตอนแรกเป็นการนำแบบทดสอบไปทดลองกับกลุ่มตัวอย่างของประชากรที่มีลักษณะคล้ายคลึงกับผู้ที่จะเป็นผู้ตอบแบบทดสอบในโอกาสต่อไป กลุ่มตัวอย่างนี้ไม่มีความจำเป็นต้องเกินกว่าจำนวน 100 ราย ความมุ่งหมายของการทดลองในขั้นนี้เพื่อทราบข้อบกพร่องทั่ว ๆ ไป การทดลองนี้ ผู้สร้างแบบทดสอบควรเป็นผู้ปฏิบัติเอง เพื่อให้สามารถเปรียบเทียบข้อบกพร่องทั้งหมดได้อย่างครบถ้วน การทดลองในตอนที่สองเป็นการวิเคราะห์รายการคำถามในแบบทดสอบ (item analysis) จำนวนของกลุ่มตัวอย่างที่ใช้ในการทดลองควรมีประมาณ 400 ราย รายละเอียดในการวิเคราะห์รายการคำถามจะมีกล่าวในหัวข้อต่อไปโดยเฉพาะ การทดลองขั้นสุดท้ายเป็นการตกแต่งแบบทดสอบทั้งหมดให้สมบูรณ์ที่สุด โดยพยายามสำรวจข้อบกพร่องเล็ก ๆ น้อย ๆ ที่อาจผ่านไปโดยมิได้สังเกตมาก่อน และในการทดลองขั้นนี้จะต้อง

¹ Guilford, J.P., *Psychometric Method*, p. 417.

วิเคราะห์ค่าความเชื่อมั่นให้เรียบร้อยด้วย เพื่อให้ตัดสินใจได้ว่า แบบทดสอบที่สร้างขึ้น มีคุณภาพใช้การได้เพียงไร การทดลองทั้งหมดนี้จะทำให้สามารถปรับปรุงแบบทดสอบให้มีลักษณะที่ถูกต้อง มีคำถามที่เลือกสรรแล้วเป็นอย่างดี และสามารถกำหนดความยาวของแบบทดสอบให้เหมาะสมและสอดคล้องกับเวลาที่กำหนดเพื่อการตอบแบบทดสอบ

แบบทดสอบที่จำกัดเวลา และที่ไม่จำกัดเวลา ข้อสอบจำกัดเวลาสอบ (speed test) คือ ข้อสอบที่บุคคลส่วนใหญ่สามารถตอบคำถามข้อต่าง ๆ ได้ถูกต้อง และค่าคะแนนที่ผู้ตอบแบบทดสอบได้รับขึ้นอยู่กับจำนวนของคำถามที่ตอบได้ถูกต้องตามเวลาที่กำหนดให้² แต่โดยทั่วไปอาจกล่าวได้ว่า แบบทดสอบที่จำกัดเวลาเป็นแบบทดสอบที่ไม่มีผู้ใดสามารถตอบคำถามได้ครบทุกข้ออย่างถูกต้องในเวลาที่กำหนด เพราะผู้สร้างแบบทดสอบได้กำหนดเวลาไว้อย่างจำกัดมาก ส่วนแบบทดสอบที่ไม่จำกัดเวลา (power test) คือ แบบทดสอบที่ผู้ตอบทุกคนมีเวลาตอบคำถามทุกข้ออย่างเพียงพอ³

ในการสร้างแบบทดสอบ มีปัญหาใหญ่ที่ศึกษากันมาเป็นเวลานานแล้วคือ ควรเลือกสร้างเป็นประเภทจำกัดเวลา หรือไม่จำกัดเวลา และประเภทใดมีส่วนดีส่วนเสียกว่ากันอย่างไร ในประเด็นเรื่องนี้อาจแยกเป็นปัญหาใหญ่ ๆ ได้ 2 ประการ คือ การตอบคำถามในเวลาจำกัดนั้นแสดงว่าสามารถตอบคำถามต่าง ๆ ในประเภทเดียวกันได้ดีจริง ๆ หรือไม่ กล่าวคือ ผู้ที่มีความสามารถตอบคำถามตามที่กำหนดในเวลาที่กำหนดให้ จะสามารถตอบคำถามทั่วไปที่ยากกว่านั้น ถึงแม้ว่าไม่มีกำหนดเวลาเลยหรือไม่ และประการที่สอง การจำกัดเวลาให้มีความนานในระยะต่าง ๆ กัน มีผลกระทบต่อคุณภาพของการวัดทางจิตวิทยาหรือไม่ ยกตัวอย่างเช่น แบบทดสอบที่มีสัมประสิทธิ์ความเชื่อมั่นสูง เมื่อจำกัดเวลาให้น้อยลง จะทำให้ค่าความเชื่อมั่นลดลงไปด้วยหรือไม่ สำหรับปัญหาประการแรก Spearman เคยยืนยันว่า ความเร็วของการตอบถูกต้องไม่ทำให้คุณภาพของการตอบปัญหาที่ยากกว่าถูกต้องเสียไป แต่อย่างไรก็ดี ในปัจจุบันนี้ได้มีการพิสูจน์แล้วว่าความเห็นของ Spearman ในเรื่องนี้ไม่ถูกต้อง⁴ เพราะโดยแท้ที่จริงรายการคำถามที่กำหนดให้ตอบอย่างจำกัดเวลานั้นมิใช่เป็นสิ่งที่เกี่ยวข้องกับรายการปัญหาซึ่งผู้ตอบอาจประสบในสภาพการณ์ทั่วไป ฉะนั้น ความสามารถในการตอบปัญหาตามปกติในแบบทดสอบจึงมิใช่เป็นความสามารถอย่างเดียวกับความสามารถในการแก้ปัญหายากขึ้นในระดับสูงขึ้นไป

² Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, p. 282.

³ Guilford, *op. cit.*, p. 368.

⁴ Nunnally, Jum C., *Psychometric Theory*, pp. 560—561.

ในเรื่องผลเสียของการจำกัดเวลาที่ต่อสัมพันธ์กับความเชื่อมั่น ได้มีการพิสูจน์ไว้ในที่หลายแห่ง⁵ ในกรณีที่มีการจำกัดเวลาตอบแบบทดสอบสัมพันธ์กับความเชื่อมั่นที่คำนวณได้อาจมีค่าสูงเกินกว่าความเป็นจริง ในบางครั้งอาจคำนวณได้สูงถึงเกือบ 1.00 ซึ่งแท้ที่จริงแล้ว แบบทดสอบฉบับนั้นมิได้มีค่าความเชื่อมั่นสูงถึงเพียงนั้น โดยเฉพาะถ้าหากคำนวณสัมพันธ์ความเชื่อมั่นด้วยวิธีแบ่งครึ่งเครื่องมือวัดแบบจัดรายการหมายเลขคี่และคู่ (odd-even coefficient) อย่างไรก็ดี โดยปกติแบบทดสอบทั้งหลายจะมีลักษณะระหว่างกลางของประเภทจำกัดเวลาและประเภทไม่จำกัดเวลา ต่างกันแต่ว่าบางฉบับจะค่อนข้างไปทางประเภทใดประเภทหนึ่งมากกว่ากัน แบบทดสอบในการวัดบุคคลส่วนใหญ่ค่อนข้างไปทางประเภทจำกัดเวลา และในการจำกัดเวลานี้ก็กระทำในลักษณะที่ให้ผู้บุคคลส่วนใหญ่สามารถตอบคำถามต่าง ๆ ได้ เมื่อได้ปฏิบัติในลักษณะเช่นนี้ก็หมายความว่า ถึงแม้จะเพิ่มเวลาขึ้นอีกก็ไม่ทำให้ค่าคะแนนที่ควรจะได้เปลี่ยนแปลงไปมากนัก ฉะนั้นประเด็นเรื่องการจำกัดเวลาจึงไม่เป็นปัญหาต่อไป

การวิเคราะห์รายการคำถาม

แบบทดสอบจะดีได้ขึ้นอยู่กับรายการคำถามทั้งหมดที่สร้างขึ้น การวางแผนที่ดีทำให้มีหลักเกณฑ์ในการสร้างแบบทดสอบที่ดี แต่ถ้าหากไม่สามารถตั้งคำถามที่ดีได้แผนการที่ร่างไว้ก็ไม่มีประโยชน์ ฉะนั้นเมื่อได้ร่างคำถามตามลักษณะที่กำหนดไว้แล้ว ก็จะต้องมีการวิเคราะห์เพื่อให้ทราบว่า คำถามเหล่านั้นมีคุณภาพถึงขนาดหรือไม่ ในการวิเคราะห์นี้ต้องการได้ข้อเท็จจริง 2 ประการ คือ ดัชนีที่จะชี้ความยากง่ายของรายการคำถามและสัมพันธ์ความเที่ยงตรง ทั้งนี้เพราะคำถามที่ตั้งขึ้น ถึงแม้ว่าจะกระทำโดยผู้เชี่ยวชาญในเรื่องนี้ ก็ยังเป็นเพียงคำถามที่ผู้สร้างแบบทดสอบคิดขึ้นด้วยตนเอง ไม่ใช่เป็นการเลือกสรรคำถามมาโดยวิธีปรนัย การวิเคราะห์จะทำให้ทราบได้ว่า คำถามใดดีที่สุดควรคงไว้ หรือบางครั้งอาจต้องแก้ไขแนวของการตั้งคำถามเสียใหม่ การเลือกคำถามให้มีเฉพาะรายการที่มีความยากตามสมควร และให้มีการกระจายคำถามที่ยากในระดับเดียวกันออกไป ตลอดจนการจัดให้คำถามแต่ละข้อมีค่าสหสัมพันธ์ระหว่างกันในระดับอันสมควร จะทำให้แบบทดสอบที่สร้างขึ้นเพียบพร้อมไปด้วยคุณสมบัติในเชิงสถิติตามที่ต้องการ กล่าวคือ มีค่าเฉลี่ยของการให้ค่าคะแนน ความแปรปรวนของผลการวัด ลักษณะการแจกแจงของค่าคะแนน ความเชื่อมั่น และความเที่ยงตรงที่ไม่บิดเบือนไปจากระดับปกติ

⁵ Thorndike, Robert L., *Personnel Selection*, p. 112. Guilford, *op. cit.*, pp. 391-392.

การดำเนินการวิเคราะห์ จำนวนผู้ตัวอย่างประชากรที่มีลักษณะคล้ายคลึงกับผู้ที่จะเป็นผู้ตอบแบบทดสอบในโอกาสต่อไป เพื่อทดลองในการวิเคราะห์รายการคำถามได้กำหนดไว้ว่า ควรมีประมาณ 400 ราย โดยหลักเกณฑ์จริงๆ นั้น อาจกำหนดเป็นหลักกว้างๆ ได้ว่า ต้องมีจำนวนไม่น้อยกว่า 5 เท่าของจำนวนคำถามที่มีในแบบทดสอบ⁶ ส่วนรายการคำถามในแบบทดสอบที่จะทดลองนี้ ก็ควรมีมากกว่าจำนวนคำถามที่กำหนดว่าจะให้มีในแบบทดสอบที่จะใช้จริงๆ ปริมาณที่มากกว่านี้อาจไม่จำเป็นที่จะต้องให้มากกว่า 50% ของจำนวนคำถามที่จะใช้จริงๆ คำถามในแบบทดสอบควรมีทั้งยากและง่ายคละกัน เพื่อให้สามารถถ่วงกรองคำถามที่ยากในระดับเดียวกันไว้ได้ในขั้นสุดท้าย ในการทดลองควรกำหนดเวลาให้มากพอที่ผู้ตัวอย่างของผู้ถูกทดลองสามารถตอบได้ทุกข้อในการให้ผู้ทดลองแต่ละคนตอบคำถามอาจมีการลำดับลำดับของคำถามให้ต่าง ๆ กัน เพื่อให้คำถามแต่ละข้อมีโอกาสที่จะอยู่ในส่วนสุดท้ายของแบบทดสอบเหมือนกัน ทั้งนี้ อาจจะมีการเพิ่มคำถามบางข้อที่ไม่จำเป็นต้องวิเคราะห์ไว้ตอนท้าย เพื่อให้ผู้ถูกทดลองที่ทำแบบทดสอบได้เร็วกว่าผู้อื่นได้มี “งานทำ” เท่า ๆ กับผู้ถูกทดลองที่ตอบช้า

ตัวเลขที่จำเป็นต้องทราบก่อนอื่น คือ สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่สร้างขึ้น ถ้าสัมประสิทธิ์ความเชื่อมั่นต่ำ แสดงว่า รายการคำถามที่สร้างขึ้นมีความผิดแผกกันภายในกลุ่ม (heterogeneity) สูง ในกรณีเช่นนั้น คำถามแต่ละข้อย่อมมีสัมประสิทธิ์สหสัมพันธ์กับรายการคำถามทั้งหมดในระดับต่ำ เมื่อเป็นเช่นนี้จำเป็นต้องนำหลักการวิเคราะห์หองศ์ประกอบที่กล่าวไว้ในบทที่ 7 มาใช้ เพื่อคำนวณว่าจะมีคำถามข้อใดที่จัดเป็นกลุ่มเดียวกันได้หรือไม่ ถ้ามี ก็อาจวิเคราะห์รายการคำถามในแต่ละกลุ่มนั้นต่อไป การคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างคำถามแต่ละข้อกับคำถามทั้งหมดก็คือการคำนวณค่าคะแนนที่ผู้ถูกทดลองได้รับในการตอบคำถามแต่ละข้อ กับค่าคะแนนรวมของทุกข้อนั่นเอง ในกรณีที่แบบทดสอบฉบับหนึ่ง ๆ แบ่งเป็นหลายภาค การคำนวณสัมประสิทธิ์สหสัมพันธ์ก็ควรกระทำระหว่างค่าคะแนนสำหรับคำถามข้อหนึ่ง ๆ กับค่าคะแนนรวมของคำถามทั้งหมดเฉพาะภาคหนึ่ง ๆ คำถามที่ได้ค่าสหสัมพันธ์เป็นศูนย์หรือใกล้ศูนย์ ควรนำมาพิจารณาอย่างละเอียดว่ามีความหมายบกพร่องประการใด คำถามที่มีลักษณะเช่นนี้มีข้อบกพร่องได้หลายประการ เช่น ยากหรือง่ายเกินไป มีความหมายกำกวม หรืออาจไม่ตรงประเด็นกับเรื่องที่จะวัด อย่างไรก็ตาม ใดๆ ก็ดี เฉพาะในแบบทดสอบความสัมฤทธิ์ผล อาจมีคำถามบางข้อที่มีค่าสหสัมพันธ์กับค่าคะแนนรวมของคำถามทั้งหมดต่ำ แต่ก็ยังมีความเที่ยงตรงได้ ถ้าหากไม่มีหลักฐานแน่ชัดว่า คำถาม

⁶ Nunnally, *op. cit.*, p. 242.

ประเภทนี้ยังมีความเที่ยงตรงอยู่ ก็ควรจะตัดทิ้งไปเสีย ผู้สร้างแบบทดสอบจะเลือกเฉพาะคำถามข้อที่มีค่าสหสัมพันธ์สูงไว้เท่านั้น

สัมประสิทธิ์สหสัมพันธ์ระหว่างคำถามแต่ละข้อกับรายการคำถามทั้งหมดอาจมีพิสัยระหว่างศูนย์ถึงประมาณ .40 ตามปกติเครื่องมือวัดความสามารถนี้ไม่มีเครื่องมือสัมประสิทธิ์สหสัมพันธ์ในเชิงนิเสธ ในกรณีที่เกิดขึ้นก็มักจะเป็นเพราะข้อบกพร่องจากการสุ่มตัวอย่าง โดยทั่วไปจึงถือว่าสัมประสิทธิ์สหสัมพันธ์ที่สูงกว่า .20 ขึ้นไป เป็นระดับที่น่าพอใจแล้ว⁷ ถ้ามีคำถามที่มีสัมประสิทธิ์สหสัมพันธ์ในระดับที่ต้องการเป็นจำนวนเพียงพอ ก็นำกลุ่มคำถามเหล่านี้ไปคำนวณค่าความเชื่อมั่นเป็นกลุ่ม ๆ ไป เพื่อจะได้มีโอกาสพิจารณาต่อไปว่าจะปรับปรุงคำถามต่อไปอีกประการใด ในกรณีที่ไม่มีจำนวนพอดีโดยประมาณกับจำนวนที่กำหนดให้มีในแบบทดสอบที่จะใช้จริง ๆ ก็จำเป็นต้องรับเอาคำถามเหล่านั้นไว้ทั้งหมดโดยไม่มีโอกาสปรับปรุง แต่ถ้าจำนวนที่น่าพอใจในขั้นทดลองมีน้อยกว่าจำนวนที่ต้องการมาก ก็อาจจำเป็นต้องเริ่มต้นประมวลคำถามทั้งหมดใหม่อีกครั้ง กลุ่มคำถามที่มีทั้งสัมประสิทธิ์สหสัมพันธ์และสัมประสิทธิ์ความเชื่อมั่นสูงพอ ก็จะนำมาพิจารณาลักษณะการแจกแจงของค่าคะแนนแต่ละข้อ แบบทดสอบมาตรฐานที่จำเป็นต้องมีการแจกแจงค่าคะแนนของแต่ละข้อในลักษณะปกติ หรือใกล้เคียงกับปกติ (approximately normal distribution) การแจกแจงที่เบ้ (skewed) ไปในทางหนึ่งทางใด หมายถึงว่าคำถามในกลุ่มที่คัดเลือกไว้นั้นยากหรือง่ายเกินไป ถ้าเป็นเช่นนั้นก็จะต้องนำคำถามที่ง่ายขึ้นหรือยากขึ้นไปใส่แทนจำนวนที่ควรตัดออก เพื่อให้การแจกแจงของค่าคะแนนมีลักษณะใกล้เคียงกับการแจกแจงปกติ

ในกรณีที่สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่สร้างขึ้นใหม่ทั้งหมดอยู่ในระดับสูงแล้ว เช่น .90 หรือสูงกว่านี้ ก็ไม่มีความจำเป็นที่จะต้องดำเนินการวิเคราะห์คำถามแต่ละข้อตามแนวในย่อหน้าก่อน การคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างคำถามแต่ละข้อกับรายการคำถามทั้งหมด หรือรายการคำถามแต่ละกลุ่ม ไม่ทำให้สามารถปรับปรุงอะไรได้เพิ่มขึ้น ในกรณีเช่นนั้น ก็อาจปรับปรุงรายการคำถามให้ดีขึ้นได้โดยใช้ดัชนีความยาก (difficulty indices) ต่อไป

ดัชนีความยาก อัตราส่วนของผู้ที่ตอบคำถามได้เป็นเครื่องชี้ว่า แบบทดสอบที่สร้างขึ้นนั้นมีความยากง่ายเพียงใด ถ้าหากว่ามี การให้คะแนนสำหรับผู้ตอบได้เป็นหนึ่งและผู้ที่ตอบไม่ได้เป็นศูนย์ อัตราเฉลี่ยของผู้ที่ได้คะแนนเท่ากับหนึ่ง เมื่อคำนวณให้เป็นค่าความน่าจะเป็น (p value) ของผู้ตอบได้ ก็จะได้ค่าดัชนีความยาก สมมุติว่าในการ

⁷ *Ibid.*, pp. 242–243.

ตอบแบบทดสอบฉบับหนึ่ง มีผู้ตอบได้เป็นอัตราร้อยละ 80 เมื่อกำหนดค่า p value ซึ่ง
เป็นดัชนีความยากของแบบทดสอบฉบับนี้ ก็จะได้ค่า .8

ดัชนีความยากที่ว่าประชากรในลักษณะหนึ่ง ๆ มีความน่าจะเป็นมากน้อยเพียงใด
ในการตอบคำถามในแบบทดสอบฉบับหนึ่ง ๆ ความน่าจะเป็นในเรื่องนี้จึงเป็นอัตราเฉลี่ย
ของความยากของคำถามสำหรับบุคคลต่าง ๆ ดัชนีนี้ไม่ได้ชี้ว่าคำถามข้อหนึ่ง ๆ นั้น
มีความยากสำหรับบุคคลแต่ละคนอย่างไร ในกรณีของผู้ที่ตอบได้ มีความหมายแต่
เพียงว่าคำถามนั้นยากน้อยกว่าความสามารถที่จะทำได้ แต่ถ้าตอบไม่ได้ ก็แปลว่ายาก
เกินกว่าความสามารถที่จะทำได้ ในเรื่องนี้มีข้อควรระวังในเชิงสถิติ 2 ประการ คือ
ประการแรก แท้ที่จริงดัชนีนี้มีความสัมพันธ์เชิงกลับกันกับความยาก ฉะนั้น ดัชนี
ความยากนี้ก็คือ เครื่องวัดความง่ายโดยตรงนั่นเอง และประการที่สอง ดัชนีนี้มิได้มี
ความสัมพันธ์ในเชิงเส้นตรง (linearly related) กับความสามารถ

จากข้อบกพร่องของการใช้ดัชนีความยากที่คำนวณจากค่าความน่าจะเป็นดังกล่าว จึง
ได้มีนักวิชาการเสนอทฤษฎีการวัดความยากหลายแบบ ที่ควรนำมากล่าวไว้ในที่นี้ คือ
แบบของ Guilford⁸ ตามทฤษฎีนี้ความน่าจะเป็นของการตอบได้นั้น ยังอาจเบี่ยงเบน
ไปเนื่องจากการเผชิญตอบถูก ฉะนั้นแทนที่จะใช้ค่าความน่าจะเป็น เป็นดัชนีความยาก
โดยตรงทีเดียว จึงควรใช้ค่าที่อาจคำนวณได้จากสูตรข้างล่างนี้แทน

$$P = \frac{kp-1}{k-1}$$

P = อัตราส่วนของคำตอบที่ตอบถูกเมื่อปรับค่ากับการเผชิญตอบถูก

p = ค่าความน่าจะเป็นที่คำนวณได้ (โดยถือจำนวนผู้ตอบจริง ๆ เป็น
จำนวนทั้งหมด)

k = จำนวนคำตอบที่ให้เลือกแต่ละคำถาม

ยกตัวอย่างเช่น คำถามข้อหนึ่งมีคำตอบให้เลือก 3 ข้อ มีผู้ตอบถูกในอัตราร้อยละ
80 อัตราส่วนของผู้ตอบถูกเมื่อปรับค่าแล้วจะคำนวณได้ดังต่อไปนี้

$$\begin{aligned} P &= \frac{3(.8)-1}{3-1} \\ &= .70 \end{aligned}$$

Guilford ได้คำนวณค่าความยากที่ปรับค่าแล้ว ในกรณีที่คำถามหนึ่ง ๆ มีคำตอบที่จะ
ให้เลือกระหว่าง 2 ถึง 5 คำถาม และมีผู้ตอบถูกในอัตราส่วนต่าง ๆ กัน ดังแสดงไว้ใน
ตารางที่ 29

⁸ Guilford, *op. cit.*, pp. 419-420.

ตารางที่ 29

ดัชนีความยากของคำถามที่ปรับค่าความเพี้ยนแล้ว

P	K				P	K			
	2	3	4	5		2	3	4	5
.99	.980	.985	.987	.9875	.59	.180	.385	.453	.4875
.98	.960	.970	.973	.9750	.58	.160	.370	.440	.4750
.97	.940	.955	.960	.9625	.57	.140	.355	.427	.4625
.96	.920	.940	.947	.9500	.56	.120	.340	.413	.4500
.95	.900	.925	.933	.9375	.55	.100	.325	.400	.4375
.94	.880	.910	.920	.9250	.54	.080	.310	.387	.4250
.93	.860	.895	.907	.9125	.53	.060	.295	.373	.4125
.92	.840	.880	.893	.9000	.52	.040	.280	.360	.4000
.91	.820	.865	.880	.8875	.51	.020	.265	.347	.3875
.90	.800	.850	.867	.8750	.50	.000	.250	.333	.3750
.89	.780	.835	.853	.8625	.49	.000	.235	.320	.3625
.88	.760	.820	.840	.8500	.48	.000	.220	.307	.3500
.87	.740	.805	.827	.8375	.47205	.239	.3375
.86	.720	.790	.813	.8250	.46190	.280	.3250
.85	.700	.775	.800	.8125	.45175	.267	.3125
.84	.680	.760	.787	.8000	.44160	.253	.3000
.83	.660	.745	.773	.7875	.43145	.240	.2875
.82	.640	.730	.760	.7750	.42130	.227	.2750
.81	.620	.715	.747	.7625	.41115	.213	.2625
.80	.600	.700	.733	.7500	.40100	.200	.2500
.79	.580	.685	.720	.7375	.39085	.187	.2375
.78	.560	.670	.707	.7250	.38070	.173	.2250
.77	.540	.655	.693	.7125	.37055	.160	.2125
.76	.520	.640	.680	.7000	.36040	.147	.2000

P	K				P	K			
	2	3	4	5		2	3	4	5
.75	.500	.625	.557	.6875	.35025	.133	.1875
.74	.480	.610	.653	.6750	.34010	.120	.1750
.73	.460	.595	.640	.6625	.33000	.107	.1625
.72	.440	.580	.627	.6500	.32000	.093	.1500
.71	.420	.565	.613	.6375	.31000	.080	.1375
.70	.400	.550	.600	.6250	.30067	.1250
.69	.380	.535	.587	.6125	.29053	.1125
.68	.360	.520	.573	.6000	.28040	.1000
.67	.340	.505	.560	.5875	.27027	.0875
.66	.320	.490	.547	.5750	.26013	.0750
.65	.300	.475	.533	.5625	.25000	.0625
.64	.280	.460	.520	.5500	.24000	.0500
.63	.260	.445	.507	.5375	.23000	.0375
.62	.240	.430	.493	.5250	.220250
.61	.220	.415	.480	.5125	.210125
.60	.200	.400	.467	.5000	.200000

แหล่งที่มา : Guilford, J.P., *Psychometric Methods*, p. 421.

ดัชนีความยากนัยมีความสำคัญ 2 ประการ คือ ประการแรก ความยากของคำถาม ทำให้ลักษณะการแจกแจงของค่าคะแนนเปลี่ยนไป ความยากเป็นส่วนหนึ่งที่กำหนดค่าเฉลี่ย การกระจาย และค่าเบี่ยงเบนมาตรฐานของค่าคะแนนที่ควรจะได้สำหรับแบบทดสอบหนึ่ง ๆ ดัชนีความยากที่ห่างจาก .5 ไม่ว่าจะค่อนข้างไปทางซ้ายหรือขวา ย่อมทำให้การแจกแจงของค่าคะแนนมีลักษณะเบ้ไป โดยเฉพาะในกรณีที่มีปริมาณคำถามน้อย เช่น น้อยกว่า 20 รายการ ในขณะที่เดียวกันค่าเบี่ยงเบนมาตรฐานก็จะน้อยลงไปด้วย ความสำคัญอีกประการหนึ่งก็คือว่า ค่าดัชนีความยากนี้สัมพันธ์กับสัมประสิทธิ์ความเชื่อมั่นอย่างใกล้ชิด ในเรื่องนี้ไม่ได้ขึ้นอยู่กับความมากหรือน้อยของค่าดัชนีความยาก คำถามที่มีดัชนีความยากต่างกันมาก ย่อมมีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันต่ำ ในขณะที่เดียวกันสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบทั้งฉบับก็จะต่ำลงไปในทาง

ปฏิบัตินั้นจึงพยายามที่จะสร้างคำถามที่มีดัชนีความยากอยู่ระหว่าง .4 ถึง .6 เพื่อให้สามารถได้แบบทดสอบความสัมฤทธิ์ผลที่มีความยากตามสมควร

ในการสร้างแบบทดสอบอื่น ๆ เพื่อการคาดคะเน เช่น แบบทดสอบเชาวน์ หรือแบบทดสอบความถนัดพิเศษ การวิเคราะห์ดัชนีความยากมีความสำคัญเป็นพิเศษ เพราะแบบทดสอบประเภทนี้ต้องไม่มีความยากง่ายเกินไป จนทำให้เสียผลการวัดตัวแปรเพื่อคาดคะเนความสามารถในการปฏิบัติงาน แต่ในการสร้างแบบทดสอบความสัมฤทธิ์ผลนี้ การวิเคราะห์ดัชนีความยากเป็นเรื่องสำคัญรองลงมาจาก การวิเคราะห์ความเที่ยงตรงในแง่เนื้อหา และการวิเคราะห์รายการคำถามในแง่สัมประสิทธิ์สหสัมพันธ์ระหว่างคำถามแต่ละข้อกับรายการคำถามทั้งหมด การสร้างความเที่ยงตรงในแง่เนื้อหาทำให้เชื่อได้พอสมควรว่า คำถามแต่ละข้อมีลักษณะและคุณสมบัติดีแล้ว การวิเคราะห์รายการช่วยในการตัดสินใจคัดเลือกคำถามที่ดีที่สุดไว้ ส่วนการคำนวณดัชนีความยากเป็นเพียงเครื่องเสริมให้สามารถยืนยันได้ว่า คำถามแต่ละข้อที่คัดเลือกไว้นั้นมีลักษณะครบถ้วนตามที่ต้องการจริง ๆ และโดยเฉพาะในกรณีการสร้างแบบทดสอบความสัมฤทธิ์ผลนี้ การสร้างและคัดเลือกคำถามส่วนใหญ่ยังจะต้องอาศัยดุลพินิจและการตัดสินใจเชิงอัตนัยของผู้สร้างเองเป็นหลัก เช่น คำถามต้น ๆ ในแต่ละภาคของแบบทดสอบอาจเป็นคำถามง่าย ๆ ซึ่งผู้สร้างตั้งใจถามเพื่อให้ผู้ถูกวัดมีกำลังใจที่จะทำแบบทดสอบต่อไป

การให้คะแนน

การให้คะแนนแบบทดสอบปรนัยและแบบทดสอบมาตรฐานในการตอบคำถามแต่ละข้อต้องมีการระบุให้แน่ชัดลงไป เพื่อให้สามารถนำไปใช้ได้ทุกโอกาสและทุกสถานที่ให้ได้ประสิทธิผลเหมือนกัน การกำหนดให้การตอบถูกมีน้ำหนักเท่ากับการตอบผิด เช่นที่ผู้ออกข้อสอบทั่วไปปฏิบัติในหลายแห่งนั้นเป็นการไม่ถูกต้อง นอกจากนั้นคำถามที่ผู้เข้าสอบไม่ตอบเพราะตอบไม่ได้ ก็มีน้ำหนักต่างกับคำถามที่ไม่ตอบเพราะยังไม่ได้อ่านคำถาม ถ้าตอบที่ถูกบางข้ออาจจะถูกเพราะความบังเอิญ ในกรณีที่มีการจำกัดเวลา ผู้เข้าสอบต้องพยายามรีบตอบ ทำให้การตอบถูกขึ้นอยู่กับความบังเอิญมากขึ้น

ความบังเอิญตอบถูกกับค่าคะแนน แบบทดสอบปรนัยที่ตั้งคำถามเป็นแบบเลือก (multiple-choice test) ทำให้มีโอกาสบังเอิญตอบถูกได้ ทั้งนั้นขึ้นอยู่กับจำนวนคำตอบที่ให้เลือกรับสำหรับคำถามแต่ละข้อ เช่น ในแบบทดสอบที่มีคำถาม 50 ข้อ ประเภทมีคำตอบให้เลือก 2 ข้อต่อ 1 คำถาม จะมีโอกาสบังเอิญตอบถูกถึง 25 คะแนน ถ้าสมมุติให้

คะแนนคำถามแต่ละข้อเท่ากับหนึ่งคะแนน ($\frac{50}{2} = 25$) และแบบทดสอบที่มีคำถาม 50 ข้อ ประเภทที่คำตอบให้เลือก 5 ข้อต่อ 1 คำถาม จะมีโอกาสบังเอิญตอบถูก 10 คะแนน ($\frac{50}{5} = 10$) ทั้งนี้จะเห็นว่าแบบทดสอบที่มีคำถามให้เลือก 2 คำตอบ ควรจะยากเป็นสองเท่าของแบบทดสอบที่มีคำถามให้เลือก 5 คำตอบ ถ้าประสพก็จะลดโอกาสบังเอิญตอบถูกนี้ลง ถ้าหากแบบทดสอบเหล่านี้มีคำถามที่ยากหรือง่ายเกินไป ก็จะต้องทำให้มีโอกาสบังเอิญตอบถูกเพิ่มขึ้นอีก

ในการพิจารณาโอกาสบังเอิญตอบถูกนี้ ยังจะต้องคำนึงถึงข้อบกพร่องจากการสุ่มตัวอย่างค่าคะแนนที่น่าจะเป็นในการตอบถูกอีกด้วย ในกรณีที่มีการแจกแจงทวินาม (binomial distribution) ค่าผิดพลาดมาตรฐานของคะแนนที่น่าจะเป็นในการตอบถูกก็คือ \sqrt{npq} โดยที่ n คือจำนวนของคำถาม p คือความน่าจะเป็นในการตอบถูก (ซึ่งมีค่าเท่ากับ $\frac{1}{k}$; k คือจำนวนคำตอบที่ให้เลือกแต่ละคำถาม) q คือความน่าจะเป็นในการตอบผิด (ซึ่งมีค่าเท่ากับ $\frac{k-1}{k}$ หรือเท่ากับ $1-p$) ในกรณีของแบบทดสอบที่มีคำถาม 50 ข้อ ให้เลือกคำตอบ 5 ข้อต่อ 1 คำถาม ค่าผิดพลาดมาตรฐานก็คือ $2.83 (\sqrt{50 \cdot (.2) \cdot (.8)})$ ถ้าหากกำหนดให้อันตรภาคความเชื่อมั่นใช้ความน่าจะเป็นเท่ากับ .95 ค่าเบี่ยงเบนก็จะสูงเท่ากับ 5.66 หรือประมาณ 6 คะแนน { $2 (2.83)$ เพราะเท่ากับค่าเบี่ยงเบนมาตรฐาน 2 ช่วง } เมื่อนำไปรวมกับโอกาสบังเอิญตอบถูก (10 คะแนน) ก็แปลว่าผู้ตอบแบบทดสอบนี้มีโอกาสบังเอิญที่จะตอบถูกถึง 16 คะแนน จากคะแนนเต็ม 50 อย่างไรก็ตาม การคำนวณคะแนนที่อาจได้จากการบังเอิญตอบถูกนี้จะต้องวิเคราะห์จากคำถามที่ได้ตอบ ไม่ใช่คำถามที่มีอยู่ทั้งหมดในแบบทดสอบ

สูตรการคำนวณคะแนน เพื่อพยายามหลีกเลี่ยงโอกาสบังเอิญตอบถูกอันจะทำให้การคิดคะแนนที่ได้ผิดไปจากที่ควรจะได้ตามความรู้ความสามารถจริง ๆ จึงได้มีผู้พยายามคิดสูตรเพื่อการคำนวณคะแนนโดยเฉพาะ สูตรนี้เรียกว่า สูตรอนุมานค่าคะแนน (A Priori Scoring Formula)⁹

$$s = r - \frac{w}{k-1}$$

s = คะแนนตามความรู้จริง ๆ

r = คำตอบที่ถูกต้อง

w = คำตอบที่ตอบผิด

k = จำนวนคำตอบที่ให้เลือกแต่ละคำถาม

⁹ *Ibid.*, p. 448.

ในกรณีตัวอย่างแบบทดสอบ 50 คำถาม ที่มีคำตอบเลือกคำถามละ 5 ข้อ ถ้าหาก นาย ก. ตอบถูก 30 ข้อ ตอบผิด 20 ข้อ แทนที่ นาย ก. จะได้คะแนน 30 คะแนน (โดยให้นำหน้าการตอบถูกเท่ากับการตอบผิด) ก็จะต้องนำมาคำนวณคะแนนตามสูตรข้างบน ซึ่งจะได้เท่ากับ 25 คะแนน $(30 - \frac{20}{(5-1)})$

ลักษณะของแบบทดสอบมาตรฐานที่วัดความสัมฤทธิ์ผล

แบบทดสอบมาตรฐานที่วัดความสัมฤทธิ์ผลในปัจจุบันอาจแยกได้เป็น 2 ประเภท คือ แบบทดสอบอาชีพ (trade test) และแบบทดสอบตัวอย่างงาน (work sample test) การใช้แบบทดสอบความสัมฤทธิ์ผลประเภทหนึ่งประเภทใดขึ้นอยู่กับลักษณะบุคคลที่สมัครเข้าปฏิบัติงาน ในกรณีบุคคลที่เพิ่งจบการศึกษามาใหม่ๆ แบบทดสอบอาชีพนับว่าเหมาะสมกว่าแบบทดสอบตัวอย่างงาน เพราะแบบทดสอบอาชีพทำให้สามารถคาดคะเนได้ดีกว่าในประเด็นที่ว่า บุคคลผู้นี้จะมีความชำนาญในการปฏิบัติงานที่ไม่เคยปฏิบัติมาก่อนเลยเพียงใด แต่ถ้าหากจะใช้คัดเลือกรับบุคคลที่เคยผ่านการปฏิบัติงานมาแล้ว ก็ควรใช้แบบทดสอบประเภทหลังมากกว่า¹⁰ ทั้งนี้จะพิจารณาเหตุผลได้จากลักษณะของแบบทดสอบแต่ละประเภท ซึ่งกล่าวไว้ข้างล่างนี้

แบบทดสอบอาชีพ แบบทดสอบประเภทนี้ทั้งแบบสอบปากเปล่าและสอบข้อเขียน เฉพาะแบบสอบปากเปล่าจะไม่ใช่คำถามชนิดให้เลือกหลายคำตอบ เช่น แบบทดสอบข้อเขียนทั่วไป แต่จะเป็นคำถามที่กำหนดให้ผู้เข้าสอบนึกคำตอบเอาเอง สำนักงาน United States Employment Service ได้สร้างแบบทดสอบมาตรฐานแบบปากเปล่าไว้แล้ว ใช้ทดสอบผู้ชำนาญงาน ผู้ฝึกหัดงานหรือผู้ช่วย และผู้ปฏิบัติอาชีพที่เกี่ยวข้องต่างๆ หลายอาชีพ เช่น ช่างติดตั้งตู้เย็น ช่างเดินเครื่องจักรกล ในการทดสอบแต่ละอาชีพจะมีคำถามเพียงประมาณ 15 คำถาม ทั้งนี้อาจใช้ตามในเวลาสอบสัมภาษณ์ทั่วไปโดยใช้เวลาเพิ่มขึ้นจากการสัมภาษณ์ตามปกติอีกไม่เกิน 20 นาที

แบบทดสอบข้อเขียนประกอบด้วยคำถามที่เกี่ยวกับการปฏิบัติงานในอาชีพหนึ่ง ๆ ในการสร้างคำถามมีหลักกว้าง ๆ ว่า จะต้องเป็นคำถามที่ผู้ปฏิบัติงานที่ดีส่วนใหญ่ในอาชีพนั้น ๆ สามารถตอบได้ และในขณะเดียวกันก็เป็นคำถามที่ผู้เริ่มปฏิบัติงานหรือผู้ไม่เคยปฏิบัติงานนั้น ๆ มาก่อนเพียงส่วนน้อยเท่านั้นที่จะสามารถตอบได้ นอกจากนี้ คำถามที่ตงขั่นนี้ ยังควรเป็นคำถามที่ผู้ปฏิบัติงานในอาชีพอื่นหรืออาชีพใกล้เคียงไม่สามารถ

¹⁰ Guion, Robert M., *Personnel Testing*, p. 419.

ตอบได้เลย การตั้งคำถามให้มีลักษณะเช่นนี้เพื่อให้แน่ใจได้ว่า บุคคลที่สามารถตอบคำถามเหล่านี้ได้สามารถปฏิบัติงานที่กำหนดไว้จริง ๆ แบบทดสอบมาตรฐานวัดความสัมฤทธิ์ผลที่มีชื่อเสียงเป็นที่กล่าวถึงหลายแห่งได้แก่ Purdue Vocational Test, Purdue Personnel Test และ Purdue Interview Aids¹¹ ซึ่ง Occupational Research Center ของมหาวิทยาลัย Purdue University เป็นผู้สร้างขึ้น แบบทดสอบนี้เป็นกลุ่มของแบบทดสอบที่ใช้วัดความรู้ และทักษะในงานช่างกล ช่างไฟฟ้า เจ้าหน้าที่อ่านพิมพ์เขียว สารระสำคัญของแบบทดสอบชุดนี้ดังนี้

1. แบบทดสอบช่างกล มี 132 คำถาม เป็นแบบเลือกคำตอบหลายข้อ และแบบจับคู่ผสมกับเนื้อหาที่ตามเกี่ยวกับงานของช่างกล (machinist) และช่างเดินเครื่อง (machine operator) สารระสำคัญที่ถามเป็นเรื่องที่ช่างกลซึ่งมีความรู้จะต้องทราบ และบางส่วนของเรื่องที่ถามมีสอนในหลักสูตรอาชีวศึกษาชั้นสูง สารระสำคัญที่ถามเหล่านี้คลุมถึงการใช้เครื่องมือเครื่องจักรต่าง ๆ เช่น เครื่องกลึง (lathe) เครื่องตัด (milling machine) เครื่องเจาะ (drill press) เครื่องบด (grinder) และเครื่องไส (shaping machine) เป็นต้น เกณฑ์การวัดที่กำหนดไว้มีทั้งประเภทที่วัดความรู้เกี่ยวกับการใช้เครื่องจักรเครื่องกลทั่วไป และความรู้ในการใช้เครื่องจักรแต่ละชนิดโดยเฉพาะ ผู้ที่มีความรู้ในเรื่องนี้จะสามารถตอบแบบทดสอบฉบับนี้เสร็จภายใน 50 นาที

2. แบบทดสอบช่างไฟฟ้า แบ่งเป็น 2 ฟอร์ม มีคำถาม 56 ข้อ กำหนดเวลาตอบ 25 นาที คำถามมีทั้งแบบเลือกคำตอบหลายข้อ และแบบจับคู่ สารระสำคัญที่ถามเกี่ยวกับวงจรไฟฟ้า หน่วยวัด สื่อ และอุปกรณ์ไฟฟ้าต่าง ๆ

3. แบบทดสอบเจ้าหน้าที่อ่านพิมพ์เขียว มี 79 คำถาม ถามในเรื่องการอ่านและการแปลความหมายพิมพ์เขียว ตลอดจนหลักการที่เกี่ยวข้อง

4. แบบทดสอบเจ้าหน้าที่คณิตศาสตร์อุตสาหกรรม มีลักษณะทำนองเดียวกับแบบทดสอบ 3 ฉบับข้างต้น คำถามเกี่ยวกับหลักคณิตศาสตร์และการคิดคำนวณในด้านอุตสาหกรรม ซึ่งผู้ปฏิบัติงานอุตสาหกรรมควรจะต้องทราบ

แบบทดสอบมาตรฐานวัดความสัมฤทธิ์ผลที่มีชื่อในการวัดบุคคลอีกฉบับหนึ่ง คือ Achievement Test in Accounting ของสถาบัน American Institute of Certified Public Accountants บริษัทธุรกิจแห่งใดสนใจที่ขอรับบริการการใช้แบบทดสอบนี้ก็อาจขอให้

¹¹ Tiffin, Joseph, and McCormick, Ernest J., *Industrial Psychology*, p. 500. Lawshe and Balma, *op. cit.*, pp. 108-109.

สถาบันดำเนินการทดสอบให้ได้ ผลการทดสอบจะเปิดเผยให้แก่กับบริษัทเจ้าของเรื่อง โดยตรงเท่านั้น ในการใช้แบบทดสอบนี้สถาบันได้ใช้ควบคู่ไปกับเครื่องมือวัดบุคลิกภาพ ด้วย คือ Strong Vocational Interest Blank และ Preference Record สารสำคัญที่แบบทดสอบนี้วัดเป็นเนื้อหาตามหลักสูตรระดับปริญญาตรี 4 ปี จากผลการวิจัย¹² ปรากฏว่า แบบทดสอบนี้มีคุณภาพในการคาดคะเนผลการสอบเพื่อขอใบรับรองคุณสมบัติของนักบัญชีอาชีพ (CPA examinations) ได้ดีพอสมควร สัมประสิทธิ์สหสัมพันธ์ระหว่างบุคคลที่สอบในชั้นปริญญาตรีสุดท้ายกับบุคคลที่สอบเพื่อขอใบรับรองคุณสมบัติของนักบัญชีอาชีพ มีพิสัยระหว่าง .46 ถึง .61 ส่วนสัมประสิทธิ์ความเที่ยงตรงในแง่การคาดคะเนความสำเร็จในการปฏิบัติงานจริงก็ปรากฏว่า สูงถึง .80

แบบทดสอบตัวอย่างงาน แบบทดสอบนี้สร้างขึ้นโดยคัดเลือกตัวอย่างที่เป็นตัวแทนของงานชิ้นนั้น ๆ มาให้ผู้ที่รับการบรรจุหรือเลื่อนขั้นทดลองปฏิบัติ ตามธรรมดาแบบทดสอบเช่นนี้ มักใช้กับงานที่ต้องใช้เครื่องมือ เครื่องจักร ในการวัดผลการทดสอบตามแนวนี้ ผู้วัดจะต้องถือเกณฑ์มาตรฐาน ซึ่งผู้สร้างแบบทดสอบกำหนดไว้เป็นแนวตัดสินใจว่า บุคคลหนึ่ง ๆ ควรจะได้รับการบรรจุ หรือเลื่อนขั้นให้ไปปฏิบัติงานนั้น ๆ หรือไม่ เกณฑ์มาตรฐานที่กำหนดไว้โดยคำนวณผลการปฏิบัติงานของผู้ที่มีความชำนาญในงานนั้น ๆ โดยเฉลี่ย ตัวอย่างของแบบทดสอบประเภทนี้ที่ควรนำมากล่าว คือ แบบทดสอบของกระทรวงกลาโหมของสหรัฐฯ ซึ่งใช้ทดสอบพนักงานขับรถบรรทุกสินค้า ผู้ที่สมัครเข้าปฏิบัติงานจะต้องทดสอบขับรถผ่านเครื่องกีดขวางตามที่กำหนดไว้เป็นแนวตายตัว ในการนี้จะมีผู้ให้คะแนนตามเกณฑ์ที่กำหนด เช่น การใส่เกียร์ต่ำเกินไป การออกกรด กระตุก การนำล้อรถครูดกับเครื่องกีดขวาง การใช้ปืนจั่นของรถยกสินค้าจากรถโดยไม่ถูกจังหวะเหล่านี้จะต้องเสียคะแนน การที่นำแบบทดสอบนี้มากล่าวไว้ในที่นี้ ก็เพราะอาจจะเห็นได้ว่างานอื่น ๆ ที่ประสงค์จะให้มีแบบทดสอบตัวอย่างงานก็อาจจะทำได้ตามแนวนี้

แบบทดสอบอาชีพเน้นความรู้ในเรื่องงาน ซึ่งผู้ที่มีความสามารถปฏิบัติจะต้องทราบ ส่วนแบบทดสอบตัวอย่างงานเน้นความชำนาญในการปฏิบัติงานนั้น ๆ ฉะนั้นแบบทดสอบตัวอย่างงาน จึงเหมาะแก่การคัดเลือกหรือเลื่อนขั้นบุคคลในแง่ประสิทธิภาพหน่วยงานหรือบริษัทธุรกิจอาจสร้างแบบทดสอบเพื่อวัดความสำเร็จของงาน ใช้ในหน่วยงานของตนโดยเฉพาะ ตามแนวแบบทดสอบทั้ง 2 ประเภทนี้ได้

¹² Guion, *op. cit.*, p. 448.

ปัญหาในการบริหารงานวัดผลบุคคล

เท่าที่กล่าวมาแล้วทั้ง 12 บท จะเห็นว่า การวัดผลและประเมินบุคคล เพื่อให้ได้บุคคลที่มีความรู้ความสามารถจริง ๆ ปฏิบัติงานในตำแหน่งหน้าที่ที่เหมาะสมนั้น เป็นเรื่องที่ต้องใช้หลักวิชาอย่างกว้างขวาง แต่การดำเนินงานในด้านนี้ ก็มีไหว้จะอาศัยความรู้ในวิชาการดังกล่าวแต่เพียงอย่างเดียวเป็นหลักในการปฏิบัติให้ได้ผลดี การที่จะจัดการวัดผลและประเมินบุคคลให้สำเร็จด้วยดีนั้น ยังต้องอาศัยความสามารถในการบริหารด้วย

การที่หน่วยราชการหรือบริษัทธุรกิจแห่งหนึ่งแห่งใด จะจัดให้มีการวัดผลบุคคลที่ดีในหน่วยงานของตนได้ จะต้องมีการดำเนินงานเป็นขั้นตอนที่ถูกต้อง การที่เพียงแต่มีบุคคลสองสามคนที่กระตือรือร้นในการวัดผลตามหลักวิชา หรือเพียงแต่มีการรวบรวมแบบทดสอบมาตรฐานที่มีจำหน่ายในท้องตลาดของต่างประเทศนั้น มิได้เป็นเครื่องประกันในการที่จะทำให้หน่วยงานนั้นมีการวัดผลที่ดี การที่เน้นความสำคัญของการบริหารงานวัดผลและประเมินบุคคลเช่นนี้ก็เพื่อที่จะชี้ว่า การวัดผลบุคคลที่ดีมิได้ขึ้นอยู่กับจำนวนแบบทดสอบมาตรฐานที่มีอยู่ในหน่วยงาน แม้ว่าในบทต่าง ๆ ได้ย้ำทฤษฎีการสร้างแบบทดสอบอย่างมากมายก็ตาม แต่การวัดผลบุคคลที่ดี จะต้องมึคุณภาพ 2 ประการคือ ประการแรก ต้องมีลักษณะเป็นเอกมัย (uniformity) คือ เป็นแบบเหมือนกัน ในการดำเนินการวัดผลโดยตลอด และประการที่สอง จะต้องมึประสิทธิภาพ (efficiency) ประสิทธิภาพในการบริหารงานวัดผลนั้นถือเอาจากการที่สามารถแสดงผลของการวัดผลโดยไม่ชักช้าอย่างหนึ่ง จากการที่สามารถใช้หลักฐานของการวัดผลในครั้งที่ผ่าน ๆ มาเป็นแนวอ้างอิงได้อย่างหนึ่ง และจากการที่สามารถประหยัดค่าใช้จ่ายในการดำเนินงานน้อีกอย่างหนึ่ง

ในการพิจารณาปัญหาของการบริหารงานเรื่องนี้ จะแยกวิเคราะห์ออกเป็น 2 ตอน เพื่อให้สะดวกแก่การทำความเข้าใจได้ง่าย คือ ตอนแรก ว่าด้วยการจัดตั้งหน่วย วัตถุประสงค์ และตอนที่สอง ว่าด้วยปัญหาต่างๆ ในระหว่างการดำเนินงานวัตถุประสงค์

หน่วยวัตถุประสงค์

การที่จะจัดการวัตถุประสงค์และประเมินบุคคลให้สำเร็จด้วยดี จำเป็นจะต้องมีหน่วยงานที่ประกอบด้วยผู้ได้รับการฝึกฝนวิชาการด้านนี้มาโดยเฉพาะ ทำหน้าที่รับผิดชอบในเรื่องนี้โดยตรง ลักษณะและคุณสมบัติของผู้ที่มีความรู้ในวิชาการด้านนี้ได้มีอธิบายในบทต่างๆ ที่ผ่านมาแล้ว จึงจะไม่กล่าวซ้ำในที่นี้อีก อย่างไรก็ตาม ในองค์การ (organization) ที่ไม่ใหญ่โตกว้างขวาง อาจจะไม่สามารถจัดตั้งหน่วยวัตถุประสงค์ขึ้นได้ ในกรณีเช่นนี้ก็อาจอาศัยผู้เชี่ยวชาญในเรื่องนี้ที่ปฏิบัติงานประจำในมหาวิทยาลัย หรือบริษัทธุรกิจที่ทำหน้าที่เรื่องนี้โดยเฉพาะ ในที่นี้จะกล่าวถึงแต่ในกรณีที่องค์การสามารถตั้งหน่วยวัตถุประสงค์ของตนเองได้ องค์การเช่นนี้อาจเป็นหน่วยราชการหรือบริษัทธุรกิจใหญ่ๆ ก็ได้

โครงสร้างของหน่วยงานวัตถุประสงค์มีความสำคัญต่อการดำเนินงานเป็นอย่างมาก ถ้าหากหน่วยวัตถุประสงค์รวมอำนาจไว้ที่สำนักงานกลางแต่แห่งเดียว¹ นักวิชาการในการวัตถุประสงค์ก็อาจมีอำนาจมากเกินไป เพราะเป็นบุคคลผู้เดียวหรือกลุ่มเดียวที่สามารถตัดสินใจว่าจะใช้แบบทดสอบฉบับใด จะรวบรวมข้อมูลเกี่ยวกับการวัตถุประสงค์ด้วยวิธีใด ตลอดจนการที่จะปกปิดหรือเปิดเผยข้อเท็จจริงประเภทใดที่เกี่ยวกับผลการทดสอบ การรวมอำนาจไว้ที่ศูนย์กลาง (centralization) เช่นนี้ อาจทำให้ผู้บริหารของหน่วยต่างๆ ในองค์การรู้สึกเกิดความลำบากใจหรือไม่สบายใจ เป็นทางให้เกิดทัศนคติไม่ดีต่อการบริหารงานการวัตถุประสงค์ อย่างไรก็ตาม ในทางปฏิบัตินั้น องค์การใหญ่ๆ เช่น ส่วนราชการของประเทศต่างๆ มิได้บริหารงานวัตถุประสงค์โดยรวมอำนาจไว้ที่ศูนย์กลางอย่างเต็มที่ ส่วนบริษัทธุรกิจใหญ่ๆ ส่วนใหญ่ก็มักบริหารงานประเภทนี้โดยการกระจายอำนาจ (decentralization) ในการกระจายอำนาจการบริหารงานวัตถุประสงค์นั้นทำได้โดยส่วนกลางกำหนดมาตรฐานในการวัตถุประสงค์ไว้ ผู้บริหารงานสนามหรือหน่วยงานในส่วนภูมิภาค มีอำนาจในการดำเนินการวัตถุประสงค์ และการแปลความหมายของผลคะแนนที่ได้โดยอาศัยมาตรฐานที่กำหนดไว้ นั้นเป็นหลัก แต่การกระจายอำนาจในการวัตถุประสงค์ก็มีผลเสียหลายอย่าง เช่น เครื่องมือวัดที่ใช้ อาจหย่อนความเที่ยงตรงลงไปกว่าที่ควรจะเป็น เพราะผู้บริหารงานสนามหรือหน่วย

¹ Pfiffner, John M., and Presthus, R. Vance, *Public Administration*, pp. 256–259.
Waldo, Dwight, ed., *Ideas and Issues in Public Administration*, pp. 166–168.

งานภูมิภาคไม่เข้าใจ หรือไม่สนใจที่จะใช้เครื่องมือวัดนั้นให้สอดคล้องกับหลักการและกฎเกณฑ์ที่กำหนดไว้ ในกรณีเช่นนี้นักวิชาการของส่วนกลางมีหน้าที่ต้องทำให้ผู้เกี่ยวข้องเข้าใจและยอมรับความสำคัญของการควบคุมการใช้เครื่องมือวัดผลตามวิธีที่ถูกต้อง แต่ปัญหาที่ยังอาจเกิดขึ้นอีกว่า นักวิชาการของส่วนกลางอาจไม่ได้รับความร่วมมือจากผู้บริหารงานสนามหรือหน่วยงานภูมิภาค ถ้าเป็นดังนี้ นักวิชาการของส่วนกลางก็อาจหาวิธีระดมการร่วมมือของผู้บริหารเหล่านี้ โดยทำการวิจัยให้ออกมาเป็นหลักฐานได้ว่า การวัดผลครั้งต่าง ๆ ที่แต่ละหน่วยได้ดำเนินการไปนั้นมีคุณค่าในเชิงคาดคะเนมากน้อยเพียงใด พร้อมทั้งมีการเปรียบเทียบผลของความเข้าใจในการใช้ตัวแปรการวัด (predictor variables) ระหว่างหน่วยงานต่าง ๆ ภายในองค์กรของตน

สิ่งสำคัญอีกประการหนึ่งที่จะต้องพิจารณา ถ้าประสงค์จะให้มีหน่วยวัดผลที่มีประสิทธิภาพ คือ การกำหนดนโยบายในเรื่องการวัดผลและประเมินบุคคลขององค์กรให้แน่ชัดลงไป เป็นที่เห็นสอดคล้องกันในบรรดาผู้ที่เคยปฏิบัติงานด้านนี้ว่า² นโยบายในเรื่องนี้นอกจากจะกำหนดให้แน่ชัดแล้ว ยังจะต้องเขียนไว้เป็นลายลักษณ์อักษร เพื่อเป็นหลักฐานอีกด้วย นโยบายที่บันทึกไว้เป็นลายลักษณ์อักษรนี้ควรจะต้องมีการปรับปรุงแก้ไขให้ทันสมัยกับกาลและเวลาด้วย การที่ต้องกำหนดให้เป็นหลักฐานลงไปเช่นนี้ก็เพื่อป้องกันความขัดแย้งระหว่างหน่วยวัดผลกับหน่วยงานอื่น ๆ ในองค์กรเดียวกันในเรื่องที่ว่าใครเป็นผู้มีอำนาจในการดำเนินการวัดผล และจะใช้ผลของการประเมินบุคคลในลักษณะใดบ้าง นอกจากนี้การที่มีนโยบายแน่ชัดในเรื่องนี้ ยังช่วยทำให้เกิดความกระจ่างในนโยบายเรื่องการบริหารงานบุคคลในแง่ต่าง ๆ เพิ่มขึ้นอีกด้วย นโยบายในเรื่องการวัดผลและประเมินบุคคล ควรจะครอบคลุมประเด็นต่าง ๆ⁴ เรื่องดังต่อไปนี้ เป็นอย่างน้อย คือ เจตนารมณ์ของโครงการวัดผล งานบุคคลในโครงการ อำนาจในการวินิจฉัยสั่งการและงบประมาณ

เจตนารมณ์ของโครงการวัดผล การกำหนดความมุ่งหมายของโครงการ เป็นสิ่งจำเป็นต่อการวัดผลบุคคล เพราะการใช้เครื่องมือวัดบุคคลนี้อาจมุ่งประโยชน์ได้หลายทางการใช้ผลการประเมินในลักษณะต่าง ๆ จะกล่าวถึงในบทต่อไป ในการกำหนดความมุ่งหมายนี้ ปัญหาหลักคือการตัดสินใจว่าจะใช้เครื่องมือวัดเพื่อช่วยในการวินิจฉัยสั่งการในเรื่องงานบุคคล หรือจะดำเนินการเรื่องนี้เพื่อประโยชน์ในการวิจัยปัญหาการบริหารงานขององค์กรนั้น ๆ เป็นส่วนรวม ถ้าจะกระทำเพื่อประโยชน์ในการบริหารงานบุคคล

² Lawshe, C.H. and Balma, Michael J., *Principles of Personnel Testing*, pp. 56—57.
Guion, Robert M., *Personnel Testing*, p. 494.

เป็นส่วนสำคัญ ก็ยังต้องพิจารณาให้แน่ชัดลงไปอีกว่า จะใช้ผลการทดสอบตาม เครื่องมือวัดให้กว้างขวางเพียงใด เช่น จะถือว่าคะแนนผ่านที่กำหนดไว้สำหรับเครื่องมือ วัดนี้ เป็นคะแนนตัดสินว่า ผู้ที่ไม่ผ่านจะไม่ได้รับการพิจารณาเลย หรือจะใช้คะแนนที่ ได้ตามเครื่องมือวัดเป็นเพียงสิ่งประกอบการพิจารณาของผู้มีอำนาจในการตัดสินคัดเลือก เท่านั้น จะรวบรวมเครื่องมือวัดบุคคลเท่าที่มีอยู่ในท้องตลาดและปรากฏว่าดีมาใช้เท่านั้น หรือจะพยายามสร้างเครื่องมือวัดใหม่ที่มีอำนาจในการคาดคะเนสูงขึ้นสำหรับใช้ในองค์ การของตน ดังนี้เป็นต้น

งานบุคคลในโครงการ นโยบายที่กำหนดจะต้องระบุคุณสมบัติของผู้ที่จะรับผิดชอบ โครงการนี้ไว้ด้วย การวัดผลบุคคลในองค์การหนึ่ง ๆ จะสำเร็จได้ด้วยดีต้องอาศัยว่าผู้ บริหารงานวัดผลเป็นบุคคลที่มีความสามารถเป็นประการสำคัญ ถ้าหากผู้รับผิดชอบ โครงการเป็นเพียงแต่บุคคลที่สนใจในกิจการวัดผลเท่านั้น ก็ไม่มีอะไรเป็นหลักประกัน ว่าเครื่องมือวัดที่จะใช้นั้น ได้คัดเลือกมาแล้วว่ามีความเที่ยงตรงตามหลักวิชาอย่างแท้จริง และถ้าการปฏิบัติงานด้านนี้ขาดหลักวิชา ก็อาจทำให้คาดได้ว่า จะดำเนินการไปได้ ไม่นาน เพราะผลงานไม่เป็นที่เชื่อถือแก่บุคคลในฝ่ายต่าง ๆ เมื่อพิจารณาหลักการของ วิชาวัดผลเท่าที่กล่าวมาแล้วในบทก่อน ๆ อาจสรุปได้ว่า ผู้บริหารหน่วยวัดผลจะต้องมี ความรู้ในเรื่องต่อไปนี้เป็นอย่างน้อย คือ เทคนิคของการวิเคราะห์งานและการประเมิน ผลงาน หลักจิตวิทยาในเรื่องลักษณะนิสัยของบุคคล ความรู้ในเรื่องประโยชน์และข้อ จำกัดของสถิติ ตลอดจนความสามารถในการแปลความหมายของผลการใช้เครื่องมือวัด แบบต่าง ๆ

นอกจากจะกำหนดคุณสมบัติของผู้บริหารหน่วยวัดผลให้แน่ชัดลงไปแล้ว ยังจะ ต้องระบุอีกด้วยว่า เจ้าหน้าที่ต่าง ๆ ในโครงการควรแบ่งงานรับผิดชอบการดำเนินงาน นี้อย่างไรบ้าง ผู้ใดควรจะเป็นผู้รับผิดชอบการทดสอบ ผู้ใดควรจะเป็นผู้รักษา ระเบียบต่าง ๆ ในการนี้อาจกำหนดเป็นตารางให้ดูง่ายขึ้นอีกก็ได้

อำนาจในการวินิจฉัยสั่งการ สิ่งที่สำคัญที่สุดในเรื่องนโยบายของการจัดหน่วยวัดผล ก็คือ การกำหนดขอบเขตอำนาจความรับผิดชอบของหน่วยงานนี้ ในเรื่องนี้เมื่อพิจารณา เฝิน ๆ อาจเห็นว่าไม่น่ามีปัญหาอะไร เพราะหน่วยวัดผลก็ต้องมีอำนาจในการวัดผล และประเมินบุคคล และต้องรับผิดชอบให้ผลการทดสอบบุคคลต่าง ๆ ตามเครื่องมือ ได้ปรากฏออกมาไม่ให้ชักช้า แต่ในการบริหารหน่วยวัดผลนั้น หากถือตามเกณฑ์ เช่นนี้ จะทำให้การดำเนินงานเรื่องนี้ไม่สำเร็จผลสมบูรณ์ เพราะการวัดผลเป็นเรื่องที่

เกี่ยวข้องกับหน่วยงานทุกหน่วยในองค์กรเดียวกัน จึงไม่สามารถที่จะถือตนว่าเป็นหน่วยอิสระ และถ้าปฏิบัติได้สองประการตามที่กล่าวข้างต้นนี้แล้ว งานวัดผลก็จะดำเนินไปด้วยดีเอง นโยบายของการจัดหน่วยวัดผลจะต้องระบุให้ชัดว่า ผู้ใดเป็นผู้มีอำนาจรับผิดชอบในการวินิจฉัยสั่งการว่า ผู้สมัครหรือผู้ที่ได้รับการทดสอบสมควรได้รับการบรรจุ หรือได้รับการฝึกอบรมเพิ่มเติม หรือได้รับการเลื่อนขั้นเลื่อนตำแหน่ง หรือยังไม่สมควรที่จะได้รับประโยชน์ข้อหนึ่งข้อใดตามที่กล่าวนี้ การมอบอำนาจในเรื่องนี้มีความสำคัญต่อความสำเร็จของหน่วยวัดผล เพราะรายละเอียดในการดำเนินงานของหน่วยงานนี้ มีความสัมพันธ์อย่างใกล้ชิดกับการที่จะวินิจฉัยสั่งการในเรื่องดังกล่าว

ปัญหาในประเด็นนี้ก็คือว่า หัวหน้าหน่วยวัดผลควรมีส่วนเสี่ยงเพียงใดในการกำหนดนโยบายและวิธีการ ในการตัดสินว่าผู้ใดมีคุณสมบัติสมควรได้รับการบรรจุหรือฝึกอบรมหรืออื่นๆ เมื่อพิจารณาในแง่ของเหตุผลที่ควรจะเป็น ก็อาจถือว่า หัวหน้าหน่วยวัดผลเป็นบุคคลที่ทราบดีที่สุดว่า ผลการทดสอบตามเครื่องมือวัดหนึ่งๆ มีความหมายในเชิงสถิติอย่างไร และในการวัดผลเท่าที่ได้กระทำไปนั้น มีตัวแปรและข้อจำกัดอันใดบ้าง ฉะนั้น ในการวินิจฉัยสั่งการของผู้บริหารองค์กรในเรื่องที่เกี่ยวกับว่า จะใช้ตัวแปรการวัดผลอันใดเป็นหลักในการประเมินบุคคล หรือนำผลการทดสอบตามเครื่องมือวัดหนึ่งๆ ไปประกอบการพิจารณาเรื่องใดนั้น หัวหน้าหน่วยวัดควรมีส่วนมีเสี่ยงในเรื่องนี้ อย่างเต็มที่ และถ้าหัวหน้าหน่วยวัดผลคัดค้านไม่เห็นสมควรกระทำอันใดในเรื่อง 2 ประการนี้ โดยแสดงเหตุผลทางวิชาการในแง่ความเที่ยงตรงของเครื่องมือวัด หรือในแง่จรรยาของผู้ใช้เครื่องมือวัด ผู้บริหารองค์กรก็ควรจะรับฟัง อย่างไรก็ดี การที่จะวินิจฉัยสั่งการขั้นสุดท้ายว่า ผู้ใดจะได้รับการบรรจุหรือเลื่อนขั้นอาจขึ้นอยู่กับนโยบายการบริหารงานขององค์กรหนึ่งๆ โดยเฉพาะ ผู้บริหารองค์กรอาจเห็นสมควรไม่ให้นำหน่วยวัดผลมีส่วนในการตัดสินเรื่องนี้ โดยเก็บอำนาจนี้ไว้กับตนเอง หรือมอบให้หัวหน้าหน่วยวัดลงไปที่ไม่ใช่หน่วยวัดผลก็ได้ หากเป็นเช่นนี้ก็ควรมีการกำหนดให้แน่ชัดเสียก่อนตั้งแต่เริ่มตั้งหน่วยวัดผล เพื่อมิให้ต้องเสียเวลาถกเถียงขัดแย้ง หรือเกิดความเข้าใจสับสนระหว่างหัวหน้าหน่วยต่าง ๆ ในชั้นหลัง แต่ถึงจะให้หน่วยวัดผลมีอำนาจร่วมในการตัดสินเรื่องเหล่านี้หรือไม่ก็ตาม หัวหน้าหน่วยวัดผลก็ควรจะได้รับสิทธิเข้าร่วมประชุมในการพิจารณาเรื่องนี้ด้วย ทางหน่วยวัดผลเองก็จะได้เข้าใจว่าตนมีอำนาจและมีส่วนมีเสี่ยงในเรื่องเหล่านี้เพียงใด และหน่วยงานอื่นๆ ในองค์กรเดียวกัน ก็จะไม่เกิดความข้องใจในการปฏิบัติงานต่าง ๆ ของหน่วยวัดผล

งบประมาณ การที่จะตั้งหน่วยวัดผลที่ยั่งยืน จำต้องมีการจัดสรรงบประมาณเพื่อ
 การนี้โดยเฉพาะ ผู้ที่ควบคุมการใช้จ่ายเงินงบประมาณเพื่อการนี้ อาจเป็นหัวหน้าหน่วย
 นี้เอง หรือผู้บริหารในระดับสูงกว่าขึ้นไป ทั้งนี้ขึ้นอยู่กับการจัดวางหน่วยงานนี้ ใน
 ระบบบริหารขององค์การนั้น ๆ หน่วยวัดผลอาจเป็นหน่วยอิสระ หรืออาจขึ้นอยู่กับฝ่าย
 บริหารงานบุคคล หรือฝ่ายแรงงานสัมพันธ์ (labor relations) ก็ได้ ทั้งนี้ขึ้นอยู่กับขนาด
 ของงานวัดผลที่ประสงค์จะให้เกิดขึ้นในองค์การ และขนาดความใหญ่โตขององค์การ
 นั้น ๆ จะเป็นในกรณีใดก็ตาม งบประมาณที่จัดสรรไว้จะต้องให้ครอบคลุมค่าใช้จ่าย
 ที่เป็นหลัก 2 ประการ อันนอกเหนือไปจากค่าจ้างเงินเดือนของผู้ปฏิบัติงานในหน่วยวัด
 ผล คือ ค่าใช้จ่ายในการซื้อแบบทดสอบมาตรฐานประการหนึ่ง และค่าใช้จ่ายในการ
 ดำเนินการวิจัยในเรื่องการวัดผลภายในองค์การอีกประการหนึ่ง ในกรณีที่หน่วย
 วัดผลขนาดใหญ่ ก็ยังคงจะต้องงบประมาณเพื่อการดำเนินงานอื่น ๆ เพิ่มเติมอีก เช่น
 ค่าใช้จ่ายในการส่งเจ้าหน้าที่ของหน่วยไปร่วมประชุมทางวิชาการในด้านวัดผลที่อาจมีขึ้น
 ค่าซื้อเอกสารทางวิชาการวัดผล และค่าจ้างผู้ชำนาญการจากภายนอกมาให้คำปรึกษาแก่
 หน่วยวัดผลเป็นครั้งคราว เป็นต้น

เพื่อให้เห็นจริงในเรื่องการกำหนดนโยบายด้านต่าง ๆ ของการจัดงานวัดผลบุคคลใน
 หัวข้อที่ 4 ที่กล่าวมาแล้ว จะได้นำตัวอย่างนโยบายขององค์การที่สมมติว่ามีขนาดค่อนข้างใหญ่
 มาแสดงไว้ในที่นี้³

“การตั้งกองวัดผลบุคคลมีความมุ่งหมายเพื่อให้หน่วยงานนี้ทำหน้าที่ประมวลข้อ
 เท็จจริง และเสนอแนะในการเลือกและบรรจุเจ้าหน้าที่ของฝ่ายต่าง ๆ ตลอดจนกำหนด
 มาตรฐานของงานในตำแหน่งหน้าที่ต่าง ๆ ด้วย”

“กองวัดผลบุคคลมีอำนาจรับผิดชอบในการประมวลข้อเท็จจริงอันเชื่อถือได้เกี่ยวกับ
 บุคคลที่สมัครเข้าทำงานและเจ้าหน้าที่ต่าง ๆ ที่กำลังปฏิบัติงานในขณะนี้ ตามแต่หัวหน้า
 ฝ่ายบริหารงานบุคคลจะกำหนด ในการประมวลข้อเท็จจริงเหล่านี้ กองวัดผลบุคคลมี
 หน้าที่ใช้เครื่องมือวัดอันเป็นปรนัยให้ถูกต้องตามหลักวิชา และมีหน้าที่รายงานผลการ
 ใช้เครื่องมือวัดเหล่านี้ แต่กองวัดผลบุคคลไม่มีหน้าที่ตัดสินว่าจะต้องรับสมัครผู้ใดหรือ
 วินิจฉัยสั่งการอันใดเกี่ยวกับการปฏิบัติงานของเจ้าหน้าที่ในฝ่ายต่าง ๆ”

“กองวัดผลบุคคลมีหน้าที่ให้คำปรึกษาแนะนำแก่หัวหน้าฝ่ายต่าง ๆ ในการประมวล
 และวิเคราะห์ข้อเท็จจริงเกี่ยวกับผู้สมัครเข้าทำงาน และเจ้าหน้าที่ที่กำลังปฏิบัติงานอยู่

³ Lawshe and Balma, *op. cit.*, pp. 58-59.

ในปัจจุบัน เพื่อประโยชน์ในการตัดสินใจ ผู้ใดมีคุณสมบัติเหมาะสม ควรบรรจุเข้าปฏิบัติงานและย้ายตำแหน่ง ตลอดจนการใช้หลักเกณฑ์อันใดเป็นมาตรฐานในการพิจารณาเกี่ยวกับตำแหน่งหน้าที่ต่าง ๆ...”

“นโยบายทั้งหมดข้างต้นนี้ จะเปลี่ยนแปลงเพิ่มเติมไม่ได้ ยกเว้นเสียแต่จะได้รับคำอนุมัติจากผู้จัดการฝ่ายโรงงาน โดยคำแนะนำเป็นลายลักษณ์อักษรของคณะกรรมการเพื่อการวัดผลบุคคลเสียก่อน คำใช้จ่ายในการดำเนินการของกองวัดผลบุคคล จะกำหนดไว้ในงบประมาณของฝ่ายบริหารงานบุคคล”

การดำเนินการวัดผล

แม้จะปรากฏว่าในท้องตลาดต่างประเทศมีแบบทดสอบมาตรฐานที่ได้ค้นคว้าทดลองแล้วมากมายหลายประเภทหลายฉบับ การที่จะดำเนินการวัดผลให้ได้ผลดีก็มิใช่จะทำได้ง่าย ๆ โดยซื้อแบบทดสอบมาตรฐานเหล่านั้นมาใช้ให้ตรงตามชื่อของแบบทดสอบแล้วเป็นการใช้ได้ จริงอยู่แบบทดสอบเหล่านั้นเป็นผลของการวิจัยที่พิสูจน์แล้วทั้งนั้น แต่การวัดลักษณะนิสัยของบุคคลในองค์การหนึ่ง ๆ มิใช่ทำได้โดยเพียงแต่นำแบบทดสอบที่วัดลักษณะนิสัยนั้น ๆ มาใช้เท่านั้น การปฏิบัติงานให้ได้ผลสำเร็จในแต่ละหน่วยงานไม่ได้ขึ้นอยู่กับความสามารถและลักษณะนิสัยเหล่านั้นแต่ประการเดียว ยังขึ้นอยู่กับภูมิหลังขององค์การนั้น ๆ ด้วย เช่น ลักษณะของการควบคุมงาน บรรยากาศภายในองค์การและความร่วมมือเป็นน้ำหนึ่งอันเดียวกันของผู้ร่วมงานอื่น ๆ เป็นต้น ฉะนั้นในการคาดคะเนศักยภาพของผู้ปฏิบัติงานในหน่วยหนึ่ง ๆ จึงต้องมีองค์ประกอบบางอย่างต่าง ๆ กันไป

ในการดำเนินการวัดผลให้ได้ผลดีสอดคล้องกับลักษณะขององค์การแต่ละแห่ง จึงควรจะต้องดำเนินเป็นขั้นตอน 3 ขั้น คือ การเลือกใช้เครื่องมือวัดที่เหมาะสมเป็นประการแรก การควบคุมการทดสอบให้ถูกต้องเป็นประการที่สอง และการรักษาระเบียบการวัดผลให้สมบูรณ์ เพื่อให้สามารถรายงานผลการทดสอบที่เชื่อมั่นได้เป็นประการสุดท้าย

การเลือกเครื่องมือวัด ในการเลือกเครื่องมือวัดบุคคล หน่วยวัดผลมีโอกาสเลือก 2 อย่าง คือ ซื้อแบบทดสอบมาตรฐานที่มีในท้องตลาดมาใช้ หรือสร้างแบบทดสอบของตนเองขึ้น

แบบทดสอบมาตรฐานในท้องตลาดนั้น แม้จะเป็นผลจากการวิจัยทดลองแล้วก็ตาม แต่ก็มีข้อบกพร่องในการที่จะนำมาใช้กับหน่วยวัดผลแห่งหนึ่งแห่งใดโดยเฉพาะ คือ

ประการแรก ถึงแม้จะมีแบบทดสอบมากมายหลายฉบับ แต่ลักษณะนิสัยที่ประสงค์จะวัดสำหรับตำแหน่งงานหนึ่งในองค์การหนึ่ง อาจเป็นลักษณะนิสัยที่นอกเหนือไปจากลักษณะนิสัยที่แบบทดสอบมาตรฐานในห้องตลาดจะวัดได้ ประการที่สอง แม้แบบทดสอบมาตรฐานเหล่านี้จะได้รับการกวัดขุ่นในการปกปิดเป็นความลับเฉพาะบริษัทผู้พิมพ์และจำหน่ายเท่านั้นก็ตาม แต่ก็อาจรั่วไหลได้ จริงอยู่จากข้อเท็จจริงที่ปรากฏ ก็ไม่มีหลักฐานว่า แบบทดสอบนี้ได้มีการรั่วไหลจนเป็นที่อึกทึกครึกโครม และประการที่สาม โดยเฉพาะสำหรับประเทศไทย การใช้แบบทดสอบของต่างประเทศอาจไม่เหมาะสมโดยตรง เพราะได้กล่าวไว้แล้วในเรื่องการวัดสติปัญญาว่า ความแตกต่างของสภาพแวดล้อมและวัฒนธรรมระหว่างชาติหนึ่ง ๆ ทำให้ความเที่ยงตรงของเครื่องมือวัดคลาดเคลื่อนไป แม้ในการวัดลักษณะทางจิตวิทยาต่าง ๆ จะไม่มีอุปสรรคเท่ากับการวัดสติปัญญา แต่แบบทดสอบมาตรฐานเหล่านี้ ก็มีลักษณะที่นำมาใช้กับบุคคลที่เป็นคนไทยทันทีไม่ได้สนิทใจนัก

ส่วนการที่หน่วยวัดผลจะสร้างแบบทดสอบของตนเองขึ้น แม้จะมีใช่เป็นการเหลือวิสัย และยังได้ประโยชน์ในแง่อุปสรรค 2 ประการของแบบทดสอบในห้องตลาดที่กล่าวมาแล้ว แต่การที่จะสร้างแบบทดสอบให้มีลักษณะลึกซึ้งโดยทำการวิจัยทดลองให้กว้างขวางอย่างแบบทดสอบมาตรฐานในห้องตลาดนั้นย่อมกระทำได้ยาก เพราะการสร้างแบบทดสอบมาตรฐานของห้องตลาดนั้นใช้กำลังคนผู้เชี่ยวชาญมากมาย ใช้เวลาเป็นแรมปีในการวิเคราะห์คำนวณค่าความเที่ยงตรงและความเชื่อมั่น กับสัมพันธ์อย่างกว้างขวาง นอกจากนี้ยังจะต้องเสียค่าใช้จ่ายลงทุนอีกเป็นจำนวนมาก ซึ่งถ้าหน่วยวัดผลแห่งหนึ่ง ประสงค์จะให้แบบทดสอบของตนเองมีคุณภาพทัดเทียมกับแบบทดสอบเหล่านี้แล้ว ก็อาจเกินกำลังของหน่วยวัดผลแต่ละหน่วยจะกระทำได้

การที่หน่วยวัดผลจะเลือกใช้แบบทดสอบประเภทใดก็ตาม ก่อนที่จะตัดสินใจให้ถูกต้องได้ในเรื่องนี้ ก็จำเป็นต้องมีการวิเคราะห์งาน และประเมินผลงานของตำแหน่งต่าง ๆ ในองค์การของตนให้เรียบร้อยเสียก่อน รายละเอียดในเรื่องนี้ได้มีกล่าวไว้แล้วในบทที่ 2 เมื่อวิเคราะห์งานแล้วก็จะทำให้ทราบว่า การปฏิบัติงานในตำแหน่งหน้าที่หนึ่ง ๆ ให้สำเร็จได้ผลดีนั้นแยกแยะรายละเอียดได้อย่างไรบ้าง และผู้ที่ปฏิบัติงานได้ตามมาตรฐานนี้ต้องมีลักษณะนิสัยสำคัญอันใดบ้าง ข้อเท็จจริงทั้งสองประการนี้จะช่วยให้สามารถพิจารณาต่อไปได้ว่า จะต้องมีการวัดผลในด้านใดบ้าง ในเรื่องอะไร ซึ่งจะทำให้ทราบว่า ควรจะใช้แบบทดสอบประเภทใดต่อไป

หากหน่วยวัดผลประสงค์จะสร้างแบบทดสอบของตนเอง ก็จะต้องดำเนินงานเป็นขั้น ๆ ตามที่กล่าวไว้ในบทที่ผ่านมาเป็นลำดับ แต่ถ้าประสงค์จะเลือกใช้แบบทดสอบมาตรฐานที่จำหน่ายในท้องตลาด ก็จะต้องพิจารณาเนื้อหาของตัวอย่างแบบทดสอบนั้น ๆ พร้อมทั้งวิเคราะห์รายละเอียดที่มีปรากฏในคู่มือประจำแบบทดสอบแต่ละฉบับอย่างรอบคอบ ในการวิเคราะห์นั้นก็นำหลักของการสร้างแบบทดสอบที่ดีมาเป็นแนวทางพิจารณานั้นเอง ผู้เลือกใช้แบบทดสอบอาจพิจารณาเป็นข้อ ๆ ได้ดังต่อไปนี้

1. เนื้อหาของแบบทดสอบ
2. เกณฑ์การวัดของแบบทดสอบ
3. ค่าความเชื่อมั่น
4. ค่าความเที่ยงตรง
5. ระยะเวลาในการตอบแบบทดสอบ และการคำนวณผล

ในเรื่องเนื้อหาของแบบทดสอบ จะต้องถามว่า แบบทดสอบนี้ใช้วัดอะไร ผู้สร้างแบบทดสอบมีความประสงค์เช่นใด วิธีดำเนินการในการสร้างแบบทดสอบสอดคล้องกับความประสงค์ที่บอกไว้หรือไม่ ลักษณะของคำถามแต่ละข้อเป็นเช่นใด ขัดแย้งกับหลักการวิเคราะห์ข้อใดข้อหนึ่งหรือไม่ มีค่าความเที่ยงตรงในแง่แนวคิดหรือไม่ (construct validity) ถ้ามี เป็นความเที่ยงตรงที่พิสูจน์โดยคำนวณสัมประสิทธิ์สหสัมพันธ์กับแบบทดสอบฉบับใด ค่าคะแนนของแบบทดสอบนั้นสอดคล้องสัมพันธ์กับเรื่องใดบ้าง กล่าวอีกอย่างหนึ่งก็คือพิสูจน์ได้หรือไม่ว่า ค่าคะแนนของแบบทดสอบนี้จะไม่ได้รับผลกระทบจากความแปรปรวน (variance) ของสิ่งไม่พึงประสงค์ ลักษณะการวัดของแบบทดสอบฉบับนั้นมีความคล้ายคลึงกันภายในกลุ่มหรือไม่ (homogeneity) ผู้สร้างได้กระทำการอันใดบ้างในการสร้างความคล้ายคลึงกันภายในกลุ่มขึ้น ข้อมูลต่างๆ ที่ผู้สร้างอ้างอิงนั้นมีลักษณะตรงตามที่อ้างไว้หรือไม่ สรุปก็คือว่า แบบทดสอบฉบับนี้จะสามารถวัดสิ่งที่ผู้เลือกประสงค์จะวัดได้หรือไม่

สำหรับเกณฑ์การวัดของแบบทดสอบนั้น จะต้องพิจารณาว่า ประชากรที่ผู้สร้างกำหนดเป็นมาตรฐานนั้น มีลักษณะสอดคล้องกับประชากรและสภาพแวดล้อมของหน่วยงานที่จะใช้วัดหรือไม่ และเกณฑ์การวัดที่กำหนดนั้นเหมาะสมกับลักษณะการจำแนกตำแหน่งภายในหน่วยงานที่จะใช้แบบทดสอบเพียงใด ในการนี้ผู้เลือกแบบทดสอบจะต้องตั้งปัญหาหลายประการ เช่น ตารางเกณฑ์การวัดของแบบทดสอบฉบับนั้นแสดงลักษณะของประชากรประเภทต่าง ๆ ไว้ชัดเจนหรือไม่ และการอ้างประชากรแต่

ละประเภทนั้น ได้สุ่มตัวอย่างไว้เป็นปริมาณเพียงพอหรือไม่ ในกรณีที่เห็นประชากรคนละประเภท ได้กำหนดเกณฑ์การวัดให้ต่างออกไปหรือเปล่า ลักษณะกว้าง ๆ ของประชากรแต่ละประเภทเหล่านี้เทียบได้กับประชากรของหน่วยงานที่จะใช้แบบทดสอบนี้หรือไม่ เกณฑ์การวัดนี้ได้คำนวณเป็นเกณฑ์ปกติ (normalized) แล้วหรือยัง พิสัย (range) ของค่าคะแนนของแบบทดสอบฉบับนี้สูงหรือต่ำ การแจกแจงค่าคะแนนเป็นไปในลักษณะปกติ (normal) เพียงใด ค่าผิดพลาดมาตรฐานของการวัด (standard error of measurement) มีปริมาณเท่าใด ปริมาณของค่าผิดพลาดมาตรฐานนี้เท่ากันหรือไม่ ระหว่างความสามารถหรือลักษณะนิสัยที่วัดได้ในระดับต่าง ๆ

ในเรื่องของลักษณะความเชื่อมั่นที่ได้มีอธิบายโดยละเอียดแล้วในบทที่ 4 แต่ก็อาจกล่าวไว้ในที่นี้เฉพาะในประเด็นที่ผู้เลือกใช้แบบทดสอบพึงถามตนเองอีกได้ คือ พิสัยของค่าความเชื่อมั่นปรากฏว่าเป็นเท่าใด ในการคำนวณค่านี้ได้อาศัยสุ่มตัวอย่างที่มีปริมาณตามสมควรหรือไม่ และสุ่มตัวอย่างนี้เป็นตัวแทนของประชากรที่คล้ายคลึงกับของหน่วยงานที่จะใช้แบบทดสอบเพียงใด ค่าความเชื่อมั่นนี้คำนวณได้โดยครอบคลุมระยะเวลาอันนานเท่าใด ระยะเวลาที่ใช้มีนานเกินไปจนความสามารถหรือบุคลิกภาพของผู้ถูกทดลองอาจเปลี่ยนไปจากสภาพการณ์ครั้งแรกที่เริ่มสร้างแบบทดสอบหรือไม่ สุ่มตัวอย่างที่นำมาใช้นั้นได้รับการคัดเลือกเป็นอย่างดีเพียงใด และผู้สร้างมีอคติ (bias) หรือไม่ในการคัดเลือกนี้ การที่ค่าความเชื่อมั่นต่างกันระหว่างกลุ่มตัวอย่าง เป็นผลของการที่ประชากรของกลุ่มสุ่มตัวอย่าง มีลักษณะต่างกันหรือเปล่า วิธีการใช้แบบทดสอบที่ผู้สร้างได้อธิบายไว้ นั้นชัดเจนพอที่จะไม่ทำให้ผลของการใช้แบบทดสอบนี้มีความเชื่อมั่นได้น้อยกว่าที่ผู้สร้างคำนวณไว้หรือเปล่า และถ้าแบบทดสอบฉบับนี้มี 2 ส่วน แต่ละส่วนมีค่าความเชื่อมั่นเท่าเทียมกันหรือไม่

ส่วนปัญหาที่น่าสงสัยเกี่ยวกับความเที่ยงตรงควรจะครอบคลุมประเด็นต่อไปนี้ คือ ผู้สร้างได้อธิบายแยกความเที่ยงตรงที่คำนวณได้ออกเป็น ประเภทความเที่ยงตรงในการคาดคะเน (predictive validity) และประเภทความเที่ยงตรงกับการวัดอื่น ๆ (concurrent validity) หรือไม่ ในการคำนวณค่าความเที่ยงตรงนี้ ได้อาศัยสุ่มตัวอย่างเป็นปริมาณเท่าใด ได้มีการวิจัยค่าความเที่ยงตรงของแบบทดสอบนี้ในระยะต่อ ๆ มาอีกหรือเปล่า และคู่มือประกอบแบบทดสอบได้นำผลการวิจัยเหล่านี้มาวิเคราะห์เปรียบเทียบไว้ด้วยหรือไม่ ถ้ามีผลการวิจัยที่ทดลองแล้วว่า แบบทดสอบนี้ขาดความเที่ยงตรง ผู้สร้างได้นำมาพิจารณาไว้ในคู่มือใช้แบบทดสอบหรือไม่ ค่าความเที่ยงตรงที่ผู้สร้างเป็นผลการคำนวณเทียบกับมาตรวัดความสำเร็จของงานอันใด (criterion measure) แบบทดสอบมี

โอกาสขาดความเที่ยงตรงเมื่อเทียบกับมาตรฐานความสำเร็จของงานอื่น ๆ หรือไม่ ถ้ามี เป็นมาตรฐานประเภทใด แบบทดสอบนี้มีความเที่ยงตรงสำหรับประชากรประเภทใด และใช้กับประชากรประเภทใดบ้างที่อาจจะทำให้ขาดความเที่ยงตรง จากข้อมูลที่คำนวณ ค่าความเที่ยงตรง อาจทำให้เชื่อได้หรือไม่ว่า ถ้านำแบบทดสอบนี้มาใช้กับประชากร ของหน่วยงานที่จะซื้อแบบทดสอบ ยังจะปรากฏคุณค่าของความเที่ยงตรงในการคาด คະเนอยู่ ในกรณีที่แบบทดสอบแบ่งออกเป็นภาค ๆ แต่ละภาคมีความเที่ยงตรงเทียบ ได้กับตัวแปรประเภทเดียวกันหรือไม่ และในขณะเดียวกัน แต่ละภาคก็เป็นอิสระ แก่กันหรือไม่

ระยะเวลาในการตอบแบบทดสอบและความยากง่ายในการคำนวณผล เป็นองค์ประ- กอบที่สำคัญอีกอย่างหนึ่งในการพิจารณาว่า ควรจะเลือกซื้อแบบทดสอบแบบมาตรฐาน ฉบับหนึ่งมาใช้หรือไม่ ในเรื่องนี้จะต้องศึกษาว่า ผู้สร้างมีข้อแม้อย่างไรในการใช้แบบ ทดสอบ เพื่อป้องกันไม่ให้ความเที่ยงตรงของแบบทดสอบด้อยลงไป ข้อบังคับในเรื่อง ระยะเวลาและในการควบคุมการสอบ มีรายละเอียดเป็นที่น่าพอใจหรือไม่ วิธีการคำนวณผล การวัดยุ่งยากสลับซับซ้อนเกินไปหรือไม่ รายละเอียดในเรื่องเหล่านี้จะเป็นเครื่องช่วย ในการพิจารณาว่าแบบทดสอบนี้ใช้ได้สะดวกหรือไม่เพียงใด

การควบคุมการทดสอบ โดยปกติแบบทดสอบมาตรฐานทุกฉบับจะมีคำอธิบายวิธี การใช้ไว้ชัดเจน การควบคุมการทดสอบโดยใช้แบบทดสอบนั้น ๆ จึงไม่น่าจะมีปัญหา อะไร และเมื่อพิจารณาโดยธรรมดาแล้ว การควบคุมผู้ถูกวัดให้ตอบแบบทดสอบ มิใช่ เป็นเรื่องยากเย็นอันใด แต่แท้ที่จริงการควบคุมการทดสอบมีความสำคัญเป็นอันมาก ผู้ควบคุมจะต้องปฏิบัติให้ผู้ถูกวัด มิฉะนั้นแบบทดสอบที่ได้คำนวณแล้วว่าจะมีความเชื่อมั่น ได้ อาจด้อยคุณภาพลงไป ผลของการวัดกลับไม่อยู่ในระดับความเชื่อมั่นที่ควรเชื่อถือ ได้ ทำให้การวินิจฉัยสั่งการในการบริหารงานบุคคลผิดพลาดไป เช่น บุคคลที่ไม่สมควร ได้รับการบรรจุกลับได้รับการแต่งตั้งให้ทำงาน บุคคลที่สมควรจะพิจารณาให้ทำงานใน หน้าที่หนึ่งกลับได้รับการพิจารณาให้ไปปฏิบัติงานอีกหน้าที่หนึ่ง ดังนี้ เป็นต้น

หลักใหญ่ของการควบคุมการทดสอบ คือ การจัดให้ผู้ถูกวัดอยู่ในสภาพที่จะแสดง ความสามารถและสติปัญญาของผู้นั้นได้อย่างเต็มที่ และผู้ถูกวัดจะต้องได้รับการกระตุ้น ให้แสดงบุคลิกภาพที่แท้จริงของตน การที่ผู้บริหารหน่วยวัดผลจะจัดให้เป็นเช่นนั้น ก็จะต้องพิจารณาดำเนินการ 5 ประการดังต่อไปนี้ คือ ⁴

⁴ Gekoski, Norman, *Psychological Testing*, pp. 227—229.
Guion, *op. cit.*, pp. 501—503.

ประการแรก จะต้องจัดให้สภาพของการสอบปราศจากสิ่งที่รบกวนสมาธิของผู้เข้าสอบ แบบทดสอบไม่ว่าฉบับใด แม้ว่าจะได้วิจัยแล้วว่ามีความสมบัติเป็นแบบทดสอบมาตรฐานที่ดี ถ้าหากนำมาใช้ในสภาพการณ์ที่ผู้ถูกวัดขาดสมาธิ ย่อมทำให้หย่อนคุณสมบัติของความเป็นมาตรฐานนี้ สภาพของการสอบที่ดี คือ สภาพการณ์ที่ผู้เข้าสอบแต่ละคนมีคู่มือสำหรับนั่งทำข้อสอบลำพังคนเดียว โดยไม่มีเสียงหรือกิจกรรมอื่นใดมารบกวน

ประการที่สอง ในการจัดสถานที่สอบ ต้องคำนึงถึงความสะดวกสบายของผู้เข้าสอบ ความสะดวกสบายนี้หมายถึงว่า ผู้เข้าสอบแต่ละคนมีโต๊ะและช่วงว่างสำหรับเขียนหนังสือได้เพียงพอ การที่ให้ผู้เข้าสอบเขียนคำตอบบนพนักรองติดเก้าอี้เช่นที่ได้จัดทำในการสอบของสถานที่หลายแห่งนั้น ไม่จัดว่าเป็นความสบายของผู้เข้าสอบ นอกจากนี้ ยังต้องจัดให้มีระบบแสง และการถ่ายเทอากาศเพื่อให้ผู้เข้าสอบไม่ต้องเกิดความอึดอัด อันอาจทำให้ผู้เข้าสอบแต่ละคนไม่ได้แสดงความสามารถเต็มที่ในการตอบแบบทดสอบ

ประการที่สาม วัสดุและอุปกรณ์ต่าง ๆ ที่จำเป็นต้องใช้ในการทดสอบต้องมีการเตรียมการไว้ให้พร้อมมูล เพื่อให้การดำเนินการในห้องสอบไม่หยุดชะงักขาดตอนไป อันอาจเป็นผลเสียหายในแง่ของค่าความเชื่อมั่นของแบบทดสอบนั้น ๆ

ประการที่สี่ ผู้ควบคุมการทดสอบมีหน้าที่กระตุ้นให้ผู้เข้าสอบใช้ความสามารถในการตอบแบบทดสอบอย่างเต็มที่ จริงอยู่ผู้เข้าสอบเองก็ประสงค์จะตอบให้ดีที่สุดอยู่แล้ว แต่ในการสอบนี้ ผู้เข้าสอบโดยทั่ว ๆ ไปย่อมมีความกังวลไม่น้อย ความกังวลนี้ทำให้ผู้เข้าสอบอาจแสดงความรู้ความสามารถไม่เต็มที่ ถ้าหากผู้ควบคุมจะได้อธิบายแนวความสำคัญที่ต้องใช้แบบทดสอบนั้น ๆ ก็อาจจะช่วยผ่อนคลายความรู้สึกเครียดของผู้เข้าสอบลงไปได้ แต่ลักษณะท่าทางของผู้ควบคุมที่แสดงว่ามีความเข้มงวดในเรื่องการจับเวลาเป็นอย่างยิ่ง อาจยังทำให้ผู้เข้าสอบเกิดความกังวลเพิ่มขึ้นอีกก็ได้

ประการสุดท้าย คือ การจัดลำดับของแบบทดสอบที่จะใช้ มีหลักกว้าง ๆ ว่า แบบทดสอบฉบับแรกควรจะง่ายและสั้น เพื่อเป็นการ “อุ่นเครื่อง” ผู้เข้าสอบ อย่างไรก็ตามถึงจะเชื่อกันว่าควรปฏิบัติเช่นนั้น ก็ไม่มีหลักฐานที่พิสูจน์ให้เห็นจริงได้ว่า การจัดเช่นนั้นเป็นประโยชน์ในการจัดการทดสอบ นอกจากนี้ยังยอมรับกันว่า หากมีแบบทดสอบทั้งประเภทวัดสติปัญญาและบุคลิกภาพ ควรให้ตอบแบบทดสอบที่วัดสติปัญญาเสียก่อน ทั้งนี้เพราะว่าถ้าผู้เข้าสอบเกิดเหนื่อยจากการตอบแบบทดสอบ ความเหนื่อยนั้นจะมีผลกระทบต่อคำตอบแบบทดสอบประเภทวัดบุคลิกภาพน้อยกว่าประเภทวัดสติปัญญา

ในกรณีที่แบบทดสอบมีลักษณะเป็นเครื่องมือวัดที่อาศัยอุปกรณ์พิเศษ ดังอธิบายไว้ในเรื่องการวัดความถนัดพิเศษ และลักษณะทางกายแล้วนั้น ผู้ควบคุมการทดสอบจะต้องจัดหลักผู้เข้าสอบให้ดี เพราะในการทดสอบเช่นนั้น ผู้เข้าสอบสามารถตอบข้อสอบได้ที่ละคน หรือที่ละเป็นจำนวนน้อย ทำให้ต้องมีการรอคอยกัน อีกประการหนึ่ง ผู้ควบคุมต้องตรวจตราความสึกหรอของอุปกรณ์ที่ใช้สอบนี้เสมอ เพราะการที่อุปกรณ์สึกหรอลงไปในนั้น อาจทำให้ผู้เข้าสอบคนหลัง ๆ สามารถตอบข้อสอบได้ง่ายขึ้น เช่น ในกรณีของการวัดทักษะของนิ้วมือ หมุด (pin) ที่ใช้กับเครื่องมือวัดอาจสึกไป ทำให้สามารถใส่ในช่องบนแผ่นไม้สะดวกขึ้น เป็นต้น ความสึกหรอของอุปกรณ์เหล่านี้ทำให้คุณภาพความเป็นมาตรฐานของเครื่องมือวัดด้อยลงไป

ระเบียบการวัดผลและรายงานผล การนำคำตอบมาคำนวณผลของการทดสอบนั้นไม่มีความสลับซับซ้อนอะไร เพราะแบบทดสอบมาตรฐานทุกฉบับย่อมมีอธิบายการคิดค่าคะแนนของคำตอบไว้ชัดเจนแล้ว อย่างไรก็ตามในเรื่องนี้ก็มีสิ่งพึงระลึก 2 ประการคือ หัวหน้าหน่วยวัดผลจะต้องควบคุมการคิดค่าคะแนนนี้ให้เป็นไปตามคู่มือการใช้แบบทดสอบนั้น ๆ อย่างเคร่งครัดประการหนึ่ง และอีกประการหนึ่ง ควรคำนวณหาค่าคะแนนนี้ให้เสร็จโดยรวดเร็วไม่ผิดพลาด ในเรื่องนี้มีหลักว่า ควรจะคิดค่าคะแนนของผู้เข้าสอบทุกคนเป็นฉบับ ๆ ไป ไม่ใช่คิดค่าคะแนนของผู้เข้าสอบแต่ละคนในการตอบแบบทดสอบทุกฉบับ

ค่าคะแนนที่คำนวณได้นี้ควรนำมาลงหลักฐานทำเป็นระเบียบขึ้น เพื่อให้สามารถติดตามผลของการวัดของแต่ละคนได้โดยสะดวก ระเบียบการวัดผลควรต้องมีข้อเท็จจริงดังต่อไปนี้เป็นอย่างน้อย⁵

1. ชื่อของแบบทดสอบ
2. ฟอรม์ของแบบทดสอบ (ในกรณีที่แบบทดสอบฉบับหนึ่ง ๆ มีหลายฟอรม์)
3. ค่าคะแนนดิบ
4. ค่าคะแนนมาตรฐาน เช่น ค่าเปอร์เซ็นต์ไต้
5. กลุ่มประชากรที่ใช้คำนวณเกณฑ์การวัดของแบบทดสอบฉบับนี้
6. วันที่ทำการวัดผล

⁵ Gekoski, *op. cit.*, p. 231.

การประมวลผลการวัดเพื่อรักษาเป็นระเบียบนี้อาจทำเป็นตารางหรือกราฟก็ได้ ดังแสดงไว้ในแผนภาพที่ 21 และ 22 ตามลำดับ

หลักฐานที่เก็บไว้เป็นระเบียบนี้ มิใช่จะต้องนำไปรายงานทั้งหมด ที่ให้เก็บหลักฐานดังกล่าวไว้ ก็เพื่อประโยชน์ในการวิเคราะห์เปรียบเทียบในระยะต่อไป และยังสามารถใช้

ชื่อ : นาย ไทย รักชาติ		วันที่ 13 มีนาคม 2514	
ชื่อแบบทดสอบ	คะแนนดิบ	เปอร์เซ็นต์ไต่	ประชากรที่เป็นเกณฑ์การวัด
Otis Employment, Higher A	69	91	ผู้ใหญ่
Minnesota Clerical Test			ผู้ใหญ่
Name Checking	104	78	
Number Checking	121	90	
Kuder Preference Record Form CH			ผู้ใหญ่
Outdoor	17	5	
Mechanical	34	22	
Computational	43	92	
Scientific	41	52	
Persuasive	56	85	
Artistic	34	89	
Literary	18	45	
Musical	8	28	
Social Service	45	58	
Clerical	64	65	
Gordon Personal Profile			นักศึกษามหาวิทยาลัย
Ascendancy	7	66	เพศชาย
Responsibility	16	99	
Emotional Stability	8	64	
Sociability	5	53	

แบบทดสอบ	คะแนน	เปอร์เซ็นต์ ไต่ล	ประชากรที่เป็น เกณฑ์การวัด	5 10 20 30 40 50 60 70 80 90 95										
													
Otis Employment, Higher A	69	91	ผู้ใหญ่											
Minnesota Clerical			ผู้ใหญ่											
Name Checking	104	78												
Number Checking	121	90												
Kuder Preference Record, CH			ผู้ใหญ่											
Outdoor	17	5												
Mechanical	34	22												
Computational	43	92												
Scientific	41	52												
Persuasive	56	85												
Artistic	34	89												
Literary	18	45												
Musical	8	28												
Social Service	45	58												
Clerical	64	65												
Gordon Personal Profile			นักศึกษา											
Ascendancy	7	66												
Responsibility	16	99	มหาวิทยาลัย											
Emotional Stability	8	64	ชาย											
Sociability	5	53												

แผนภาพที่ 22 ผลการวัดบุคคลแสดงเป็นกราฟ

ประกอบในการวิจัยที่อาจจัดให้มีขึ้นในโอกาสต่าง ๆ ตามธรรมดาหน่วยวัดผลไม่ใช่เป็นหน่วยงานที่จะเป็นผู้ตัดสินสั่งการขั้นสุดท้ายในเรื่องการบริหารงานบุคคล ฉะนั้น หน่วยวัดผลจึงมีหน้าที่ต้องรายงานผลการวัดนี้ต่อผู้มีอำนาจในเรื่องนี้ ข้อเท็จจริงที่จะรายงานขึ้นไป ควรเป็นตัวเลขที่มีความหมายช่วยให้ผู้มีอำนาจตัดสินใจในเรื่องนี้ได้ ผลการวัดที่จะรายงานนี้ จึงควรจะเป็นค่าคะแนนมาตรฐาน อันจะทำให้ผู้บริหารชั้นเหนือขึ้นไปสามารถเปรียบเทียบคุณสมบัติของบุคคลต่าง ๆ ที่ถูกวัดได้โดยสะดวก คะแนนมาตรฐานที่มักนิยมใช้เพื่อการนี้ มักจะเป็นค่าเปอร์เซ็นต์ได้ เพราะเป็นค่าที่เข้าใจง่าย และเป็นค่าที่ใช้เทียบกับเกณฑ์มาตรฐานของแบบทดสอบต่าง ๆ ได้ อย่างไรก็ดี การที่จะให้รายงานผลการวัดนี้มีความหมายสมบูรณ์ ควรมีตัวเลขนอกจากค่าเปอร์เซ็นต์เพิ่มเติมอีก คือ ค่าตัวเลขที่แสดงความน่าจะเป็นของโอกาสสำเร็จในการปฏิบัติงาน ค่าเปอร์เซ็นต์ได้สามารถอธิบายได้แต่เพียงคุณภาพของผู้ถูกวัดคนหนึ่ง ๆ เมื่อเทียบกับคนอื่น ๆ ยังไม่มีความหมายในการคาดคะเนความสำเร็จของงาน ดังนั้นในรายงานผลการวัดถ้าหากมีการอธิบายอีกด้วยว่า บุคคลหนึ่ง ๆ นั้นมีโอกาส 75 ใน 100 ที่จะปฏิบัติงานในขั้นดีมาก หรือมีโอกาส 50 ใน 100 ที่จะปฏิบัติงานพอใช้ได้ ผู้บริหารชั้นเหนือขึ้นไปย่อมได้รับประโยชน์ในการพิจารณาตัดสินใจยิ่งขึ้น

การใช้ผลการวัดในงานบุคคล

ค่าคะแนนที่ได้โดยเครื่องมือวัดที่ถูกต้องตามหลักเกณฑ์ อาจใช้ประโยชน์ในการบริหารงานบุคคลได้หลายทาง คือ ในการคัดเลือก หรือ เลือกรับ (selection and placement) ในการวิเคราะห์ความจำเป็นที่จะต้องมีการฝึกอบรมและความสำเร็จผลของการฝึกอบรม และในการพิจารณาผู้ที่มีคุณสมบัติความสามารถเหมาะสมที่จะเลื่อนชั้นเลื่อนตำแหน่ง นอกจากนี้ยังอาจใช้วิเคราะห์ปัญหาขององค์การได้อีก เช่น วิจัยความถูกต้องของคำอธิบายงาน (job description) โดยเทียบค่าความสัมพันธ์ระหว่างความสำเร็จกับความไม่สามารถในการปฏิบัติงานในตำแหน่งงานหนึ่ง ๆ การเทียบค่าความสัมพันธ์นี้ทำได้โดยการวิเคราะห์ค่าสมการถดถอย (regression) ของตัวแปรการวัดผล (predictor variable) กับมาตรวัดความสำเร็จของงาน (criterion measure) อย่างไรก็ตาม ในบทนี้ จะจำกัดการพิจารณาไว้เฉพาะแต่ในเรื่องการใช้ผลของการวัดในงานบุคคลเท่านั้น

ความสัมพันธ์ระหว่างข้อวินิจฉัยสั่งการกับการใช้เครื่องมือวัด

ข้อวินิจฉัยสั่งการในเรื่องงานบุคคล (personnel decisions) นานัปการ เมื่อพิจารณาในแง่การวัดผลบุคคล อาจยื่นข้อให้เหลือแต่เพียงประเภทเดียว คือ จะให้บุคคลหนึ่ง ๆ ปฏิบัติงาน หรือไม่ให้ปฏิบัติงาน หรือให้รอคอยเพื่ออาจให้ปฏิบัติงานในเวลาต่อไป เมื่อพิจารณาตามแนวนี้ การใช้เครื่องมือวัดที่ถูกต้องตามหลักเกณฑ์ย่อมมีประโยชน์เป็นอันมากในการช่วยผู้บริหารให้สามารถตัดสินใจในเรื่องนี้ได้ถูกต้องเที่ยงธรรม และให้ได้บุคคลที่เหมาะสมจริง ๆ ปฏิบัติงานในตำแหน่งหนึ่ง ๆ แต่เครื่องมือวัดก็มีไม่ใช่เป็นตัวแทน

¹ Cronbach, Lee J., and Gleser, Goldine C., *Psychological Tests and Personnel Decisions*, pp. 15–17.

ของผู้บริหารในการวินิจฉัยสั่งการเรื่องนี้ เพราะเมื่อพิจารณาลักษณะและวิธีการของการสร้างเครื่องมือวัดดังกล่าวมาแล้วทั้งหมด ก็จะเห็นว่า เครื่องมือวัดบุคคลมิใช่ของวิเศษที่สามารถตัดสินในเรื่องนี้อย่างได้ผลเป็นที่น่าพอใจ ถึงแม้ว่าเครื่องมือวัดที่ได้พยายามสร้างขึ้นให้ถูกต้องตามหลักเกณฑ์จะมีคุณสมบัติในเรื่องความเที่ยงตรง (validity) และความเชื่อมั่น (reliability) ในระดับเป็นที่น่าพอใจก็ตาม ค่าคะแนนที่วัดได้โดยอาศัยเครื่องมือวัดที่ถูกต้องตามหลักเกณฑ์เหล่านี้ เป็นเพียงข้อสนับสนุนในเชิงปริมาณ (quantitative data) อันอาจใช้ประกอบการวินิจฉัยสั่งการเท่านั้น

เพื่อให้สะดวกแก่การพิจารณาว่า จะใช้ผลการวัดได้อย่างไรในการจัดงานบุคคล จึงขอแยกข้อวินิจฉัยสั่งการในเรื่องนี้ออกเป็น 3 ประเภท ตามแนวของ Thorndike² คือ การรับหรือไม่รับเข้าทำงาน (selection-rejection) การไม่รับเข้าทำงานสำหรับตำแหน่งงานหลายตำแหน่ง (multiple selection-rejection) และการจัดให้ปฏิบัติงานในตำแหน่งที่เหมาะสม (classification)

การรับหรือไม่รับเข้าทำงาน ข้อวินิจฉัยสั่งการที่จัดเข้าอยู่ในประเภทนี้ คือข้อวินิจฉัยสั่งการที่มีลักษณะทำนองว่า จะให้บุคคลหนึ่ง ๆ เข้าปฏิบัติงานหรือไม่ให้เข้าปฏิบัติงานในตำแหน่งหนึ่ง ๆ ปัญหาที่เกี่ยวข้องในการที่ต้องมีการตัดสินใจในเรื่องนี้ คือ การที่มีตำแหน่งหนึ่ง ๆ ว่าง จำเป็นต้องมีผู้ปฏิบัติงาน และอาจเลือกผู้ปฏิบัติงานในตำแหน่งนั้น ๆ ได้จากบุคคลเป็นจำนวนหลาย ๆ คน

ในกรณีที่ข้อวินิจฉัยสั่งการมีลักษณะเป็นเช่นนี้ ผู้บริหารมีความจำเป็นแต่เพียงที่จะเลือกบุคคลที่มีลักษณะและคุณสมบัติตามที่ต้องการไว้เท่านั้น เครื่องมือวัดบุคคลที่ใช้ก็มุ่งในแง่ที่จะคัดบุคคลที่ไม่เข้าข่ายออกจากบุคคลที่จะสามารถปฏิบัติงานได้ ดังนั้นในการวัดสติปัญญาที่ดี หรือความสัมฤทธิ์ผลที่ดี หรือในการวัดอื่นใดที่ดี เมื่อจะใช้ประกอบการวินิจฉัยสั่งการในลักษณะเช่นนี้ ก็ไม่มีความจำเป็นที่จะต้องแสดงความน่าจะเป็นของความสำเร็จในการปฏิบัติงานของค่าคะแนนแต่ละตัว หากเครื่องมือวัดสามารถกำหนดค่าคะแนนขั้นต่ำที่จำเป็นจะต้องได้ในการปฏิบัติงานในตำแหน่งหนึ่ง ๆ ก็นับว่าเพียงพอแล้วที่ผู้บริหารจะพิจารณาตัดสินใจ จากคะแนนสูงต่ำที่บุคคลแต่ละคนได้เรียงลำดับกัน

การไม่รับเข้าทำงานสำหรับงานหลายตำแหน่ง สภาพการณ์นี้หมายถึงว่า บุคคลแต่ละคนได้รับการพิจารณาความเหมาะสมในการปฏิบัติงานสำหรับตำแหน่งงานหลาย

² Thorndike, Robert L., *Personnel Selection*, pp. 294—296.

ตำแหน่งพร้อม ๆ กันไปที่เดียว ผู้บริหารต้องการทราบว่า บุคคลแต่ละคนเหล่านั้นไม่มีคุณสมบัติพอที่จะปฏิบัติงานในตำแหน่งใดบ้าง เพื่อให้สามารถตัดสินใจได้ว่า ตำแหน่งงานที่ว่างหลาย ๆ ตำแหน่ง มีผู้ที่มีคุณสมบัติและความสามารถสมควรปฏิบัติงานได้เป็นจำนวนพอเพียงทุก ๆ ตำแหน่งหรือไม่ การที่เกิบบัญหาทำให้ต้องพิจารณาเช่นนี้ก็เพราะว่า ตำแหน่งที่ว่างหลาย ๆ ตำแหน่งนั้นต้องการบุคคลที่มีคุณสมบัติและความสามารถคล้าย ๆ กัน การที่จะเลือกผู้ใดให้ปฏิบัติในตำแหน่งใด จึงต้องทำให้แน่ใจว่า ผู้นั้นขาดคุณสมบัติที่จะปฏิบัติงานในตำแหน่งอื่น ๆ

ในกรณีเช่นนี้ เครื่องมือนัดจะต้องสามารถแสดงรายละเอียดมากกว่าในกรณีของการรับหรือไม่รับเข้าทำงาน กล่าวคือ ต้องสามารถแสดงได้ว่า ค่าคะแนนแต่ละค่าที่วัดได้ มีความน่าจะเป็นของความสำเร็จในการปฏิบัติงานตำแหน่งหนึ่ง ๆ สูงต่ำเพียงไร ความน่าจะเป็นในเรื่องนี้ คือ ดัชนีในเชิงปริมาณของการปฏิบัติงานได้ (quantitative indicator of success) ค่าคะแนนแต่ละค่าที่ผัดกันไปย่อมมีความน่าจะเป็นของความสำเร็จในการปฏิบัติงานเป็นปริมาณต่างกัน แม้จะมีข้อมูลเป็นรายละเอียดในเชิงปริมาณ เช่นนี้ ผู้บริหารก็ยังตัดสินใจในเรื่องนี้ได้ลำบากกว่าการวินิจฉัยสั่งการในกรณีแรก ดังจะกล่าวต่อไป

การจัดให้ปฏิบัติงานในตำแหน่งที่เหมาะสม ในบางโอกาสมีปัญหาก็จะต้องแต่งตั้งให้ผู้ที่รับเข้าปฏิบัติงานไปดำรงตำแหน่งที่เหมาะสม ในการนี้ผู้บริหารจำเป็นต้องมีผลการวัดบุคคลเหล่านั้นทุกแง่ทุกมุม ผลการวัดบุคลิกภาพหรือความสัมพันธ์ผลจะต้องสามารถแสดงได้ว่าบุคคลหนึ่ง ๆ เหมาะสมกับงานในตำแหน่งหนึ่ง ๆ มากกว่าบุคคลอื่น ๆ อย่างไร หรือกล่าวอีกอย่างหนึ่งก็คือว่า ทุกคนที่จะปฏิบัติงานนั้นจะเหมาะสมที่สุดที่จะปฏิบัติงานในตำแหน่งใด การที่จะพิจารณาเช่นนี้ได้จำเป็นต้องมีการเทียบกันระหว่างบุคคลทุกคน เครื่องมือนัดที่ใช้ก็จำเป็นต้องสามารถแสดงดัชนีในเชิงปริมาณของการปฏิบัติงานได้ เช่นเดียวกับในกรณีที่สอง

การใช้เครื่องมือวัดเพื่อให้มีดัชนีในเชิงปริมาณของการปฏิบัติงานได้ มีลักษณะที่ทำให้ผู้บริหารต้องประสบความยุ่งยากในการวินิจฉัยสั่งการ เครื่องมือที่ใช้ในสภาพการณ์เช่นนี้ มีคุณสมบัติเป็นตัวแปรการวัดผลที่ต่อเนื่อง (continuous quantitative variable) ค่าคะแนนทุกค่าแสดงโอกาสที่จะปฏิบัติงานได้ด้วยกันทั้งสิ้น ต่างกันแต่ปริมาณของความน่าจะเป็นในเรื่องนี้เท่านั้น ถ้าหากผู้บริหารตัดสินใจแต่เพียงว่า จะให้บุคคลคนหนึ่ง ๆ ปฏิบัติงานหรือไม่ให้ปฏิบัติงานก็ย่อมเป็นการง่าย ในเมื่อมีการกำหนดค่า

คะแนนขั้นต่ำที่ควรจะต้องได้ไว้แล้ว ผู้ที่ได้ต่ำกว่าเกณฑ์ก็ไม่ได้รับการพิจารณา ส่วนผู้ที่ได้รับการพิจารณาก็อาจพิจารณาเป็นลำดับลดหลั่นได้โดยสะดวก แต่การพิจารณาเครื่องมือวัด ในแง่ตัวแปรการวัดผลที่ต่อเนื่อง ต้องประสบความสำเร็จอย่างถึง 2 ประการ คือ นอกจากจะต้องพิจารณาว่าบุคคลคนหนึ่ง ๆ ควรปฏิบัติงานตำแหน่งใดจึงจะสามารถใช้ศักยภาพ (potentialities) ของผู้นั้นได้เต็มที่ แล้วยังต้องวิเคราะห์อีกด้วยว่า เมื่อพิจารณาเทียบกับบุคคลอื่น ๆ แล้ว ยังมีบุคคลที่ดีกว่าบุคคลผู้นั้นอีกหรือไม่ ในการที่จะมอบหมายให้ผู้นั้นปฏิบัติงานในตำแหน่งนี้ โดยใช้ศักยภาพอย่างเต็มที่ แต่ความยุ่งยากในเรื่องนี้ ไม่ควรเป็นเหตุผลที่ทำให้ฝ่ายบริหารหรือผู้บริหารหันไปปฏิบัติแต่วิธีที่สะดวก คือ เพียงแต่ตัดสินใจว่าจะให้บุคคลหนึ่ง ๆ ปฏิบัติงาน หรือไม่ให้ปฏิบัติงานในตำแหน่งหนึ่ง ๆ เพราะการปฏิบัติงาน ๆ เช่นนี้ มิใช่เป็นการช่วยให้สามารถใช้ศักยภาพของบุคคลแต่ละคนได้อย่างเต็มที่

ตัวแปรการวัดผลที่ไม่ใช่เครื่องมือวัด ในแง่ของการวิจัย องค์ประกอบต่าง ๆ ที่เกี่ยวข้องกับบุคคลที่จะปฏิบัติงานอาจเป็นตัวแปรการวัดผลทั้งสิ้น เพราะเป็นเรื่องที่เกี่ยวข้องกับการที่บุคคลหนึ่ง ๆ จะปฏิบัติงานได้ดีเพียงไร แต่ที่ได้ย้ำเรื่องตัวแปรการวัดผลในแง่ของเครื่องมือวัดโดยตลอดมา ก็เพราะว่าเครื่องมือวัดเป็นสิ่งที่ได้วิเคราะห์ความเที่ยงตรงและความเชื่อมั่นมาแล้ว ในการนี้ได้มีการวัดความแปรปรวนของผลการใช้เครื่องมือวัดตามหลักคณิตศาสตร์โดยสมบูรณ์ ส่วนองค์ประกอบอื่น ๆ ที่เกี่ยวกับตัวบุคคลที่จะปฏิบัติงาน เช่น อายุ เพศ สภาพการสมรส ตลอดจนประสบการณ์ต่าง ๆ เป็นสิ่งที่ยังไม่ได้รับการวิเคราะห์เหมือนอย่างเครื่องมือวัด ฉะนั้นจึงนำมาใช้ในการคาดคะเนความสำเร็จของการที่จะปฏิบัติงานที่แท้จริงไม่ได้ อย่างไรก็ตาม เครื่องมือวัดที่ได้พยายามสร้างขึ้น ตลอดจนเครื่องมือวัดที่อาจสร้างขึ้นได้ตามทฤษฎีที่มีอยู่ ก็ยังไม่สามารถครอบคลุมองค์ประกอบทุกตัวที่เกี่ยวข้องกับบุคคลที่จะปฏิบัติงาน ฉะนั้นองค์ประกอบอื่น ๆ ที่นอกเหนือจากเครื่องมือวัด จึงควรได้รับการพิจารณาประกอบไปกับเครื่องมือวัด เมื่อมีการวินิจฉัยสั่งการในเรื่องงานบุคคล

การนำข้อเท็จจริงอื่น ๆ อันไม่ใช่ผลการใช้เครื่องมือวัดมาพิจารณานั้น ควรทำให้อยู่ในขอบเขต เพราะข้อเท็จจริงไม่มีคุณสมบัติในการคาดคะเนเท่าเทียมกับเครื่องมือวัดดังกล่าวแล้ว การใช้ข้อเท็จจริงต่าง ๆ ประกอบเครื่องมือวัดนี้ อาจกระทำได้เป็น 3 วิธี คือ³ ใช้ข้อเท็จจริงเป็นเกณฑ์ขั้นต้นตัดสินโอกาสที่จะพิจารณาต่อไป ใช้ประกอบ

³ Ibid., p. 416.

การพิจารณาเป็นราย ๆ ในกรณีที่ผลของการวัดยังไม่ถึงระดับ และใช้ควบคุมไปกับผลการวัด

การใช้ข้อเท็จจริงที่เกี่ยวกับตัวบุคคลเป็นเกณฑ์ขั้นต้นในการพิจารณานั้นควรพยายามใช้ให้น้อยที่สุด เพราะข้อเท็จจริงที่เกี่ยวกับตัวบุคคล เช่น อายุ หรือสาขาวิชาของการศึกษามีใช่เป็นเครื่องตัดสินโดยเด็ดขาดว่าบุคคลจะมีหรือขาดความสามารถที่แท้จริงในการปฏิบัติงานหนึ่ง ๆ อย่างไรก็ดี การกำหนดว่าจะรับพิจารณาเฉพาะคนที่มีลักษณะเช่นนั้นเช่นนั้น เช่น กำหนดอายุระดับหนึ่ง ต้องเป็นเพศชายหรือเพศหญิง เป็นต้นมีส่วนช่วยให้ฝ่ายบริหารได้รับความสะดวกมากขึ้นในการพิจารณา แต่หากว่ามีการใช้องค์ประกอบต่าง ๆ หลายประการทำนองนี้เป็นเครื่องขัดกั้นในการพิจารณา อาจทำให้ไม่ได้บุคคลที่ดีที่สุดมาปฏิบัติงานในตำแหน่งที่ต้องการ

แทนที่จะใช้องค์ประกอบเกี่ยวกับตัวบุคคลอันไม่ใช่ผลการวัดเน้นหนักไปในลักษณะประการแรก อาจใช้ข้อเท็จจริงนี้ประกอบการพิจารณาผลของการวัดผลเป็นราย ๆ ไป เช่น กำหนดว่า ผู้ที่สำเร็จการศึกษาวิชาในสาขาหนึ่งอาจไม่ได้รับการพิจารณา ยกเว้นแต่ว่ามีความสามารถที่วัดได้ถึงระดับหนึ่ง วิธีนี้อาจทำให้มีโอกาสมากขึ้นในการเลือกสรรบุคคลที่มีความสามารถจริง ๆ แต่อย่างไรก็ตามก็เป็นการยากที่จะพิสูจน์ว่าข้อแม้อันที่กำหนดขึ้นสำหรับบุคคลประเภทหนึ่ง ๆ นั้น สมเหตุสมผลเป็นประโยชน์ต่อการปฏิบัติงานในตำแหน่งนั้น ๆ อย่างแท้จริง

วิธีการใช้ข้อเท็จจริงเกี่ยวกับตัวบุคคลประกอบในการวัดผลอีกวิธีหนึ่ง คือ นำข้อเท็จจริงเหล่านี้มาถือเป็นองค์ประกอบที่จะต้องพิจารณาควบคุมไปกับผลการวัด ในการพิจารณาตามวิธีนี้ ฝ่ายบริหารจะต้องสนับสนุนให้มีการวิจัยหาความสัมพันธ์ระหว่างข้อเท็จจริงแต่ละรายการกับโอกาสที่จะปฏิบัติงานในตำแหน่งหนึ่ง ๆ ได้สำเร็จ และเมื่อทราบความสัมพันธ์เช่นนั้นแล้วก็นำข้อเท็จจริงแต่ละรายการมาให้ค่าเป็นสัดส่วนกับความสัมพันธ์ที่มีอยู่ในเรื่องนี้ ส่วนข้อเท็จจริงอื่น ๆ ที่มีมีความเกี่ยวข้องกับโอกาสที่จะปฏิบัติงานได้สำเร็จ ก็ไม่นำมาพิจารณา ผู้ที่มีคุณสมบัติตามรายการที่กำหนดนี้ ก็จะได้รับค่าคะแนนลดหลั่นไปตามลำดับของปริมาณคุณสมบัติที่มีอยู่ ผลของการพิจารณาเช่นนั้น ก็สามารถนำไปประกอบการพิจารณาผลของการวัดได้ เช่น ถ้าทราบว่าประสพการณ์ในการปฏิบัติงานในระยะเวลาหนึ่ง จะทำให้มีโอกาสปฏิบัติงานในตำแหน่งหนึ่งได้ผลดี ในการพิจารณาผลการวัดก็จะนำระยะเวลาที่ต่างกันที่บุคคลแต่ละคนมีประสพการณ์ในการทำงานมาเทียบค่าคะแนน แล้วนำค่าคะแนนนี้ไปประเมิน

ร่วมกับค่าคะแนนที่บุคคลแต่ละคนเหล่านั้นได้จากเครื่องมือวัดหนึ่ง ๆ ส่วนข้อเท็จจริงอื่น ๆ เช่น อายุหรือเพศ ก็ไม่นำมาพิจารณาในเรื่องนี้ การจัดลำดับที่ ที่สมควรจะได้รับการพิจารณาของแต่ละบุคคลอาจยุ่งยากซับซ้อน เพราะผู้ที่ได้ค่าคะแนนสูงในการวัดผลอาจมีคุณสมบัติในเรื่ององค์ประกอบอื่น ๆ ต่ำ หรืออาจกลับกันในบรรดาผู้ที่ได้ค่าคะแนนต่ำในการวัดผลก็ได้ แต่ความยุ่งยากในเรื่องนี้เป็นเครื่องประกันว่าฝ่ายบริหารได้พิจารณาเหตุผลต่าง ๆ โดยรอบคอบแล้ว จึงได้มีการวินิจฉัยสั่งการในเรื่องงานบุคคล

เครื่องมือวัดในขั้นตอนต่าง ๆ ของงานบุคคล

ได้กล่าวไว้ในตอนต้นของบทนี้ว่า ผลของการใช้เครื่องมือวัดเป็นประโยชน์ต่อการจัดงานบุคคลในขั้นตอนต่าง ๆ นับตั้งแต่การคัดเลือกไปจนถึงการบรรจุ และการเลื่อนขั้นเลื่อนตำแหน่ง ในตอนนี้นำรายละเอียดในเรื่องนี้มาพิจารณาต่อไป

การคัดเลือกและการเลือกบรรจุ เมื่อหน่วยงานประสงค์จะรับบุคคลเข้าปฏิบัติงาน อาจจัดให้มีการคัดเลือก (selection) หรืออาจรับโดยเลือกบรรจุ (placement) ก็ได้ การคัดเลือกหมายถึงการเลือกบุคคลแต่ละคนให้ปฏิบัติหน้าที่ที่เหมาะสมที่สุดในบรรดาคำแหน่งงานที่มีอยู่ ในการคัดเลือกนั้นต้องเน้นความแตกต่างระหว่างบุคคล (interpersonal differences) เป็นประการสำคัญ ส่วนในการเลือกบรรจุนี้เน้นความแตกต่างในเรื่องคุณสมบัติและความสามารถต่าง ๆ ภายในตัวบุคคลแต่ละคนเอง (intrapersonal differences)

ในการคัดเลือกบุคคลโดยทั่วไป การวัดผลในเรื่องสติปัญญา ความสัมฤทธิ์ผล และความถนัดพิเศษหรือทักษะทางกายมีความสำคัญมาก เพราะบุคคลแต่ละคนมีคุณสมบัติในเรื่องเหล่านี้แตกต่างกันไป โดยเฉพาะในการคัดเลือกบุคคลในระดับผู้บริหารนั้น ระดับสติปัญญาที่มีความสำคัญมาก ดังจะกล่าวถึงในหัวข้อต่อไปอีกโดยเฉพาะ การวัดผลเพื่อคัดเลือกบุคคลเข้าปฏิบัติงานในระดับต่าง ๆ อาจเน้นหนักความสามารถในเรื่องเหล่านี้ต่างกันอย่างออกไป เช่น ในงานฝ่ายเลขานุการ อาจเน้นในเรื่องความสัมฤทธิ์ผล และมีการวัดความถนัดพิเศษและสติปัญญาประกอบไปด้วย แต่สำหรับงานในตำแหน่งช่างฝึกหัด หรือช่างผู้ช่วย อาจเน้นเรื่องความถนัดพิเศษและวัดความสัมฤทธิ์ผลเป็นการประกอบ การที่จะเน้นการวัดในเรื่องใดก็อาศัยหลักการวิเคราะห์ความเที่ยงตรงดังกล่าวไว้แล้วในบทที่ 5 เป็นแนวในการพิจารณา

ในระยะเวลาที่ผ่านมา ได้มีผู้สร้างเครื่องมือวัดบุคลิกภาพเพื่อใช้ประกอบในการคัดเลือกบุคคลเป็นจำนวนหลายฉบับ อย่างไรก็ตามก็ดี จากผลการวิจัยของ Kirkpatrick⁴ ในปี ค.ศ. 1956 ปรากฏว่า เครื่องมือวัดบุคลิกภาพต่างๆ ไม่มีคุณสมบัติในเชิงภาคคะเนความสำเร็จในการปฏิบัติงานในระดับเป็นที่น่าพอใจ ฉะนั้น เครื่องมือวัดที่สร้างเพื่อประโยชน์ในการคัดเลือกบุคคลเหล่านี้ จึงอาจเหมาะแก่การเลือกบรรจุมากกว่า

ในการเลือกบรรจุนี้ มีความหมายว่า ผู้บริหารยอมรับบุคคลหนึ่ง ๆ เข้าปฏิบัติงาน แต่ยังไม่ได้กำหนดว่าจะให้ปฏิบัติงานในตำแหน่งใด การที่จะทราบว่าบุคคลหนึ่ง ๆ เหมาะแก่หน้าที่ตำแหน่งอันใด จึงต้องพิจารณาความแตกต่างในเรื่องความสามารถต่างๆ ภายในตัวบุคคลนั่นเอง การที่กล่าวว่าบุคคลหนึ่ง ๆ มีความแตกต่างภายในตนเอง หมายความว่าลักษณะนิสัยและความสามารถอย่างหนึ่ง ๆ ของบุคคลคนเดียวกันมีระดับไม่เท่ากัน เช่น บุคคลที่มีความสามารถสูงในวิชาคณิตศาสตร์อาจดี้อยู่ความสามารถในด้านภาษา หรือบุคคลที่มีความคิดริเริ่มสูง อาจขาดความชำนาญงานบางอย่างก็ได้ การด้อยคุณสมบัติของลักษณะนิสัยและความสามารถอย่างหนึ่ง ๆ นี้ อาจด้อยทั้งในแง่เมื่อเทียบกับเกณฑ์ปกติของบุคคลทั่วไป หรือด้อยเฉพาะเมื่อเทียบระหว่างประเภทต่างๆ ของความสามารถภายในตนเอง แต่ไม่ด้อยกว่าเกณฑ์ปกติของบุคคลทั่วไปก็ได้ ดังจะสังเกตได้จากตัวอย่างในตารางที่ 30

ตารางที่ 30

ค่าสมมติของความสามารถของบุคคลสามคนเป็นเปอร์เซ็นต์ได้

ความสามารถ	คนที่หนึ่ง	คนที่สอง	คนที่สาม
ทักษะ ด้านภาษา	72	72	72
ทักษะ ด้านตัวเลข	1	39	97
ความชำนาญในงานธุรการ	68	11	25
ความสามารถในการหาเหตุผล	19	41	72
ความชำนาญในการพิมพ์ดีด	49	38	25

จากตารางที่ 30 จะเห็นว่าบุคคลที่หนึ่ง มีความสามารถในด้านสติปัญญาดีด้อยกว่าผู้อื่นอย่างเห็นได้ชัด แต่มีความสามารถในเรื่องที่เกี่ยวกับการปฏิบัติงานในสำนักงาน

⁴ Guion, Robert M., *Personnel Testing*, p. 318.

โดยตรง คือ มีความรู้ในเรื่องภาษา และชำนาญในด้านธุรการและพิมพ์ดีด ส่วนบุคคลที่สอง มีความสามารถในด้านสติปัญญาสูงกว่า แต่ถนัดในเรื่องงานที่เกี่ยวกับสำนักงานน้อยกว่า คนที่สามนั้นยังมีความสามารถในด้านสติปัญญาสูงขึ้นไปอีก และความสามารถในเรื่องอื่นๆ ก็สูงกว่าคนที่สองทั้งสิ้น ยกเว้นแต่ในเรื่องพิมพ์ดีด

จากตัวอย่างนี้ แสดงว่า ความสามารถในเรื่องการพิมพ์ดีดของบุคคลคนที่หนึ่ง ซึ่งมูลค่า 49 เปอร์เซนต์ อันเป็นระดับเกณฑ์เฉลี่ยของบุคคลทั่วไปกลับเป็นความสามารถที่เด่น ทั้งนี้ เพราะในการพิจารณาความแตกต่างภายในตัวบุคคลเอง ได้นำความสามารถนี้ไปเทียบค่ากับทักษะด้านตัวเลขและความสามารถในการหาเหตุผล ในขณะที่เดียวกัน บุคคลคนที่สามซึ่งมีความสามารถในเรื่องงานธุรการและการพิมพ์ดีด เป็นค่า 25 เปอร์เซนต์ ก็ถือว่าด้อยความสามารถในเรื่องงานที่เกี่ยวกับสำนักงาน ทั้งนี้เพราะพิจารณาเทียบกับความสามารถในด้านตัวเลขและด้านอื่นๆ ของบุคคลคนเดียวกันนี้ จากการเทียบความแตกต่างภายในตัวบุคคลเองเช่นนี้ ย่อมทำให้เกิดแนวพิจารณาที่จะจัดให้บุคคลเหล่านี้ปฏิบัติงานในหน้าที่ตำแหน่งที่เหมาะสม เพื่อให้แต่ละคนได้ใช้ศักยภาพของตนอย่างเต็มที่ ได้ เช่น บุคคลคนที่หนึ่งควรจะปฏิบัติงานที่มีแบบต้องทำซ้ำเป็นประจำ และมีผู้บังคับบัญชาควบคุมใกล้ชิด ส่วนบุคคลที่สองและที่สาม ควรได้รับมอบหมายให้ปฏิบัติหน้าที่ที่มีโอกาสคิดเป็นอิสระและเป็นงานที่ต้องใช้สติปัญญาในการหาเหตุผล ดังนี้เป็นต้น

การฝึกอบรม การฝึกอบรมเป็นขบวนการในการเสริมสร้างทักษะ อุปนิสัย ความรู้ และ ทักษะคติ ในหมู่ผู้ปฏิบัติงาน เพื่อให้บุคคลเหล่านั้นสามารถปฏิบัติงานในตำแหน่งหน้าที่ปัจจุบัน หรือในตำแหน่งหน้าที่ที่จะได้รับมอบหมายต่อไปได้ดีขึ้น การที่จะจัดการฝึกอบรมให้ได้ผลตามที่มุ่งหมายเช่นนี้ นอกจากจะต้องดำเนินการหลายอย่างแล้ว สิ่งหนึ่งที่ขาดไม่ได้ก็คือ การคัดเลือกผู้เข้ารับการอบรมและการประเมินผลการอบรม ซึ่งทั้งสองอย่างนี้ต้องอาศัยการวัดผลบุคคลที่ถูกต้องตามหลักเกณฑ์เป็นหลัก

ในการคัดเลือกผู้เข้ารับการอบรม เครื่องมือวัดบุคคลช่วยให้ทราบว่า ผู้ที่จะควรจะได้เข้ารับการอบรมนั้นมีความสามารถถึงระดับที่จะเรียนรู้วิชาการที่จะฝึกอบรม ได้หรือไม่ ความสามารถถึงระดับที่จะเรียนรู้ประกอบด้วย สติปัญญา ความสัมพันธ์ผลหรือความรู้เดิม ความถนัดพิเศษ หรือทักษะทางกาย แล้วแต่ประเภทของวิชาที่จะฝึกอบรม การที่จะย้ายความสามารถด้านหนึ่งด้านใดในเรื่องต่างๆ เหล่านี้ ก็เป็นเรื่องของความมุ่งหมายของการจัดการฝึกอบรมแต่ละครั้ง หากได้มีการวัดอย่างถูกต้องตามหลักเกณฑ์ ก่อนที่จะ

จัดการอบรมแล้ว ย่อมทำให้การฝึกอบรมที่จัดขึ้นได้รับผลคุ้มค่า เพราะบุคคลที่มีความสามารถไม่ถึงระดับที่จะเรียนรู้วิชาการที่จะฝึกอบรมนั้น ย่อมทำให้เกิดการเสียเวลาทั้งฝ่ายผู้รับการอบรมและผู้ให้การอบรม โดยเฉพาะในเรื่องการฝึกอบรมก่อนประจำการ (pre-service training) ควรจะต้องคัดเลือกรับเฉพาะผู้ที่มีความรู้เป็นพื้นฐานและมีความถนัดพิเศษในระดับหนึ่งแล้วเป็นหลัก ส่วนการฝึกอบรมระหว่างประจำการ (in-service training) นั้น แม้จะถือปฏิบัติว่าผู้ที่ดำรงตำแหน่งหรือปฏิบัติหน้าที่หนึ่ง ๆ อยู่แล้ว เป็นผู้ที่เหมาะสมรับการฝึกอบรมโดยอัตโนมัติก็ตาม แต่การที่จะนำตัวแปรการวัดผลมาวิเคราะห์เลือกผู้สมควรรับการฝึกอบรม ก็ยังจะเป็นประโยชน์ต่อการฝึกอบรมนั้น ๆ การฝึกอบรมในระดับผู้บังคับบัญชาชั้นสูง เช่น การศึกษาตามหลักสูตรของวิทยาลัยป้องกันราชอาณาจักร และการสัมมนานักบริหารของสถาบันบัณฑิตพัฒนบริหารศาสตร์ก็ได้ถืออายุของผู้เข้ารับการอบรมเป็นตัวแปรการวัดผลอย่างหนึ่ง โดยกำหนดเกณฑ์อายุของผู้มีสิทธิรับการอบรมไว้ แม้ผู้ที่มีคุณสมบัติอย่างอื่น ๆ ครบถ้วน แต่มีอายุไม่เข้าเกณฑ์ก็ไม่มีโอกาสได้เข้ารับการฝึกอบรม การกำหนดเช่นนี้ย่อมช่วยให้มีเครื่องประกันได้ว่าผู้ที่จะเข้ารับการอบรมจะสามารถรับวิชาการในการฝึกอบรมได้เต็มที่ และสามารถนำวิชาความรู้ที่ได้ไปปฏิบัติงานต่อไปให้เป็นประโยชน์หรือไม่ อย่างไรก็ดี หากได้มีการวิเคราะห์ตัวแปรการวัดผลให้กว้างขวางตามหลักเกณฑ์ที่กล่าวไว้ในหัวข้อเรื่องตัวแปรการวัดผลที่ไม่ใช่เครื่องมือวัด ก็อาจจะทำให้สามารถได้บุคคลที่เหมาะสมยิ่งขึ้นเข้ารับการอบรมดังกล่าว

ผลของการใช้เครื่องมือวัดเพื่อคัดเลือกบุคคลเข้ารับการอบรม ทำให้การฝึกอบรมแต่ละครั้งได้ผลดีนั้น มีตัวอย่างแสดงไว้แล้วในที่หลายแห่ง ในที่นี้จะได้ยกตัวอย่างของบริษัท Boeing Airplane Company ซึ่งได้ใช้แบบทดสอบสติปัญญา ชื่อ Adaptability Test เพื่อคัดเลือกผู้เข้ารับการอบรมก่อนปฏิบัติงาน ปรากฏว่าผู้ที่ได้ค่าคะแนนสูงจากแบบทดสอบคัดเลือกก่อนเข้ารับการอบรม สามารถรับวิชาการที่จัดอบรมให้เป็นอย่างดีกว่าผู้ที่ได้ค่าคะแนนต่ำในแบบทดสอบคัดเลือก ดังปรากฏรายละเอียดในตารางที่ 31

Adaptability Test เป็นแบบทดสอบมาตรฐานที่ Charles H. Lawshe และ Joseph Tiffin ร่วมกันสร้างขึ้น ในปี 1942 และได้มีการแก้ไขเพิ่มเติมอีกในปี 1954 ในการสร้างแบบทดสอบนี้ได้อาศัยสัมตัวอย่างองค์ประกอบต่าง ๆ ที่เกี่ยวกับสติปัญญามาเป็นแนวในการพิจารณาความเที่ยงตรงของการวัด แบบทดสอบนี้นับเป็นแบบทดสอบที่นิยมแพร่หลายฉบับหนึ่งในการวัดสติปัญญา จากตารางที่ 31 จะเห็นว่า ผู้ที่มีสติมีปัญญาตามค่านิยมของแบบทดสอบนี้สูงกว่าผู้อื่น จะอบรมได้รับความรู้ดีกว่าผู้ที่มีสติปัญญาอ่อนกว่า ดังตัวเลขที่แสดงว่า ผู้ที่มีสติปัญญาสูงในระดับที่ 1 ถึงที่ 25 ในอัตรา 100 คน

ตารางที่ 31

คะแนนคัดเลือกกับผลการอบรม

Adaptability Test	อัตราการยลที่สูงกว่าผู้อื่นหลังการอบรม
25 % แรก	86
25 % ที่สอง	64
25 % ที่สาม	29
25 % สุดท้าย	21

แหล่งที่มา: Guion, Robert M., *Personnel Testing*, p.9.

เมื่ออบรมแล้วจะได้รับความรู้ดีกว่าผู้อื่นเป็นจำนวนถึงร้อยละ 86 ส่วนผู้ที่มีสติปัญญา ก่อนการอบรมรองๆ ลงไป ก็จะมีโอกาสได้รับความรู้ดีกว่าผู้อื่นในปริมาณที่ลดหลั่นตามกันลงไป ดังนี้ จึงเป็นของแน่นอนว่า ถ้าได้คัดเลือกบุคคลเสียก่อนที่จะให้เข้ารับการฝึกอบรม ย่อมเป็นประโยชน์ต่อการจัดอบรมเป็นอย่างดี

การอบรมความรู้วิชาการไม่ว่าในเรื่องใด เช่น การอบรมหน้าที่ของผู้บริหารที่ดี หรือการอบรมวิธีการปฏิบัติงานในสำนักงานที่ดี ย่อมมุ่งหวังให้ผู้รับการอบรมสามารถไปปฏิบัติงานหลังจากการอบรมได้ดีขึ้น ในการที่สนองความมุ่งหมายนี้ การวัดผลที่ถูกต้องตามหลักเกณฑ์ย่อมเป็นเครื่องมือที่มีประโยชน์ เพราะจะทำให้ทราบได้ว่า ผู้รับการอบรมมีความสัมฤทธิ์ผลสมตามโครงการฝึกอบรมที่จัดขึ้นหรือไม่ โดยเฉพาะอย่างยิ่งในกรณีที่มุ่งหวังจะให้ผู้รับการอบรมได้เปลี่ยนไปปฏิบัติงานในหน้าที่ หรือตำแหน่งที่ต้องใช้ความรู้ความสามารถสูงขึ้น นอกจากจะให้มี การวัดความสัมฤทธิ์ผลแล้ว ยังอาจให้วัดความสามารถอื่นๆ โดยใช้หลักการวัดบุคคลเพื่อเลือกบรรจุ (placement) มาอนุโลมใช้ได้อีกด้วย

การเลื่อนขั้นเลื่อนตำแหน่ง การที่ฝ่ายบริหารตัดสินใจเลื่อนขั้นตำแหน่งแก่ผู้หนึ่งผู้ใดที่กำลังปฏิบัติงานอยู่ มักจะทำให้เกิดความเดือดร้อนใจ (grievances) แก่ผู้ปฏิบัติงานอื่นๆ อยู่เสมอ เพราะผู้ที่ไม่ได้รับเลื่อน ซึ่งถือตนว่ามีความรู้ความสามารถไม่ด้อยไปกว่าผู้ที่ได้รับการพิจารณาเลื่อน ต่างก็ได้รับความไม่พอใจในเรื่องนี้ ระบบการเลื่อนโดยอาศัยหลักอาวุโส (seniority) เป็นวิธีทำให้ผู้ปฏิบัติงานส่วนใหญ่พอใจ โดยเหตุที่ผู้ที่ไม่ได้รับการเลื่อนหลายคนก็มีความเกรงใจผู้ที่ปฏิบัติงานมานาน และยังคงเกิดความหวังว่าเมื่อตนมีอาวุโสถึงระดับ ก็จะเป็นโอกาสของตนเช่นเดียวกัน แต่การใช้ระบบนี้ไม่ได้ทำ

ให้ฝ่ายบริหารได้บุคคลที่มีคุณสมบัติที่ดีที่สุด และเหมาะสมที่สุดเท่าที่มีอยู่ปฏิบัติงานในตำแหน่งหน้าที่ที่ว่างลง เพราะความรู้ความสามารถในการปฏิบัติงานในตำแหน่งปัจจุบันไม่จำเป็นต้องเป็นเครื่องประกันว่า จะทำให้มีความรู้ความสามารถปฏิบัติงานในตำแหน่งที่สูงขึ้นได้ดีเท่าเดิม หรือได้เหนือกว่าบุคคลอื่น ๆ ที่มีได้ปฏิบัติงานในตำแหน่งที่บุคคลผู้นั้นดำรงอยู่

วิธีการเลื่อนชั้นเลื่อนตำแหน่งผู้ปฏิบัติงานที่จะช่วยให้ฝ่ายบริหารได้บุคคลที่เหมาะสมที่สุด สำหรับตำแหน่งหน้าที่ที่ว่างลง จึงควรจะต้องอาศัยการวัดผลบุคคลประกอบการวัดความสัมฤทธิ์ผลตามวิธีมาตรฐาน เช่น แบบทดสอบอาชีพ (trade test) หรือแบบทดสอบตัวอย่างงาน (work sample test) ดังได้อธิบายไว้ในบทที่ 12 เป็นสิ่งที่ฝ่ายบริหารควรยึดปฏิบัติ เพราะเป็นวิธีการที่จะทำให้ฝ่ายบริหารทราบได้ว่า บุคคลหนึ่ง ๆ ได้เรียนรู้เนื้อหาของงาน ในตำแหน่งที่จะเลื่อนชั้นไปปฏิบัติมากน้อยเพียงไร จริงอยู่อาจมีผู้ค้านว่า การปฏิบัติงานในตำแหน่งที่จะเลื่อนชั้นไปนั้น ไม่ได้อาศัยความรู้ในเรื่องเนื้อหาของงานเป็นหลัก แต่อาศัยความสามารถส่วนบุคคลซึ่งต้องใช้การสังเกตติดตามเผื่อใจในระยะเวลาที่ปฏิบัติงานผ่านมาเป็นเรื่องสำคัญ แต่เมื่อพิจารณาตามหลักวิชาที่กล่าวไว้ในหนังสือเล่มนี้แล้ว ย่อมจะเห็นได้ว่า การสังเกตเผื่อใจนั้นเป็นสิ่งที่เป็นเรื่องอัตนัย (subjective) และอีกประการหนึ่ง ผู้บังคับบัญชาที่ไม่สามารถสังเกตเผื่อใจผู้อยู่ใต้บังคับบัญชาได้อย่างใกล้ชิดตัวต่อตัวโดยแท้จริงตลอดเวลา และในการสังเกตเผื่อใจนี้ ก็ไม่ได้ให้ความเป็นธรรมแก่ผู้ปฏิบัติงานทุกคน โดยสังเกตเผื่อใจบุคคลเหล่านั้นอย่างใกล้ชิดเท่า ๆ กัน ฉะนั้นข้อวินิจฉัยสั่งการโดยถือเอาการสังเกตเผื่อใจเป็นมาตรฐานย่อมเป็นข้อวินิจฉัยสั่งการที่ขาดความเที่ยงตรง อย่างไรก็ดี ในการที่สนับสนุนการวัดผลบุคคล ในการตัดสินใจเรื่องการเลื่อนชั้นตำแหน่งของผู้ปฏิบัติงานนี้ ก็มีใช้หมายความว่า ฝ่ายบริหารควรถือเอาผลการวัดเป็นเครื่องพิจารณาแต่ประการเดียว ผู้บริหารอาจใช้ค่าคะแนนขั้นต่ำที่กำหนดไว้สำหรับการวัดเหล่านี้ เป็นเครื่องช่วยให้สามารถคัดผู้ที่ไม่สมควรที่จะได้รับการพิจารณาออก และอาศัยผลการสังเกตประกอบการวัด และข้อเท็จจริงอื่น ๆ ที่เกี่ยวกับผู้ปฏิบัติงานคนหนึ่ง ๆ เป็นหลักในการวินิจฉัยสั่งการในเรื่องนี้

โดยเฉพาะการเลื่อนตำแหน่งงานในระดับงานที่ไม่ใช่ระดับผู้บังคับบัญชาแบบทดสอบตัวอย่างงานและแบบทดสอบอาชีพมีประโยชน์อย่างเห็นได้ชัด⁵ การทดลองใช้แบบทดสอบอาชีพเป็นหลักในการวินิจฉัยว่า ผู้ปฏิบัติงานคนใดสมควรได้รับการเลื่อน

⁵ Lawshe, C.H. and Balma, Michael J., *Principles of Personnel Testing*, pp. 106–107. Guion, *op. cit.*, pp. 406–407.

ตำแหน่ง ในบริษัท Collins Radio Company ในรัฐ Iowa ปรากฏว่า ทำให้จำนวน ผู้ได้รับเลื่อนตำแหน่งที่ไม่สามารถปฏิบัติงานได้ดังที่เคยมีมาแต่ก่อนลดลงเหลือเพียงร้อยละ 2 เท่านั้น บริษัท International Harvester Company ถึงกับได้สร้างแบบทดสอบ ตัวอย่างงานเพื่อใช้ในการพิจารณาเลื่อนชั้นเจ้าหน้าที่ธุรการโดยเฉพาะ แบบทดสอบนี้ ชื่อว่า Clerical Qualification Tests ใช้วัดบุคคลที่จะได้รับการพิจารณาเลื่อนตำแหน่งงาน 30 ตำแหน่ง

การวัดบุคคลในด้านอื่น ๆ เช่น บุคลิกภาพ หรือสติปัญญา อาจใช้เป็นเครื่อง ประกอบร่วมไปกับการวัดความสัมฤทธิ์ผล จริงอยู่ บุคลิกภาพมีความสำคัญต่อการ ปฏิบัติงานในตำแหน่งที่สูงขึ้นไป แต่ปัญหาในเรื่องนี้มีอยู่ว่า ยังขาดเครื่องมือและวิธีการ ที่จะสร้างเครื่องมือวัดลักษณะนิสัยด้านนี้ของบุคคลให้มีความเที่ยงตรง และความเชื่อมั่น ในระดับที่น่าพอใจเท่าเทียมกับการวัดลักษณะนิสัยด้านอื่น ๆ ของบุคคล ส่วนการที่ กล่าวถึงว่าอาจวัดสติปัญญาประกอบนั้น ก็เพราะว่า ในการเลื่อนชั้นตำแหน่งนี้ ควร จำกัดความของปัญหาว่า จะเลือกผู้ที่มีคุณสมบัติความสามารถเหมาะสม หรือจะเลือก ผู้ที่อาจจะเป็นผู้มีความสามารถเหมาะสม ถ้าปัญหาอยู่ที่ประเด็นแรก ก็หมายความว่า ต้องเลือกผู้ที่มีความรู้และสามารถปฏิบัติงานในตำแหน่งใหม่ทันที ส่วนผู้ที่มีสติปัญญา แต่ไม่มีความรู้ในการปฏิบัติงาน เป็นผู้ที่มีความสามารถซึ่งอาจปฏิบัติงานในตำแหน่ง นั้น ๆ ได้ หากได้รับโอกาสเวลาสักกระยะหนึ่งก่อน

การวัดเจ้าหน้าที่สำนักงาน

เพื่อที่จะเป็นแนว ในการใช้ เครื่องมือวัดให้เป็นประโยชน์ต่อการจัดงานบุคคลอย่าง เต็มที่ ในทันทีจะได้นำผลการวิจัยในเรื่องผลดีผลเสียของการใช้เครื่องมือวัดเพื่อวินิจฉัย สิ่งการในการจัดงานบุคคลระดับเจ้าหน้าที่สำนักงาน (office personnel) กับระดับผู้บริหาร งาน (managerial personnel) มาเป็นแนว เมื่อจะพิจารณาการใช้เครื่องมือวัดบุคคลกับ เจ้าหน้าที่ฝ่ายอื่น ๆ เช่น เจ้าหน้าที่วิชาการหรือเจ้าหน้าที่อื่น ๆ ในโรงงาน ก็อาจอนุมาน หลักการที่จะกล่าวถึงสำหรับเจ้าหน้าที่ทั้งสองระดับนี้ได้

ลักษณะงานของเจ้าหน้าที่สำนักงาน งานของเจ้าหน้าที่สำนักงาน ประกอบด้วยงาน หลายอย่างหลายประเภท งานระดับนี้อาจแยกอย่างหยาบได้ 3 ประเภท คือ ประเภท งานเสมียน ประเภทบริการด้านธุรการ และประเภทควบคุมงาน ตำแหน่งงานประเภท แรกได้แก่พนักงานเก็บเอกสาร พนักงานรับโทรศัพท์ เจ้าหน้าที่ต้อนรับ ตำแหน่งงาน ประเภทบริการด้านธุรการได้แก่ เลขานุการส่วนตัว พนักงานพิมพ์ดีดชวเลข พนักงาน

เดินหนังสือ ส่วนตำแหน่งงานประเภทควบคุมงานได้แก่ พนักงานลงบัญชี พนักงานควบคุมรายการพัสดุ L.L. Thomas ได้ทำการวิจัยลักษณะของงานของเจ้าหน้าที่สำนักงาน 112 ตำแหน่ง ในปี 1952 และค้นพบว่า⁶ งานระดับนี้อาจแยกเป็นกลุ่มได้ 8 กลุ่มคือประเภทพิมพ์ดีด ชวเลข และควบคุมเครื่องกลในสำนักงาน ประเภทจัดทำบัญชีและประมวลข้อเท็จจริง ประเภทสื่อสาร ประเภทวางแผนและควบคุมงาน ประเภทเก็บเอกสาร ประเภทจัดการวัสดุครุภัณฑ์ ประเภทงานเสมียนที่ต้องปฏิบัติประจำวัน และประเภทคำนวณตัวเลข

ลักษณะของงานในสำนักงานที่บรรยายมานี้ แสดงว่า จะใช้ชุดเครื่องมือวัด (test battery) ใด ๆ โดยเฉพาะได้ยาก เพราะประกอบด้วยงานหลายอย่างต่าง ๆ กัน อีกทั้งจะใช้เครื่องมือวัดบุคคลในด้านสติปัญญาเพื่อคาดคะเนความสำเร็จของผู้ที่จะปฏิบัติงานก็เป็นการไม่ตรงประเด็น (valid) ที่เดียว เครื่องมือที่จะใช้ในการทดสอบบุคคลที่จะปฏิบัติงานประเภทนี้จึงควรเป็นแบบทดสอบที่วัดความเร็วและความแม่นยำ ทั้งในเรื่องตัวเลขและในด้านภาษา

จากตารางที่ 27 ในบทที่ 8 ซึ่งแสดงเกณฑ์เฉลี่ยของความสามารถในการคาดคะเนของแบบทดสอบเขาวัวที่ Ghiselli และ Brown ได้ทำการวิจัยไว้ จะเห็นว่าผลการวัดสติปัญญาเพื่อคาดคะเนความสำเร็จของผู้ที่จะปฏิบัติงานในระดับเสมียนไม่มีคุณภาพดีเท่าผลการวัดสติปัญญาเพื่อคาดคะเนความสำเร็จของผู้ที่จะปฏิบัติงานในระดับผู้บริหารหรืองานประเภทเดินตลาด (sales) ในการปฏิบัติงานในสำนักงาน ความสามารถในการหาเหตุผลและในการเรียนรู้ นับว่ามีความสัมพันธ์เพียงส่วนน้อยกับการปฏิบัติงานในหน้าที่เสมียน

E.N. Hay⁷ ได้ทำการวิจัยเปรียบเทียบคุณภาพในการคาดคะเนความสำเร็จของการปฏิบัติงานในสำนักงานระหว่างเครื่องมือวัดสติปัญญา และเครื่องมือวัดความเร็วและความแม่นยำของสัญชาตญาณ (perception) ในปี 1954 ในการวิจัยนี้ได้ใช้แบบทดสอบชื่อ Wonderlic Personnel Test, SRA Clerical Test และ Hay Clerical Battery โดยทดลองกับกลุ่มเจ้าหน้าที่ในสำนักงานของบริษัทประกันภัยแห่งหนึ่ง Wonderlic Personnel Test เป็นแบบทดสอบที่ปรับปรุงดัดแปลงมาจาก Otis Self Administering Tests of Mental Ability คำถามที่มีในแบบทดสอบเกี่ยวกับศัพท์ประเภทต่างๆ การหาเหตุผล

⁶ *Ibid.*, p. 416.

⁷ Lawshe and Balma, *op. cit.*, pp. 182—183.

โดยอาศัยสามัญสำนึก การคิดหาเหตุผลในเรื่องตัวเลขและการคำนวณตัวเลข ทักษะในเชิงสัญชาตญาณ (perceptual skill) ความสัมพันธ์ของมิติ (spatial relations) ลักษณะของชุดตัวเลขและตัวอักษร แบบทดสอบฉบับนี้เป็นที่นิยมกันแพร่หลายว่าสามารถวัดสติปัญญา ในเรื่องทั่ว ๆ ไปที่จำเป็นต้องมีในการปฏิบัติงาน SRA Clerical Test นั้น เป็นแบบทดสอบความสามารถในการวัดงานเสมือนในสำนักงาน ประกอบด้วยคำถามในเรื่องศัพท์ที่ต้องใช้ในสำนักงาน การคำนวณตัวเลขและการตรวจสอบบางอย่างที่เจ้าหน้าที่ในสำนักงานต้องปฏิบัติ คำถามในเรื่องศัพท์คือคำถามที่วัดความเข้าใจในเรื่องภาษา (verbal comprehension) นั่นเอง ส่วนคำถามเกี่ยวกับการคำนวณตัวเลข มิได้เป็นการบวกเลขตามธรรมดา แต่เป็นการวิเคราะห์ปัญหาตามเรื่องที่กำหนด การทดสอบตัวเลขนี้จึงมิใช่การทดสอบความคล่องในการใช้ตัวเลข (numerical) แต่เป็นการทดสอบความสามารถในการหาเหตุผลทั่ว ๆ ไป (general reasoning) มากกว่า และคำถามในเรื่องการตรวจสอบในสำนักงานนั้นก็เป็นการทดสอบในแง่ความจำมากกว่าอย่างอื่น เพราะผู้ตอบคำถามจะต้องจัดคู่ของคำสามพยางค์ให้ตรงกับตัวเลขที่กำกับไว้แต่ละคำให้สำเร็จภายในเวลาที่กำหนดให้ ความสามารถของงานเสมือนที่ทดสอบนี้ จึงเป็นความสามารถในเชิงสติปัญญามากกว่าความสามารถในแง่ความเร็วของสัญชาตญาณ (perceptual) อันควรเป็นความสามารถหลักของเจ้าหน้าที่ในสำนักงาน ส่วน Hay Clerical Battery นั้น เป็นชุดของแบบทดสอบที่ผู้วิจัยสร้างขึ้นเองเพื่อวัดความสามารถของเจ้าหน้าที่ในสำนักงานโดยเฉพาะ แบบทดสอบชุดนี้ประกอบด้วยคำถามเกี่ยวกับสัญชาตญาณในเชิงตัวเลข (number perception) ซึ่งมีลักษณะทำนองเดียวกับคำถามในแบบทดสอบ Minnesota Clerical Test นอกจากนี้ยังมีคำถามที่มุ่งหวังจะวัดความคล่องในเรื่องตัวเลข และความแม่นยำในการเปรียบเทียบคำที่ใช้ในภาษา ผลของการวิจัยปรากฏว่า ผู้ที่ได้ค่าคะแนนสูงในผลการวัดของแบบทดสอบ Hay Clerical Battery เป็นเจ้าหน้าที่ที่ปฏิบัติงานดีในสำนักงาน ส่วนแบบทดสอบ Wonderlic Personnel Test และแบบทดสอบ SRA Clerical Test ภาควัดความเข้าใจในเรื่องภาษาและความคล่องในการใช้ตัวเลข นับว่ายังขาดความเที่ยงตรงในแง่การวัดบุคคลผู้ปฏิบัติหน้าที่ในสำนักงาน

ในการใช้เครื่องมือวัดเพื่อวัดบุคคลที่จะปฏิบัติหน้าที่ในสำนักงานนั้น Guion⁸ ได้ทำการค้นคว้าเปรียบเทียบผลการวิจัยของผู้ทรงคุณวุฒิหลายคน เช่น J.T. Rusmore, G.J. Toorenaar, E.B. Knauft และ W.J. McNamara เป็นต้น และได้สรุปว่า หากจะใช้เครื่องมือวัดผู้ปฏิบัติงานในหน้าที่ระดับนี้ให้ได้ผล ควรจะต้องดำเนินการในเรื่องต่อไปนี้ให้

⁸ Guion, *op. cit.*, p. 424.

ครบทั้ง 4 ประการ คือ ประการแรก ควรใช้เครื่องมือวัดความเร็วและความแม่นยำของ
 สัญชาตญาณ เพื่อคาดคะเนความสำเร็จผลของการปฏิบัติงานของเจ้าหน้าที่ในสำนักงานใน
 ตำแหน่งงานทั่ว ๆ ไป แต่ในตำแหน่งงานที่ต้องรับผิดชอบสูงขึ้น ควรใช้เครื่องมือวัด
 ทักษะในด้านภาษาและตัวเลข และความสามารถในการคิดหาเหตุผลประกอบไปด้วย
 นอกจากนี้ยังควรใช้แบบทดสอบตัวอย่างงาน ซึ่งเป็นเครื่องมือวัดความสัมฤทธิ์ผลเพิ่ม
 เต็มเฉพาะแต่ละตำแหน่ง เพื่อให้สามารถได้ผลการคาดคะเนที่สมบูรณ์ยิ่งขึ้น ประการ
 ที่สอง ควรใช้ข้อเท็จจริงจากแบบฟอร์มประวัติส่วนตัวของผู้ปฏิบัติงานหรือผู้สมัครงาน
 ประกอบด้วย เพื่อให้ทราบทัศนคติและความสนใจในการที่จะปฏิบัติงาน แบบฟอร์ม
 ประวัติส่วนตัวนี้หมายถึงแบบรายการที่กำหนดเป็นพิเศษเพื่อให้มีข้อมูลเกี่ยวกับภูมิหลัง
 ในเรื่องการปฏิบัติงานโดยละเอียดเก็บไว้ในสำนักงาน มิใช่เป็นแบบใบสมัครงานทั่ว ๆ
 ไป ในเรื่องนี้อาจใช้การสัมภาษณ์ซักถามประวัติการทำงานแทนแบบฟอร์มนี้ก็ได้อีก การ
 ที่กระทำเช่นนั้นก็เพื่อให้สามารถคาดคะเนอัตราร้อยละของบุคคลที่จะต้องบรรจุแทนผู้ที่ลา
 ออกไปในโอกาสหนึ่ง ๆ ในวาระต่อ ๆ ไป ประการที่สาม ในการพิจารณามาตรฐานวัดความ
 สำเร็จของงานควรจะต้องถือความพอใจในการปฏิบัติงาน (job satisfaction) ของผู้ปฏิบัติ
 งานเป็นเกณฑ์ ทั้งนี้อาจพิจารณาได้จากแบบฟอร์มประวัติส่วนตัว ประการสุดท้ายจะ
 ต้องพยายามใช้มาตรฐานวัดความสำเร็จของงานหลาย ๆ ด้าน ประกอบการพิจารณาในการวัด
 บุคคลผู้ปฏิบัติงานในระดับนี้ เพราะสัมประสิทธิ์สหสัมพันธ์ระหว่างการปฏิบัติงานใน
 สำนักงานกับมาตรฐานวัดความสำเร็จของงานในด้านหนึ่ง ๆ นั้นมักมีระดับต่ำ ฉะนั้น การ
 จำกัดการวัดให้แคบเฉพาะมาตรฐานวัดความสำเร็จของงานด้านหนึ่งด้านใด จึงทำให้เครื่อง
 มื่อวัดสามารถวัดผลให้เป็นที่พอใจได้แต่เพียงด้านเดียวไม่ครบถ้วน

การวัดเจ้าหน้าที่ระดับผู้บริหารงาน

การใช้เครื่องมือวัดประกอบการวินิจฉัยสั่งการในการจัดงานบุคคลเป็นเรื่องที่โต้แย้ง
 กันมาช้านานแล้วว่าเป็นของที่เชื่อถือได้จริงหรือไม่ และสมควรใช้หรือไม่ โดยเฉพาะ
 การใช้เครื่องมือวัดในการจัดงานบุคคลระดับผู้บริหารนั้นยังมีผู้คัดค้านเป็นจำนวนไม่น้อย
 ในปัจจุบันก็ไม่เชื่อว่าทุกคนจะปลงใจเหมือนกันหมดว่าควรใช้เครื่องมือวัดในการจัดงาน
 บุคคลระดับนี้ Sands⁹ ได้พิมพ์ผลงานวิจัยในเรื่องนี้โดยเฉพาะในปี ค.ศ. 1963 ผู้
 วิจัยได้ศึกษาบริษัทธุรกิจขนาดใหญ่ของสหรัฐอเมริกาซึ่งมีปริมาณขาย (sales volume)
 ตั้งแต่หนึ่งพันล้านเหรียญอเมริกันขึ้นไปเป็นจำนวน 82 บริษัท บริษัทธุรกิจเหล่านี้มีหลาย

⁹ Sands, Edith, *How to Select Executive Personnel*, p. 109.

ประเภท กล่าวคือ มีทั้งบริษัทประกันภัย บริษัทที่ประกอบการพาณิชย์โดยเฉพาะ ธุรกิจ
ในด้านสาธารณูปโภคและธนาคารต่างๆ ผลปรากฏว่า ธุรกิจในด้านสาธารณูปโภค และ
ธนาคารต่างๆ เกินกว่ากึ่งหนึ่ง ไม่แน่ใจว่าการใช้เครื่องมือวัดจะมีประโยชน์ต่อการจัด
งานบุคคลของตนทั้งๆ ที่ตนก็เคยทดลองใช้เครื่องมือวัดต่างๆ มาแล้ว แต่บริษัทประกัน
ภัยและบริษัทที่ประกอบการพาณิชย์โดยเฉพาะหลายแห่งกลับมีความเห็นเป็นตรงกันข้าม
ว่า เครื่องมือวัดมีประโยชน์ในการคัดเลือก และเลือกสรรผู้เจ้าหน้าที่ระดับผู้บริหารงาน
อย่างไรก็ดี ถึงแม้ว่าบุคคลทุกฝ่ายและทุกคนจะยังไม่ยอมรับประโยชน์ของเครื่องมือวัด
บริษัทธุรกิจต่างๆ ก็ได้ใช้เครื่องมือวัดเพื่อประโยชน์ในการจัดงานบุคคลทุกระดับ รวมทั้ง
ระดับผู้บริหารงานมาเป็นเวลาไม่น้อยกว่าสองทศวรรษ (decade) แล้ว ดังปรากฏผลใน
รายงานการวิจัยซึ่งได้ดำเนินการก่อนหน้าการวิจัยของ Sands เสียอีก คือ ในปี ค.ศ. 1957
ได้มีการวิจัยแบบสำรวจ (survey research)¹⁰ ด้วยการศึกษานักบริหารชั้นสูงสุด (top-level
executives) เป็นจำนวน 116 นาย ปรากฏว่าบุคคลเหล่านี้เป็นจำนวนกว่าหนึ่งในสามได้
ผ่านการทดสอบโดยเครื่องมือวัดมาแล้ว บุคคลที่ต้องผ่านการทดสอบโดยเครื่องมือวัด
ก่อนที่จะได้ดำรงตำแหน่งนักบริหารงานเหล่านี้มีถึงระดับประธาน (president) ของบริษัท
ฉะนั้นปัญหาที่จะต้องพิจารณาในเรื่องนี้ก็คือว่า เครื่องมือวัดมีคุณค่าและอาจใช้ประโยชน์
ในการจัดงานบุคคลระดับผู้บริหารงานอย่างไร แต่ก่อนที่จะพิจารณาประเด็นนี้ควรจะ
วิเคราะห์ลักษณะของนักบริหารและลักษณะของงานระดับนี้ เพื่อให้เกิดความเข้าใจ
เบื้องต้นเสียก่อน

ลักษณะของนักบริหารและลักษณะของงานบริหาร บทบาทของนักบริหารอาจนิยาม
สั้นๆ ได้ว่า เป็นการจัดและส่งเสริมความพยายามและการปฏิบัติงานของบุคคลต่างๆ
ในกลุ่มให้ประสานกัน เพื่อให้บรรลุเป้าหมายที่กำหนดไว้¹¹

บุคคลที่มีหน้าที่รับผิดชอบในฐานะนักบริหารจะต้องจัดกิจการของหน่วยงานให้ดำเนิน
ไปในทิศทางที่เป็นเหตุเป็นผล และเป็นประโยชน์ในการที่จะฝ่าฟันอุปสรรคและ
ปัญหาต่างๆ อันเกิดขึ้นจากสภาพการณ์หนึ่งๆ นักบริหารจะต้องวิเคราะห์และแปลความ
หมายของข้อมูลทั้งหมดที่มีอยู่ พร้อมกันนั้นก็จะต้องชี้แจงนำหนักความสำคัญของหลัก
ฐานต่างๆ ที่เกี่ยวข้องกับกรปฏิบัติงาน เพื่อที่จะวินิจฉัยสั่งการให้เหมาะสมสอดคล้อง
กับกรณีการปฏิบัติงานแต่ละกรณี จึงกล่าวกันว่านักบริหารเป็นผู้ที่คิดในแง่การปฏิบัติ

¹⁰ Lawshe and Balma, *op. cit.*, pp. 208–209.

¹¹ Simon, Herbert A., *Administrative Behavior*, pp. 8, 217–218, 145.

Pfiffner, John M., and Prestus, R. Vance, *Public Administration*, pp. 3–5, 74, 548–560.

(action-oriented) มากกว่าที่จะเป็นนักปรัชญา การคิดในแง่ปฏิบัตินี้หมายความว่าต้องเน้นความสนใจของตนในปัญหาสำคัญที่อาจแก้ไขได้ และในขณะเดียวกันต้องยอมรับสภาพความเป็นจริงว่า ปัญหาบางอย่างเป็นสิ่งที่ไม่ได้ เพราะเป็นเรื่องของอิทธิพลของสิ่งแวดล้อมอันอยู่นอกเหนืออิทธิพลของการบริหาร อย่างไรก็ตาม ปัญหาในเรื่องอิทธิพลของสิ่งแวดล้อมเหล่านี้ก็จะต้องมีลักษณะเปลี่ยนแปลงไปตามกาลเวลาและสภาพของสังคม นักบริหารก็จะศึกษาแนวคิด (concept) ของความเปลี่ยนแปลงเหล่านี้ เพื่อให้สามารถปฏิบัติการได้อย่างเหมาะสมอยู่เสมอ ในการวิเคราะห์สภาพการณ์ต่าง ๆ นักบริหารจะพยายามค้นหาว่าปัญหาใดเป็นปัญหาหลัก เพื่อให้ทราบว่าตนต้องการข้อมูลในลักษณะใด ประเภทใด ในการพิจารณาการปฏิบัติงานอยู่ตลอดเวลา

ในการอธิบายบทบาทของนักบริหาร มีตัวอย่างซึ่งขมกวมอ้างอิงกันแพร่หลาย คือ ตัวอย่างของบุคคลสองคนที่ช่วยกันเคลื่อนก้อนหินก้อนหนึ่งซึ่งหนักเกินกว่าที่จะยกเพียงคนเดียว การกระทำนี้มีความมุ่งหมายคือ การเคลื่อนก้อนหิน และถือได้ว่าเป็นการปฏิบัติที่ต้องประสานกัน เพราะคนใดคนหนึ่งเพียงคนเดียวเคลื่อนก้อนหินไปไม่ได้ การเคลื่อนก้อนหินนี้อาจทำได้หลายวิธี เช่น ดึงหรือผลัก ใช้คานไม้หรือคานเหล็กงัด ใช้เชือกผูกหรือใช้ลูกรอกกลาก หรืออาจทุบหินให้แตกแล้วช่วยกันยกไปเป็นชิ้น ๆ นักบริหารในกรณีนี้ คือ ผู้ที่ตัดสินใจจะใช้วิธีใดในการเคลื่อนก้อนหิน จะดำเนินการอย่างไรในการกำหนดว่าผู้ใดจะต้องปฏิบัติอะไรบ้าง ในการปฏิบัตินั้นจะแบ่งงานกันอย่างไร แต่ละคนจะรู้ได้อย่างไรว่าตนจะต้องปฏิบัติอะไรบ้าง ในทำนองใด และผู้ปฏิบัติงานเหล่านี้จะมีความรู้ในการปฏิบัติงานได้อย่างไร หรือกล่าวรวม ๆ ว่าจะประสานความพยายามและการปฏิบัติงานของทุกคนได้อย่างไรนั่นเอง

ผลการศึกษาวิจัยทางการบริหารได้ทำให้ทราบหลักการหลายอย่างที่จะใช้เป็นเครื่องมือประกอบการพิจารณาสร้างความประสานนี้ เช่น การวางแผน การควบคุมโดยประมุข-ศิษฐ์ วิธีกาส่งข่าวสาร และการจัดหน่วยงาน เป็นต้น ในกรณีของก้อนหินที่ยกมาเป็นตัวอย่างก็อาจสร้างความประสานได้หลายวิธี เช่น ผู้ที่จะเคลื่อนก้อนหินสนทนาแลกเปลี่ยนความคิดจนตกลงกันได้ว่า จะเคลื่อนก้อนหินไปโดยวิธีใด หรืออาจมีการว่าจ้างเกิดขึ้นโดยบุคคลคนหนึ่งเป็นผู้กำหนดวิธี และส่งให้อีกบุคคลหนึ่งช่วยในการเคลื่อนย้าย การปฏิบัติงานทุกอย่างในสังคมไม่ว่าจะเป็นการสร้างถนน การต่อสู้คดีก็ภัย การขายสินค้า หรืออื่น ๆ ก็จะต้องอาศัยหลักทางการบริหาร เพื่อช่วยในการสร้างความประสานตามแนวคิดนี้ทั้งนี้

การวินิจฉัยสั่งการในเชิงบริหารเพื่อประสานความพยายามและการปฏิบัติงานของบุคคลเป็นงานสำคัญของนักบริหาร แต่นักบริหารมิได้รับผิดชอบงานนี้เพียงอย่างเดียว นักบริหารจะต้องวินิจฉัยสั่งการในแง่เนื้อหาของงานที่จะต้องปฏิบัติด้วยการวินิจฉัยสั่งการในแง่เนื้อหาของงาน กระทำโดยการสร้างเจตนาธรรมณ์ของหน่วยงานขึ้น เพื่อเป็นแนวในการตัดสินใจของผู้ปฏิบัติงานต่าง ๆ การที่จะสร้างเจตนาธรรมณ์นี้ได้ก็หมายความว่าผู้บริหารจะต้องสามารถเข้าใจวิชาการที่เกี่ยวข้องกับเนื้อหาของงานที่จะปฏิบัติถึงระดับที่จะแลกเปลี่ยนความคิดเห็นกับผู้ปฏิบัติงานในสาขาต่าง ๆ ได้เป็นอย่างดี และผู้ปฏิบัติงานเหล่านี้เห็นว่าผู้บริหารมีความเข้าใจในข้อเสนอแนะของตนพอที่ตนจะให้ความยกย่องว่าเป็นผู้นำได้ นอกจากนี้นักบริหารยังจะต้องรับผิดชอบในด้านกรงบประมาณ ซึ่งก็หมายความว่านักบริหารจะต้องเข้าใจเนื้อหาของงานพอที่จะทราบได้ว่าควรเน้นหนักงานประเภทใด เพื่อกำหนดให้การปฏิบัติงานเป็นไปในทิศทางที่ถูกต้อง อย่งไรก็ดี อาจกล่าวได้ว่านักบริหารที่อยู่ในระดับที่ยิ่งสูงขึ้นไป ก็ยิ่งเกี่ยวข้องกับงานด้านบริหารมากขึ้น และเกี่ยวข้องกับงานด้านวิชาการน้อยลงตามส่วน ทั้งนี้ เพราะนักบริหารระดับสูงวินิจฉัยสั่งการโดยคำนึงถึงความมุ่งหมายของหน่วยงาน และสภาพปัญหาของสถานการณ์ในวงกว้างกว่านักบริหารในระดับรองลงไป

โดยสรุปนักบริหารมีหน้าที่ 3 ประการ คือ :-

1. กำหนดเจตนาธรรมณ์หรือแนวนโยบายของหน่วยงาน
2. จัดหน่วยงานและกำหนดทิศทางของการดำเนินงานของหน่วยงาน
3. วางแนวปฏิบัติ กำหนดระเบียบหรือหลักการ วินิจฉัยสั่งการพร้อมๆ กับปรับปรุงแนวปฏิบัติ และระเบียบ หรือหลักการให้สอดคล้องกับความเปลี่ยนแปลงของสถานการณ์ในสังคมและกฎหมายต่างๆ

คณะกรรมการศึกษาวิเคราะห์ข้าราชการของสหรัฐอเมริกา (Commission of Inquiry on Public Service Personnel) ได้แยกความรับผิดชอบของนักบริหารระดับกระทรวงไว้ว่ามี 6 ประการ ดังต่อไปนี้¹² :-

1. ตีความกฎหมาย และระเบียบต่างๆ ที่เกี่ยวข้องกับการปฏิบัติงาน
2. แบ่งงานให้ทุกคนปฏิบัติในลักษณะที่จะทำให้บรรลุผลตามนโยบายที่กำหนดไว้
3. รักษาความคงเส้นคงวา (consistency) ในการปฏิบัติงานภายในกระทรวง และกับกระทรวงอื่นๆ

¹² Stone, Donald C., "Applying Management Knowledge", in Marx, Fritz Morstein, *Elements of Public Administration*, pp. 414-415.

4. รักษาความต่อเนื่อง (continuity) ในการปฏิบัติในระยะที่ผ่านมากับระยะปัจจุบันและอนาคต ยกเว้นแต่ในกรณีที่มีการเปลี่ยนแปลงนโยบาย
5. ส่งเสริมให้ผู้ปฏิบัติงานตื่นตัวและให้ผลงานมีประสิทธิภาพ
6. รายงานต่อประชาชนและพร้อมที่จะอธิบายการปฏิบัติงานของกระทรวง

แนวคิดของคณะกรรมการดังกล่าวเป็นเรื่องเฉพาะของนักบริหารงานรัฐ ส่วนนักบริหารหน่วยงานอื่น ๆ ที่ไม่ใช่ของรัฐ เช่น นักบริหารธุรกิจ และนักบริหารงานสังคมสงเคราะห์ ต้องปฏิบัติงานในสภาพแวดล้อมที่ต่างออกไป ลักษณะหน้าที่ความรับผิดชอบ จึงต้องอธิบายในลักษณะที่ผิดแผกออกไปบ้าง อย่างไรก็ตามอย่างไรก็ดี *mutatis mutandis* บทบาทของนักบริหารทั่ว ๆ ไปก็อาจอนุมานตามแนวคิดของคณะกรรมการนี้ จากการสัมมนาวิจัย (research conference) ระหว่าง American Society for Public Administration, International City Managers' Association, Fels Institute of Local and State Government และ University of Pennsylvania ได้สรุปว่านักบริหารจะต้องรับผิดชอบต่อเรื่องต่าง ๆ ดังต่อไปนี้

1. ต้องมีการสนองตอบ (responsiveness) ต่อผู้บังคับบัญชาและกลุ่มผู้ใช้ผลประโยชน์จากการปฏิบัติงานของหน่วยงาน (clientele) กล่าวคือ ปฏิบัติงานในด้านที่จะช่วยผู้บังคับบัญชาให้ดำเนินงานอย่างมีประสิทธิภาพ และสร้างความเข้าใจอันดีต่อนโยบายและโครงการดำเนินงานในหน่วยงานในระหว่างผู้ใช้ผลประโยชน์จากการปฏิบัติงานของหน่วยงาน

2. ช่วยในการวางเป้าหมายและโครงการ นักบริหารต้องสามารถนำเสนอแนะของผู้อยู่ใต้บังคับบัญชามาพิจารณาในการสร้างเป้าหมายและโครงการของหน่วยงานเพื่อนำเสนอผู้บังคับบัญชาพิจารณาอีกชั้นหนึ่ง และจะต้องประสานกิจกรรมของบุคคลและหน่วยงานที่มีได้อยู่ภายใต้การบังคับบัญชาของตน แต่มีผลกระทบเกี่ยวข้องกับดำเนินงานของหน่วยงานของตนให้สอดคล้องเป็นประโยชน์ในการแก้ปัญหาอันอาจเกิดจากการปฏิบัติงานตามเป้าหมาย

3. จัดหน่วยงานและควบคุมทรัพยากร ทรัพยากรของนักบริหารคือ คน เครื่องมือ วัสดุ และเงิน ซึ่งนักบริหารมีหน้าที่จัดในลักษณะที่จะเป็นประโยชน์ต่อหน่วยงานอย่างเต็มที่ ทั้งนี้ ทำได้โดยกำหนดมาตรฐานในการดำเนินงาน มีการมอบหมายอำนาจ และจัดการบริหารงานบุคคลให้เป็นไปตามหลักวิชา

4. **อำนาจการและนิเทศงาน** นักบริหารงานจะต้องสร้างการส่งข่าวสาร (communication) ที่รัดกุม สามารถปกครองผู้อยู่ใต้บังคับบัญชาโดยลักษณะของผู้นำที่ดี เพื่อเป็นทางให้ผู้ปฏิบัติงานทุกคนยอมรับและเกิดความกระตือรือร้นในการปฏิบัติงานตามนโยบายและโครงการที่วางไว้ อีกทั้งจะได้ส่งเสริมให้นักวิชาการบำเพ็ญตนให้เป็นประโยชน์ต่อหน่วยงานอย่างเต็มที่

เมื่อพิจารณารวม ๆ จึงถือกันว่างานของนักบริหารก็คือ การทำตนเป็นผู้นำของหน่วยงาน เพราะการปฏิบัติงานของหน่วยงานจะลุล่วงไปได้ด้วยความสำเร็จก็ต่อเมื่อมีผู้นำที่จะคอยช่วยให้กลุ่มบุคคลที่ปฏิบัติงานร่วมกันมีเป้าหมายตรงกัน มีการวางแผนและโครงการในการปฏิบัติงานและคอยช่วยปรับปรุงโครงสร้างของหน่วยงานให้สอดคล้องกับความเปลี่ยนแปลงของสภาพแวดล้อมและปัญหาที่เกี่ยวข้อง

เพื่อให้ทราบว่าบุคคลที่มีลักษณะเช่นใดจึงจะสามารถปฏิบัติหน้าที่รับผิดชอบที่กล่าวมาทั้งหมดได้ผลดี นักวิชาการหลายคนจึงได้พยายามทำการวิจัยเรื่องนี้ ผลการวิจัยของ Wald และ Doty¹³ ในปี ค.ศ. 1960 นับว่ามีรายละเอียดพอสมควรที่จะนำมากล่าวไว้โดยสรุปในที่นี้ คณะผู้วิจัยได้ศึกษากรณีตัวอย่างของนักบริหารชั้นสูงสุดที่มีชื่อเสียงว่ามีความสามารถเป็นต้นฉบับถือกันทั่วไปเป็นจำนวน 33 นาย ปรากฏว่า บุคคลเหล่านี้มีลักษณะต่าง ๆ 13 ประการข้างล่างนี้ คือ

1. เป็นผู้มีชีวิตครอบครัวในวัยเยาว์อย่างเหมาะสม
2. เป็นผู้ที่ห่วงใยและใช้ชีวิตกับครอบครัวในปัจจุบันอย่างใกล้ชิด
3. มีวุฒิการศึกษาสูงกว่าเกณฑ์เฉลี่ยโดยทั่วไป
4. ขวนขวายหาวิชาความรู้เสมอแม้ในปัจจุบัน
5. มีส่วนร่วมในกิจการสังคมและองค์การที่จัดตั้งขึ้นเพื่อประโยชน์ทางสังคม
6. ถือเอาศาสนาเป็นเครื่องยึดทางศีลธรรมและทางจิตใจ
7. สุขภาพอนามัยดี
8. สนใจบุคคลทั่วไป ตลอดจนการแลกเปลี่ยนความรู้ความคิด และปฏิบัติงานร่วมกับผู้อื่น โดยอาศัยหลักมนุษยสัมพันธ์
9. มีความฉลาดกว่าบุคคลทั่วไป
10. เอาใจใส่ต่อการทำงาน และเป็นผู้ที่ตระหนักในหน้าที่ความรับผิดชอบ
11. มีพลังความเข้มแข็งทางใจที่จะมุ่งมั่นในการปฏิบัติงาน

¹³ Guion, *op. cit.*, p. 458.

12. เป็นคนตรงไปตรงมา และพิจารณาสิ่งต่าง ๆ ในแง่เหตุผล ไม่เอาแต่ความรู้สึกของตนในการปฏิบัติต่อผู้อื่น
13. มีความทะเยอทะยาน แต่ความทะเยอทะยานนี้สอดคล้องกับเป้าหมายการบริหารงานของหน่วยงานที่สังกัด

ประเภทของเครื่องมือวัดที่ใช้กับนักบริหาร จากรายละเอียดของลักษณะนักบริหาร และลักษณะงานของนักบริหารที่ได้ประมวลไว้ข้างบนนี้โดยสังเขป เฉพาะสาระสำคัญ จะเห็นว่ามาตรวัดความสำเร็จของงานนักบริหาร (criterion measure) จะต่างกันไประหว่างนักบริหารในหน่วยงานหนึ่ง ๆ ซึ่งอาจเป็นส่วนราชการหรือธุรกิจเอกชนประเภทต่าง ๆ ยิ่งไปกว่านั้นมาตรวัดความสำเร็จของงานในหน้าที่ระดับผู้บริหารยังอาจวิเคราะห์ได้โดยยาก ทั้งนี้เพราะงานในระดับผู้บริหารมีลักษณะพิเศษเฉพาะ (unique) ไม่เหมือนกับงานระดับอื่น ๆ ตรงที่ว่า ความสำเร็จของการปฏิบัติงานมิใช่ขึ้นอยู่กับการปฏิบัติงานให้เสร็จสิ้นแต่อย่างเดียว แต่ยังขึ้นอยู่กับคนที่ผู้ปฏิบัติงานได้บังคับบัญชาสามารถและมีกำลังใจที่จะปฏิบัติงานให้สำเร็จตามนโยบายอีกด้วย ในการวัดบุคคลระดับผู้บริหาร จึงได้มีการกำหนดแนวคิด (concepts) เพื่อเป็นตัวแปรในการวัดได้หลายอย่าง เช่น กำลังใจร่วมในการปฏิบัติงาน (morale) ความพอใจในงาน (satisfaction) แรงกระตุ้น (motivation) ความสามารถในการใช้กำลังคน (quality of manpower utilization) และ ปริมาณของผลิตผลของงานที่นิเทศ (quantity of the work output of operations supervised) เป็นต้น ทั้งนี้เพื่อหลีกเลี่ยงข้อบกพร่องของวิธีการให้ค่า (rating scale) ซึ่งนิยมใช้กันมาเป็นเวลานานแล้ว ดังได้อธิบายรายละเอียดไว้ในบทที่ 6 โดยเฉพาะในระดับนักบริหารชั้นสูงสุด (top level executives) ได้มีการกำหนดแนวคิดเพื่อเป็นตัวแปรในการวัดที่สำคัญ 2 อย่าง คือ ผลการปฏิบัติงานของทุกคนทั้งหมด ทำให้งานสำเร็จ บรรลุถึงเป้าหมายของหน่วยงานที่รับผิดชอบหรือไม่ และผู้บริหารได้ทำให้ผู้ปฏิบัติงานทุกคนร่วมมือกันมุ่งหน้าทำงานให้ตนเพียงไร อย่างไรก็ตาม การกำหนดแนวคิดดังกล่าวนี้ แม้จะอาศัยหลักเหตุและผลทำให้เป็นที่เชื่อได้ว่า การวัดบุคคลระดับผู้บริหารจะมีลักษณะเป็นปรนัย (subjective) ยิ่งขึ้น แต่การวิเคราะห์มาตรวัดความสำเร็จของงานตามแนวนี้ก็มีความลำบาก ต้องใช้เวลา และจำนวนผู้ปฏิบัติงานในระดับผู้บริหารในหน้าที่ความรับผิดชอบของงานที่เหมือนหรือคล้ายคลึงกันในหน่วยงานหนึ่ง ๆ มิได้มีเป็นจำนวนมากเพียงพอที่จะทำให้คุ้มแก่ค่าใช้จ่ายและเวลาที่ต้องลงทุนในการดำเนินการวิเคราะห์เช่นนั้น แต่ถึงแม้จะมีความลำบากหลายอย่างเช่นนี้ ผลของการวิจัยเครื่องมือวัด

งานบุคคลก็ยังมีหลักฐานชี้ให้เห็นว่า มีแบบทดสอบบางประเภทที่อาจใช้คาดคะเนความสามารถของผู้ที่จะปฏิบัติงานระดับน้อยอย่างได้ผล

Lawshe¹⁴ ได้เผยแพร่ผลการวิจัยในปี ค.ศ. 1949 ผู้วิจัยได้ทดลองใช้แบบทดสอบเชาวน์ Purdue Adaptability Test กับเจ้าหน้าที่ระดับผู้บริหารจำนวน 176 คน ของบริษัทที่สุ่มตัวอย่างแล้ว จากบริษัทธุรกิจต่าง ๆ ทั่วประเทศสหรัฐอเมริกา เป็นจำนวน 44 บริษัท ปรากฏว่าเครื่องมือวัดเชาวน์ชุดนี้สามารถคาดคะเนผู้ที่อาจปฏิบัติงานในหน้าที่นักบริหารได้ ดังแสดงรายละเอียดไว้ในตารางที่ 32

ตารางที่ 32

ผลการวัดเชาวน์กับความสามารถในหน้าที่นักบริหาร

ผลการวัดเชาวน์	อัตราร้อยละที่สามารถปฏิบัติงานนักบริหารได้
20% ที่ได้คะแนนสูงสุด	73
40% ”	66
60% ”	61
80% ”	56

แหล่งที่มา : Lawshe, C.H., and Balma, Michael J., *Principles of Personnel Testing*, p. 214.

ในปี 1955 Ghiselli และ Brown¹⁵ ได้ทดลองการใช้เครื่องมือวัดเชาวน์ในการจัดงานบุคคลระดับผู้บริหาร ผลการวิจัยก็สนับสนุนงานของ Lawshe คือ เครื่องมือวัดเชาวน์ นี้มีความเที่ยงตรงในการวัดบุคคลระดับผู้บริหารมากกว่าเครื่องมือวัดชนิดอื่น ๆ ยิ่งผู้ปฏิบัติงานต้องรับผิดชอบงานบริหารในระดับสูงขึ้นไปเท่าใด ก็ยิ่งต้องใช้สติปัญญาในการคิดและปฏิบัติงานมากขึ้นเท่านั้น ทั้งนี้เพราะนักบริหารต้องสามารถเข้าใจสาระสำคัญของรายงานและปัญหาต่าง ๆ ในเวลารวดเร็ว การที่สามารถเข้าใจรับฟังเรื่องราวต่าง ๆ และสามารถวินิจฉัยสั่งการได้นั้นต้องอาศัยสติปัญญาอย่างสูง สติปัญญานี้ไม่ใช่หมายถึงความสามารถของนักวิชาการในการอ่านตำรับตำรา หรือในการคิดค้นทฤษฎีอื่นใด แต่หมายถึงความสามารถในการใช้ภาษา ในการคิดหาเหตุผล ในการ

¹⁴ Lawshe and Balma, *op. cit.*, p. 211.

¹⁵ Guion, *op. cit.*, p. 471.

วิเคราะห์ในด้านความจำตลอดจนความคล่องตัวในการคิด อันเป็นองค์ประกอบเฉพาะด้านของสติปัญญา

ส่วนเครื่องมือวัดความสัมฤทธิ์ผลและบุคลิกภาพนั้น แม้จะมีส่วนในการคาดคะเนความสามารถของผู้ปฏิบัติงานระดับผู้บริหาร แต่ในปัจจุบันก็ยังปรากฏว่ามีค่าความเที่ยงตรงกับงานระดับนี้ไม่ชัดเจนเหมือนอย่างเครื่องมือวัดเชาวัน ในการปฏิบัติงานของนักบริหารระดับต่ำสุด เช่น หัวหน้าคนงาน หรือหัวหน้างานขั้นต้น (foreman or first-line supervisor) ความรู้ในเนื้อหาของงานมีส่วนช่วยให้สามารถบริหารงานได้¹⁶ แต่เครื่องมือวัดความสัมฤทธิ์ผลที่จะใช้กับบุคคลที่บริหารงานระดับนี้ ก็ควรจะต้องสร้างเป็นพิเศษ เพราะมิได้มุ่งให้เนื้อหาของความรู้เช่นเดียวกับผู้จะต้องปฏิบัติงานโดยตรง ในเรื่องบุคลิกภาพนั้น เป็นสิ่งที่เห็นจริงแล้วตามหลักจิตวิทยาที่กล่าวไว้ในบทที่ 1 ว่า บุคคลที่รับผิดชอบงานระดับผู้บริหาร ก็ต้องอาศัยแรงกระตุ้นทางใจ เช่นเดียวกับผู้ปฏิบัติงานในหน้าที่อื่น ๆ จึงจะสามารถปฏิบัติตนเช่นนักบริหารที่ดีได้ การวัดบุคลิกภาพของผู้ที่อาจเป็นนักบริหารจึงจะต้องดำเนินการให้ครอบคลุมเรื่องต่างๆ ต่อไปนี้ คือ ความมั่นคงในด้านอารมณ์ ความเชื่อมั่นในตนเอง ความทะเยอทะยาน ตลอดจนความสามารถวางตัวในที่สาธารณะ แต่เป็นที่น่าเสียดายว่า ยังไม่มีผลการวิจัยอันใดที่จะพิสูจน์ค่าความคะเนของเครื่องมือวัดบุคลิกภาพในการจัดงานบุคคลระดับผู้บริหารให้ได้เป็นที่พอใจ¹⁷

การประเมินบุคคลระดับผู้บริหารในปัจจุบัน อาศัยการสังเกตและความรู้สึกของผู้บังคับบัญชาชั้นสูงขึ้นไปเป็นหลัก แต่การปฏิบัติเช่นนี้ แม้จะพยายามให้เที่ยงตรงอย่างที่สุด โดยพิจารณาข้อมูลเกี่ยวกับตัวบุคคล เช่น ประวัติการทำงานและการสัมภาษณ์ ก็ยังถือว่าเป็นสิ่งที่ยังขาดคุณภาพในแง่การวัดผลบุคคล เพราะไม่มีความเที่ยงตรงตามหลักวิทยาศาสตร์ ดังนั้นในกรณีของหน่วยงานที่มีขนาดใหญ่พอสมควร ซึ่งทำให้การสร้างเครื่องมือวัดที่ถูกต้องตามหลักเกณฑ์คุ้มกับค่าใช้จ่าย และเวลาในการวิจัย การใช้เครื่องมือวัดเพื่อประกอบการวินิจฉัยสั่งการในเรื่องงานบุคคลระดับนี้ จึงเป็นสิ่งที่ควรส่งเสริมและปฏิบัติ

¹⁶ Lawshe and Balma, *op. cit.*, p. 215.

¹⁷ *Ibid.*, p. 214.

บรรณานุกรม

ภาษาไทย

- สำนักงาน ก.พ., คู่มือการสอบข้อเขียน, โรงพิมพ์คุรุสภา, กรุงเทพฯ ฯ, 2512, 24 หน้า.
- สุคนธ์รังษี, สวัสดิ์ และ จุลละนันท์, สมศรี, ระเบียบวิธีเบื้องต้นในการวิจัย, โรเนียว, สถาบันบัณฑิตพัฒนบริหารศาสตร์, กรุงเทพฯ ฯ, 2511, 171 หน้า.

ภาษาอังกฤษ

- Amrine, Harold T.; Ritchey, John A.; and Hulley, Oliver S., **Manufacturing Organization and Management**, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1966, 568 pp.
- Berdie, Ralph F.; Layton, Wilbur L.; Swanson, Edward O.; and Hagonah, Theda, **Testing in Guidance and Counselling**, McGraw-Hill Book Co., N.Y., 1963, 288 pp.
- Chruden, Herbert J., and Sherman, Jr., Arthur W., **Personnel Management**, South-Western Publishing Co., Cincinnati, 1968, 761 pp.
- Gronbach, Lee J., **Essentials of Psychological Testing**, Harper & Brothers, N.Y., 1949, 475 pp.
- Cronbach, Lee J., and Gleser, Goldine, **Psychological Tests and Personnel Decisions**, University of Illinois Press, Ill., 1965, 347 pp.
- Drucker, Peter F., **The Practice of Management**, Harper & Brothers, N.Y., 1954, 404 pp.
- Ferber, Robert, and Verdoorn, P.J., **Research Methods in Economics and Business**, The Macmillan Co., N.Y., 1962, 573 pp.
- French, Wendell, **The Personnel Management Process**, Houghton Mifflin Co., Boston, 1964, 624 pp.
- Fruchter, Benjamin, **Introduction to Factor Analysis**, New York, D. Van Nostrand Co., 1954, XII + 280 pp.
- Gekoski, Norman, **Psychological Testing**, Charles C. Thomas, Ill., 1964, 300 pp.
- Ghiselli, Edwin E., **Theory of Psychological Measurement**, McGraw-Hill, Inc., N.Y., 1964, 408 pp.
- Good, Carter V., editor, **Dictionary of Education**, McGraw-Hill, Inc., N.Y., 1959, 676 pp.
- Good, Carter V., and Scates, Douglas E., **Methods of Research**, Appleton-Century-Crofts, Inc., N.Y., 1954, 920 pp.
- Guilford, J.P., **Personality**, McGraw-Hill, Inc., N.Y., 1959, 562 pp.
- Guilford, J.P., **Psychometric Methods**, McGraw-Hill, Inc., N.Y., 1954, 597 pp.
- Guion, Robert M., **Personnel Testing**, McGraw-Hill, Inc., N.Y., 1965, 585 pp.
- Hillway, Tyrus, **Introduction to Research**, Houghton Mifflin Co., Boston, 1964, 303 pp.
- Jackson, Douglas N., and Messick, Samuel, eds. **Problems in Human Assessment**, McGraw-Hill Book Co., New York, 1967, 873 pp.

- Jucius, Michael J., **Personnel Management**, Richard D. Irwin, Inc., Homewood, Ill, 1967, 593 pp.
- Kerlinger, Fred N., **Foundations of Behavioral Research**, Holt, Rinehart and Winston, Inc., N.Y., 1964, 739 pp.
- Lawshe, C.H., and Balma, Michael J., **Principles of Personnel Testing**, McGraw-Hill, Inc., 1966, 427 pp.
- Lyman, Howard B., **Test Scores and What They Mean**, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963, 223 pp.
- McFarland, Dalton E., **Personnel Management**, The Macmillan Co., N.Y., 1968, 694 pp.
- Marx, Fritz Morstein, ed., **Elements of Public Administration**, Prentice-Hall, Inc., N.J., 1959, 572 pp.
- Mee, John F., ed., **Personnel Handbook**, The Ronald Press Co., N.Y., 1958, 1165 pp.
- Negro, Felix A., **Public Personnel Administration**, Holt, Rinehart and Winston, Inc., N.Y., 1963, 499 pp.
- Nunnally, Jum C., **Psychometric Theory**, McGraw-Hill, Inc., N.Y., 1967, 640 pp.
- Patton, Arch, "How to Appraise Executive Performance," **Harvard Business Review**, 38: 63-70, January-February, 1960.
- Pfiffner, John M., and Presthus, R. Vance, **Public Administration**, The Ronald Press, Company, New York, 1953, 626 pp.
- Pigors, Paul, and Myers, Charles A., **Personnel Administration**, McGraw-Hill Book Co. Inc., N.Y., 1961, 749 pp.
- Rethlingshafer, Dorothy, **Motivation as Related to Personality**, McGraw-Hill, Inc., N.Y., 1963, 388 pp.
- Remmers, H.H., et. al., **Measurement and Evaluation**, Harpers, & Row, New York, 1965, 390 pp.
- Rummel, J. Francis, and Ballaine, Wesley C., **Research Methodology in Business**, Harper & Row, N.Y., 1963, 359 pp.
- Sands, Edith, **How to Select Executive Personnel**, Reinhold Publishing Corp., New York, 1963, 215 pp.
- Schuessler, Karl F., **Social Research Methods**, Institute of Public Administration, Thammasat University, Bangkok, 1963, 144 pp.
- Shaw, Marvin E., and Wright, Jack M., **Scales for the Measurement of Attitudes**, McGraw-Hill, Inc., N.Y., 1967, 604 pp.
- Simon, Herbert A., **Administrative Behavior**, The Macmillan Co., New York, 1957, 259 pp.
- Stahl, O. Glenn, **Public Personnel Administration**, Harper & Row, N.Y., 1962, 531 pp.
- Thorndike, Robert L., **Personnel Selection**, John Wiley & Sons, Inc., N.Y., 1964, 358 pp.
- Thorndike, Robert L., and Hagen, Elizabeth, **Measurement and Evaluation in Psychology and Education**, John Wiley & Sons, Inc., N.Y., 1957, 575 pp.
- Thurstone, L.L., **The Measurement of Values**, The University of Chicago Press, Chicago, 1967, 322 pp.

- Tiffin, Joseph, "Merit Rating: Its Validity and Techniques", in Doohar, N. Joseph, and Marquis, Vivienne, editors, **Rating Employee and Supervisor Performance**, American Management Association, N.Y., 1950, 192 pp.
- Tiffin, Joseph, and McCormick, Ernest J., **Industrial Psychology**, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1958, 584 pp.
- Travers, Robert M.W., **Educational Measurement**, The Macmillan Co., N.Y., 1955, 420 pp.
- Waldo, Dwight, ed., **Ideas and Issues in Public Administration**, McGraw-Hill Book Company, Inc., New York, 1953, 462 pp.
- Witkin, H.A., et. al., **Psychological Differentiation**, John Wiley and Sons, Inc., New York 418 pp.

พิมพ์ที่ บริษัทโรงพิมพ์ไทยวัฒนาพานิช จำกัด ๘๘๑ ถนนพระราม ๑ กรุงเทพมหานคร
นางบุญพริ้ง ต. สุวรรณ ผู้พิมพ์ผู้โฆษณา พ.ศ. ๒๕๑๗