

สติปัญญาของผู้ปฏิบัติงานกับมาตรวัด ความสำเร็จของงาน

สวัสดิ์ สุภนธรังษี*

ความนำ

ข้อวิพากษ์วิจารณ์การปฏิบัติงานของข้าราชการไทยในปัจจุบัน มีมากมายหลายลักษณะ เช่น การทุจริตต่อหน้าที่ ทำให้ประชาชนได้รับความเดือดร้อน การขาดความกตัญญู และการไม่อุทิศตนให้ในการปฏิบัติหน้าที่ความรับผิดชอบ ทำให้ข้าราชการดำเนินไปด้วยความล่าช้า และการเล่นพรรคเล่นพวกในการบรรจุและเลื่อนขั้นตำแหน่ง ทำให้ข้าราชการที่มีคุณและความสามารถลาออกจากราชการไปประกอบอาชีพอย่างอื่น เป็นต้น ในการวิพากษ์วิจารณ์เรื่องเหล่านี้ก็ได้มีการยกตัวอย่างประกอบ เช่น ในเรื่องทุจริต ก็มีตัวอย่างของประชาชนที่มาขอรับบริการจากทางราชการต้องเสียเงินให้แก่ข้าราชการที่ปฏิบัติหน้าที่เกี่ยวข้อง นอกเหนือไปจากระเบียบที่ทางราชการกำหนดไว้ ในเรื่องไม่อุทิศตนในการปฏิบัติราชการจะเห็นได้จากการมาทำงานสาย และการรีบออกจากสถานที่ทำงานก่อนเวลาเลิกทำงาน ส่วนการเล่นพรรคเล่นพวกแม้จะไม่มีตัวอย่างที่จะชี้แจงไปชัดเจนก็อาจสังเกตได้จากการแสดงความไม่พอใจในการเลื่อนขั้นเลื่อนตำแหน่งในหมู่ข้าราชการด้วยกันเองที่มีความรู้สึกที่ไม่ได้รับความเป็นธรรม ต้นเหตุของข้อบกพร่องทั้งหมดในเรื่องนี้มิได้อยู่ที่ระเบียบปฏิบัติของทางราชการที่บางคนอ้างว่ายังมีช่องโหว่อยู่ หรืออยู่ที่ความยากจน และความเดือดร้อนทางการเงินของข้าราชการ ดังที่มีการอ้างว่าเงินเดือนของข้าราชการนั้นมีระดับต่ำกว่าผู้ปฏิบัติงานในวิสาหกิจเอกชน จริงอยู่ที่อ้างนั้นนับได้ว่าเป็นเหตุที่ทำให้เกิดความบกพร่องในการปฏิบัติงานของข้าราชการไทย แต่ก็มีใช่เป็นต้นเหตุที่แท้จริง เป็นเพียงสาเหตุร่วมที่ช่วยส่งเสริมให้ข้าราชการปฏิบัติตนบกพร่องดังกล่าวเท่านั้น ต้นเหตุที่แท้จริงนั้นคือ ความหละหลวมของผู้บริหารงานซึ่งมีอำนาจในการบังคับบัญชาข้าราชการทั้งหลาย ความหละหลวมนี้เกิดขึ้นในการปฏิบัติงานหลายชั้นหลายตอนของผู้บังคับบัญชา นับตั้งแต่การมอบหมายให้ข้าราชการในบังคับบัญชาปฏิบัติหน้าที่ความรับผิดชอบ

*สวัสดิ์ สุภนธรังษี, Ph.D. ผู้ช่วยศาสตราจารย์ คณะรัฐประศาสนศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

การส่งเสริมให้กำลังใจแก่ข้าราชการที่ปฏิบัติงานในระดับต่าง ๆ ตลอดไปจนถึงการควบคุมติดตามผลงานให้ประสานกันระหว่างข้าราชการทุกฝ่ายที่เกี่ยวข้อง ข้อเท็จจริงในเรื่องนี้จะพิจารณาเห็นได้จากความเหมาะสมของตัวบุคคลในตำแหน่งข้าราชการต่าง ๆ และผลของงานที่ขาดความรัดกุม อันเกิดขึ้นอยู่เสมอ ๆ เมื่อมองลักษณะการภายนอกของการปฏิบัติราชการขณะนี้ปรากฏว่า การแต่งตั้งให้ข้าราชการดำรงตำแหน่งหน้าที่ทั้งหลายโดยทั่วไปในปัจจุบันเป็นไปตามกฎเกณฑ์และระเบียบของทางราชการอย่างหาข้อบกพร่องไม่พบ ส่วนการปฏิบัติงานต่าง ๆ ของทางราชการตามปกตินั้นก็มียางานและการพิมพ์เผยแพร่ผลงานเป็นรายปีโดยหาข้อตำหนิไม่ได้ว่าผิดแบบแผนหรือกฎหมายอันใดบ้าง แต่แท้ที่จริงแล้วปัญหาที่ซ่อนเร้นอยู่ในขณะนี้ก็คือว่า ข้าราชการที่ได้รับการแต่งตั้งให้ดำรงตำแหน่งหน้าที่ต่าง ๆ นั้นเป็นบุคคลที่เหมาะสมที่สุดกับงานในความรับผิดชอบนั้น ๆ หรือไม่ และการปฏิบัติราชการที่เป็นไปตามระเบียบแบบแผนดังเห็นอยู่นี้เป็น การสอดคล้องกับหลักการบริหารงานที่แท้จริง ซึ่งถือว่าผู้บริหารต้องมีหน้าที่เป็นผู้นำของหน่วยงานและต้องสามารถประสานความพยายามและความคิดของบุคคลทุกฝ่ายหรือเปล่า

เป็นที่ทราบกันดีว่า หน้าที่ของผู้บริหารนั้นมีหลายประการซึ่งไม่จำเป็นต้องนำมากล่าวทั้งหมดไว้ในที่นี้ แต่ที่สมควรย้ำ ณ ที่นี้ก็คื การจัดให้บุคคลได้ปฏิบัติงานในหน้าที่ความรับผิดชอบที่เหมาะสมที่สุดกับคุณสมบัติและความสามารถของผู้นั้น นับเป็นหน้าที่สำคัญที่สุดประการหนึ่งของผู้บริหาร การบรรจุแต่งตั้งและเลื่อนขั้นเลื่อนตำแหน่งที่เป็นไปโดยไม่ขัดต่อระเบียบแบบแผนของทางราชการในปัจจุบันไม่ได้เป็นเครื่องประกันว่า ผู้ที่ได้รับการมอบหมายให้ปฏิบัติในตำแหน่งหน้าที่หนึ่งจะเป็นผู้เหมาะสมที่สุดกับตำแหน่งหน้าที่นั้น ๆ ความเหมาะสมที่สุดนี้มีได้หมายแต่เพียงว่ามีคุณสมบัติและความสามารถตรงกับงานที่ต้องรับผิดชอบเท่านั้น แต่ยังไม่หมายความว่า ไม่มีบุคคลอื่นที่จะเหมาะสมกับงานนั้น ๆ อีกด้วย เจตนารมณ์และหลักการของทางราชการในปัจจุบันเป็นที่ทราบกันดีว่า ยึดระบบคุณธรรมเป็นแนว แต่ในทางปฏิบัตินั้นยังเป็นที่น่าสงสัยว่า ผู้บริหารได้ยึดมั่นในอุดมการณ์ของระบบนี้เพียงใด บรรดาข้าราชการและพนักงานที่แสดงความรู้สึกว่าเขาไม่ได้รับความเป็นธรรมมักจะได้รับคำตอบจากผู้บังคับบัญชาว่า ความยุติธรรมที่แท้จริงไม่มีในโลก คำกล่าวเช่นนี้ไม่ใช่ข้ออ้างที่จะนำมาใช้ในการจัดงานบุคคล เพราะถ้ายอมให้เป็นไปตามที่อ้างนี้ได้ ก็แปลว่าผู้บังคับบัญชาไม่ได้ปฏิบัติหน้าที่ผู้บริหารที่ดี ดังได้ยกมาข้างต้น ผลก็คือ ข้าราชการเกิดความไม่สบายในการปฏิบัติหน้าที่ความสามารถในการคิดริเริ่มสร้างสรรค์ (creativity) ต้องเสื่อมถอยลงไป และนับเป็นการทำลายกำลังใจในการปฏิบัติงานร่วมกัน (morale) ของผู้ปฏิบัติงานทุก ๆ ฝ่ายที่เกี่ยวข้อง

ความจริง การจัดให้บุคคลได้ปฏิบัติงานในหน้าที่ความรับผิดชอบที่เหมาะสมที่สุดกับคุณสมบัติและความสามารถของผู้นั้นมิใช่เป็นเรื่องที่ปฏิบัติได้ง่าย เรื่องนี้มีใช่เป็นเรื่องที่เพียงแค่ตั้งความมุ่งหมายหรือตั้งใจไว้แล้วว่าจะปฏิบัติก็จะทำให้ทำได้สำเร็จ การที่หยิบยกปัญหาข้อบกพร่องของผู้บริหารขึ้นมากล่าวในเรื่องนี้มิได้มีความประสงค์จะติเตียนว่า ผู้บังคับบัญชาในหน่วยราชการทั่วไปต่างตั้งใจที่จะทำให้เกิดความเสียหายแก่ทางราชการโดยมอบหมายให้บุคคลที่ไม่ได้เหมาะสมที่สุดปฏิบัติงานในตำแหน่งราชการต่าง ๆ เพราะถึงแม้ว่าผู้บังคับบัญชาจะมีความสุจริตใจในการแต่งตั้งเลื่อนขึ้นเลื่อนตำแหน่งผู้อยู่ใต้บังคับบัญชาของตนก็ยิ่งอาจทำให้เกิดสภาพการณ์ที่มีบุคคลที่ไม่เหมาะสมที่สุดปฏิบัติงานในตำแหน่งงานของทางราชการได้ ความสุจริตใจนี้จะต้องอาศัยความเข้าใจหลักเกณฑ์ที่ถูกต้องในเรื่องการวัดบุคคลประกอบ จึงจะทำให้สามารถจัดงานบุคคลให้เป็นไปตามระบบคุณธรรมได้อย่างแท้จริง

การวัดว่าบุคคลสามารถทำอะไรได้บ้าง เป็นการวัดความสามารถ 2 ประเภทใหญ่ คือ ความถนัดทางธรรมชาติซึ่งได้แก่ 1. ความสามารถที่จะเรียนรู้เพื่อไปปฏิบัติงานในระยะต่อไป และความสัมฤทธิ์ผลซึ่งได้แก่ ผลการเรียนรู้ที่จะนำไปใช้ในการปฏิบัติ เฉพาะในเรื่องความถนัดทางธรรมชาตินี้ยังแยกออกไปได้อีก 2 อย่างคือ ความสามารถทางสมอง (intellectual ability) และความถนัดพิเศษทางธรรมชาติ (special aptitude) เพื่อที่จะให้ผู้บริหารเกิดแนวคิดในการจัดบุคคลให้เหมาะสมกับงานโดยละเอียดตามหลักเกณฑ์ที่ถูกต้อง ในที่นี้จะกล่าวแต่เฉพาะความสามารถทางสมองเท่านั้น เพราะถึงแม้ความสามารถด้านอื่นจะมีลักษณะแตกต่างออกไป แต่แนวคิดในการวิเคราะห์ก็เป็นไปโดยหลักของเหตุและผลในเส้นทางเดียวกัน

เมื่อกล่าวถึง ความสามารถทางสมอง อาจทำให้นึกถึงศัพท์อื่น ๆ ที่มีความหมายในลักษณะเดียวกับคำนี้ได้หลายคำ เช่น เซอว์น (intellect) สติปัญญา (intelligence) ความเฉลียวฉลาด (sagacity) ความมีสมองดี (brains) เป็นต้น เซอว์นทำให้นึกถึงความสามารถในการคิดหาเหตุผลและเข้าใจสิ่งต่าง ๆ โดยฉับพลัน สติปัญญามักหมายถึงความสามารถในการศึกษาหาความรู้ ความเฉลียวฉลาดทำให้เกิดความรู้สึกว่าเป็นเรื่องความสามารถในการหยั่งรู้และการมองเห็นการณ์ไกล ในการใช้ดุลยพินิจตัดสินเรื่องต่าง ๆ ส่วนความมีสมองดีอาจทำให้นึกถึงความสามารถทั่ว ๆ ไปที่ต้องใช้ความคิดและความเข้าใจซึ่งเป็นศักยภาพของสมอง ศัพท์ทั้ง 4 คำที่ยกมาเป็นตัวอย่างนี้มีความแพร่หลายในการใช้ในภาษาเขียนไม่เท่ากัน ศัพท์ที่ใช้แพร่หลายมากกว่าคำอื่น ๆ ก็คือ เซอว์นหรือสติปัญญา โดยเฉพาะในระเบียบและกฎข้อบังคับ

ในการบริหารงานบุคคลในประเทศนิยมใช้คำว่า เซวน์ เป็นหลัก ส่วนในต่างประเทศใช้คำ intelligence เป็นหลัก ในทางวิชาการจัดงานบุคคลซึ่งบัญญัติศัพท์ใหม่ว่า ความสามารถทางสมอง (intellectual ability) เพื่อให้มีความหมายครอบคลุมความสามารถด้านนี้ทั้งหมด และนิยามไว้สั้น ๆ ว่าเป็นความสามารถทั่วไปในการเรียนรู้ ในที่นี้จึงจะถือเอาศัพท์ความสามารถทางสมองเป็นหลัก แต่ก็ใช้คำว่าเซวน์ซึ่งแพร่หลายในภาษาไทยสลับกันไปด้วย โดยให้ความหมายแทนความสามารถทางสมอง เพื่อให้ผู้อ่านได้เกิดมโนภาพ เพราะเป็นศัพท์กระต๊อกรักและรู้จักกันกว้างขวางอยู่แล้ว

ความเป็นมาของทฤษฎีการวัดความสามารถทางสมอง

การที่จะเข้าใจพื้นฐานของทฤษฎีการวัดความสามารถทางสมอง ควรทราบประวัติการวัดผลบุคคลโดยสังเขปเสียก่อน การวัดผลบุคคลเป็นวิชาการที่เกิดหลังจากคณิตศาสตร์และจิตมิติได้มีความเจริญก้าวหน้าขึ้น วิชาคณิตศาสตร์ที่วัดด้วยความน่าจะเป็นได้เริ่มศึกษากันในคริสต์ศตวรรษที่ 17 Bernoulli เป็นบุคคลแรกที่พิมพ์เอกสารทางวิชาการเกี่ยวกับเรื่องนี้ ในศตวรรษที่ 18 De Moivre ได้คิดทฤษฎีการแจกแจงปกติ Gauss เป็นบุคคลสำคัญในศตวรรษที่ 19 ที่ได้คิดเส้นโค้งปกติขึ้น บุคคลสำคัญคนแรกที่ได้นำเอาหลักคณิตศาสตร์มาประยุกต์กับจิตวิทยาคือ Sir Francis Galton Karl Pearson ได้มีส่วนทำให้งานของ Galton สมบูรณ์ขึ้น โดยคิดวิธีหาสัมประสิทธิ์สหสัมพันธ์และค่าคะแนนมาตรฐาน บุคคลสำคัญในศตวรรษที่ 20 หลายคนได้ร่วมกันคิดค้นวิธีการวัดให้เป็นที่เชื่อถือได้ยิ่งขึ้น เช่น James M. Cattell, E. L. Thorndike, Alfred Binet และ Lewis Terman แต่การค้นคว้าในเรื่องนี้เป็นไปเพื่อประโยชน์ในแง่จิตวิทยาโดยตรงเป็นประการสำคัญ ต่อเมื่อเกิดสงครามโลกขึ้นจึงได้นำความรู้นี้มาประยุกต์กับการบริหารงานบุคคล ในสมัยสงครามโลกครั้งที่หนึ่ง Binet และ Cattell ได้นำวิธีการวัดนี้มาทดลองใช้ในการคัดเลือกแยกประเภทบุคคลที่จะเข้ารับราชการทหารในสมัยสงครามโลกครั้งที่สอง กองทัพสหรัฐอเมริกาได้สนับสนุนให้มีการวิจัยการวัดบุคคลอย่างกว้างขวาง และได้สร้างเครื่องมือวัดที่แพร่หลายขึ้นเพื่อใช้ในการจัดงานบุคคล เช่น Army Alpha และ Army Beta นอกจากนี้ สถาบันอุดมศึกษาอเมริกันก็ยังได้มีบทบาทสำคัญในการวิเคราะห์ค้นคว้าเรื่องการวัดบุคคลอย่างกว้างขวางอีกด้วย

เฉพาะในด้านกรวัดความสามารถทางสมองนั้น Galton เป็นผู้ริเริ่มศึกษาหลักเกณฑ์การวัดในเรื่องนี้ Galton เชื่อว่าความฉลาดและบุคลิกภาพเป็นเรื่องของกรรมพันธุ์ จึงได้

พยายามสร้างเครื่องมีวัดความฉลาดของบุคคล และได้คิดศัพท์เรียกเครื่องมีนี้ว่า แบบทดสอบ
 ความฉลาด (mental test) วิธีวัดความฉลาดตามแนวคิดของ Galton ทำให้โดยวัดความ
 สามารถทางประสาท (sensory discrimination) เพราะถือตามแนวคิดของ Locke ว่า บุคคล
 ที่มีความสามารถทางประสาทสูง เป็นผู้ที่มีพรสวรรค์ดีและสามารถเรียนรู้ได้ดี ความสามารถทาง
 ทางประสาทที่วัดได้แก่ ความสามารถในการฟังระดับเสียง ความไวของไขว้ของจักขุ ความสามารถในการ
 การได้ยิน เป็นต้น¹ ผลงานของ Galton ก็ได้ประมวลวิเคราะห์ในเชิงสถิติทำให้ยืนยันได้ว่า
 บุคคลแต่ละคนมีความแตกต่างกัน แต่ก็มีได้กำหนดแบบมาตรฐานอันใดในการวัดความฉลาด
 ของบุคคล ฉะนั้นในสหรัฐฯ จึงได้มีผู้นำเอาสถิติของ Galton ไปทำการวิจัยความฉลาดของนัก
 ศึกษาวิทยาลัย Columbia University ผู้นี้คือ James McKeen Cattell² ในระหว่างปี 1890-69
 Cattell ได้ทดลองแบบทดสอบความฉลาดที่วัดความสามารถทางประสาทในด้านต่าง ๆ เช่น
 ช่วงระยะเวลาที่บุคคลมีปฏิกิริยาโต้ตอบ เวลาที่ใช้ในการตัดสินใจ ความสามารถในการฟังเสียง
 สูงต่ำ การพิจารณาความหนักของสิ่งของโดยไม่ตองชั่ง ความสามารถจำอักษรได้ เป็นต้น
 แต่ผลปรากฏว่าแบบทดสอบเหล่านี้มีสัมประสิทธิ์สหสัมพันธ์ระหว่างกันต่ำมาก และยังไม่มีความ
 สัมพันธ์กับผลการศึกษาศึกษาบุคคลอีกด้วย

อย่างไรก็ดีในปี 1904 Charles Spearman ชาวอังกฤษก็ได้ชี้แจงให้เห็นว่าการที่
 สัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบความสามารถทางประสาทต่ำก็เพราะแบบทดสอบ
 แต่ละชุดขาดความเชื่อมั่น หรือมีสัมประสิทธิ์ความเชื่อมั่นต่ำ เมื่อหักค่าผิดพลาดของการวัด
 (Standard error of measurement) ผลของการวัดด้วยเครื่องมือวัดความสามารถทางประสาท
 ชุดต่าง ๆ จะมีค่าความแปรปรวนร่วมกันในระดับค่อนข้างสูง ทั้งยังมีความสัมพันธ์กับผลการ
 ศึกษาอีกด้วยซึ่งทำให้ Spearman เชื่อว่าในเรื่องความสามารถทางสมองจะต้องมีอะไรอย่างหนึ่ง
 ที่เป็นตัวกลาง จึงได้ทำการทดลองวัดผลการศึกษาในวิชาสาขาต่าง ๆ ของบุคคล เช่น ภาษาฝรั่งเศส
 รัสเซีย ภาษาอังกฤษ คณิตศาสตร์ ดนตรี ถ้าปรากฏว่า ผลการศึกษาวិชาที่ไม่เหมือนกันเหล่านี้มีสิ่ง
 เป็นตัวกลางจริง ผลการวิจัยนี้สรุปได้ว่า เครื่องมือวัดความฉลาดของบุคคลจะมีตัวกลางที่เป็น
 เครื่องแสดงความสามารถทั่วไป (general ability) กำกับเป็นหลักอยู่ และจะมีลักษณะบาง

¹ Jum C. Nunnally, *Psychometric Theory* (New York : McGraw-Hill, Inc, 1967), pp.428-429.

² Douglas N. Jackson, and Samuel Messick, eds., *Problems in Human Assessment* (New York: McGraw-Hill, Inc, 1967) ,pp. 413-414

อย่างที่มีคุณสมบัติพิเศษเฉพาะเครื่องมือวัดแต่ละอันประกอบอยู่ด้วย ลักษณะพิเศษนี้เรียกว่า ความสามารถพิเศษเฉพาะ แนวคิดนี้คือหลักของทฤษฎีองค์ประกอบสองตัวของ Spearman

ในระยะที่ Spearman พยายามค้นคว้าทฤษฎีการวัดบุคคล Alfred Binet ชาวฝรั่งเศส ก็ได้สร้างเครื่องมือวัดสติปัญญาสำเร็จ Binet ได้ร่วมมือกับ Theodore Simon สร้างแบบทดสอบซึ่งประกอบด้วยปัญหา 30 รายการเสร็จในปี 1905 ตามแนวคิดนี้ การวัดความสามารถทางสมองของบุคคลนั้นไม่จำเป็นต้องวัด “ส่วนต่างๆ” (elements) ของความสามารถ แต่ควร จะวัดผลการแสดงความสามารถซึ่งอาจเรียกว่า สติปัญญา (intelligence) คำถามในแบบทดสอบนี้จึงเกี่ยวกับความคิด ความเข้าใจ และการตัดสินใจ ในเรื่องต่างๆ ภายในสังคมเช่น ความสามารถในการอธิบายลักษณะของ บ้าน พ่อแม่ ชื่อนช่อม และม้า หรือความสามารถในการ พูดย่อประโยคที่มีคำถึง 15 คำเมื่อได้ยินเพียงครั้งเดียว ในปี 1908 Binet ได้ขยายแบบ ทดสอบให้มีความยากง่ายเป็นหลายระดับโดยให้สอดคล้องกับเกณฑ์เฉลี่ยของผู้มีอายุต่าง ๆ กันไป Lewis M. Terman และ Maude A. Merrill ชาวอเมริกันได้นำแบบทดสอบของ Binet มา แปลเป็นภาษาอังกฤษ และปรับปรุงอีกหลายครั้ง จนแบบทดสอบนี้เป็นที่แพร่หลายและเป็น แม่แบบของเครื่องมือวัดความสามารถทางสมองในปัจจุบัน กล่าวคือได้มีการแก้ไขในปี 1916 ปี 1937 และ 1960 และแบ่งสติปัญญาของบุคคลออกเป็น 20 ระดับ ใช้ทดสอบกับเด็ก ตั้งแต่อายุ 2 ขวบไปจนถึงผู้ใหญ่ แบบทดสอบที่ใช้กับเด็กเป็นคำถามง่าย ๆ ใช้รูปภาพและ วัตถุเป็นหลักฐานที่ใช้กับผู้ใหญ่เป็นคำถามเกี่ยวกับแนวความคิดและความหมายต่างๆ ในภาษา สารสำคัญของเนื้อหาในแบบทดสอบมีตัวอย่างข้างล่าง³

ระดับอายุสองขวบครึ่ง :

1. การรู้จักสิ่งของเครื่องใช้ (มีสิ่งของ 6 อย่าง)
“จงชี้ว่าอันไหนสำหรับต้ม” ฯลฯ
2. อวัยวะ และส่วนต่าง ๆ ของร่างกาย (มีตุ๊กตากระดาษตัวใหญ่)
“ชี้ว่าผมอยู่ตรงไหน” ฯลฯ

³ Robert L. Thorndike, and Elizabeth Hagen, *Measurement and Evaluation in Psychology and Education* (New York: John Wiley & Sons, Inc, 1957), pp.210-211.
Nunnally, *op. cit.*, p. 430.

3. การเรียกชื่อ (ของเล่น 4 อย่างเช่น แก้ว รัถยนต์ เป็นต้น)
“สิ่งนี้เรียกที่ชื่อว่าอะไร” ฯลฯ
4. คำศัพท์ (มีรูปภาพ 18 รูป เป็นรูปสิ่งต่างๆ ที่รู้จักกันทั่วไป)
“รูปนี้ รูปอะไร” ฯลฯ
5. พุทธชาเลขสองตัว
“จงพูดตาม เลข 4 เลข 7” ฯลฯ
6. การจับของใส่ช่อง (มีแผ่นกระดาษที่เจาะช่องหลาย ๆ ช่องเป็นสี่เหลี่ยม วงกลม และสามเหลี่ยม พร้อมทั้งแท่งไม้ที่จะใส่ไปในช่องเหล่านี้ได้พอดี)
“ให้โยนแท่งไม้ใส่ช่อง”

ระดับผู้ใหญ่ :

1. คำศัพท์ (มีคำศัพท์ 45 คำ เรียกตามความยากง่าย)
“ให้อธิบายความหมายของศัพท์ว่า ขวดหาว (flaunt)” ฯลฯ
2. การหาเหตุผล (2 ภาค)
“ให้เหตุผล 3 ประการว่า ทำไมบุคคลที่ก่ออาชญากรรมอย่างร้ายแรงควรได้รับการลงโทษ”
3. พุทธชาเลขแปดตัว
4. ความหมายของสุภาษิต
“จงอธิบายความหมายของประโยคว่า สิบเบี้ยใกล้มือ (bird in the hand)”
5. ลักษณะที่คล้ายกันระหว่างสิ่งที่ต่างกัน (6 ภาค)
“ความหนักและความเบาเหมือนกันอย่างไร” ฯลฯ
6. พุทธข้อความซ้ำ
เมื่ออ่านข้อความต่อไปนี้จบแล้ว จงพูดซ้ำให้มากที่สุดเท่าที่จะมากได้”
ผู้ที่ทำแบบทดสอบได้ในระดับอายุของตนหรือระดับอื่นๆ ก็สามารถคำนวณสติปัญญาได้ตามวิธี William Stern เช่นสมมติว่า 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

นำจำนวนนี้มากำหนดค่าเฉลี่ยตามสูตร :

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

โดย I.Q. = อัตราร้อยละของเซวาร์
M.A. = จำนวนปีของสติปัญญาคิดเป็นเดือน
C.A. = จำนวนปีของอายุจริงคิดเป็นเดือน
∴ I.Q. = $\frac{80}{240} (100)$
= 33.34

การที่นาย ก. มี I.Q. เพียงเท่านี้หมายความว่า นาย ก. มีระดับเซวาร์ต่ำ
คะแนน I.Q. ที่จะได้จากแบบทดสอบ Binet นี้ อาจแยกตามจำนวนผู้ที่ควรจะได้ ตาม
อัตราดังต่อไปนี้

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
ต่ำกว่า 60	0.6	0.6

เครื่องมือวัดของ Binet สามารถแสดงผลของการมีความสามารถทางสมอง เช่น ความชำนาญด้านภาษา (language mastery) ความสามารถในการหาเหตุผล (reasoning) และสามารถในการประมวลข้อเท็จจริง (assimilation of facts) เป็นต้น แต่แนววิเคราะห์ของ Binet มิได้ช่วยให้นักวัดบุคคลทราบว่าเขาวินหรือความสามารถทางสมองนี้ ประกอบด้วยอะไรบ้าง ส่วนทฤษฎีองค์ประกอบสองตัวของ Spearman นั้น แม้มิได้ช่วยให้ได้เครื่องมือวัดเขาวิน แต่ก็ช่วยปูพื้นฐานในการวิเคราะห์เขาวินอันจะเป็นทางให้สามารถสร้างเครื่องมือวัดที่สมบูรณ์อย่างแท้จริงต่อไป

สหรัฐอเมริกาได้รับแนวความคิดในเรื่องทฤษฎีองค์ประกอบจากอังกฤษ และวิธีการวัดผลที่เกิดจากความสามารถทางสมองของ ฟรังเศสมาดำเนินการวิจัยต่ออย่างกว้างขวาง บุคคลสำคัญชาวอเมริกันคือ L.L. Thurstone ได้ค้นคว้าทฤษฎีกลุ่มขององค์ประกอบในปี 1938 ซึ่งค้นทฤษฎีองค์ประกอบสองตัวว่าหายากเกินไป การค้นทฤษฎีเดิมอาจทดลองได้โดยแบบทดสอบที่ใช้ภาษา (verbal test) 3 ชุด กับบุคคล 2 กลุ่ม สัมประสิทธิ์สหสัมพันธ์ระหว่างชุดภายในกลุ่มจะสูงมากกว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างกลุ่ม เช่น สัมประสิทธิ์สหสัมพันธ์ระหว่างกลุ่มมีค่า 0.30 สัมประสิทธิ์สหสัมพันธ์ภายในกลุ่มอาจสูงถึง 0.70 การที่เป็นเช่นนี้จะถือว่ามีองค์ประกอบร่วมที่มีผลกระทบต่อเครื่องมื่อวัดทุกชุดอยู่หนึ่งตัว และมีองค์ประกอบพิเศษอีก 6 ตัว ย่อมไม่เป็นเหตุผล Thurstone ได้พิสูจน์ว่าความสามารถของบุคคลอาจแยกเป็นองค์ประกอบได้ถึง 9 ตัว ตั้งแต่นั้นมา ก็ได้มีการค้นคว้าองค์ประกอบของเขาวินมาเรื่อยๆ ในปัจจุบันได้มีการค้นพบองค์ประกอบเป็นร้อย ๆ ตัว⁴ จากการวิเคราะห์องค์ประกอบนี้จึงทำให้เกิดทฤษฎีที่ยึดถือในปัจจุบันอย่างแพร่หลายคือ ทฤษฎีโครงสร้างของเขาวิน (structure of intellect) ของ J.P. Guilford ชาวอเมริกัน

ทฤษฎีโครงสร้างของเขาวิน

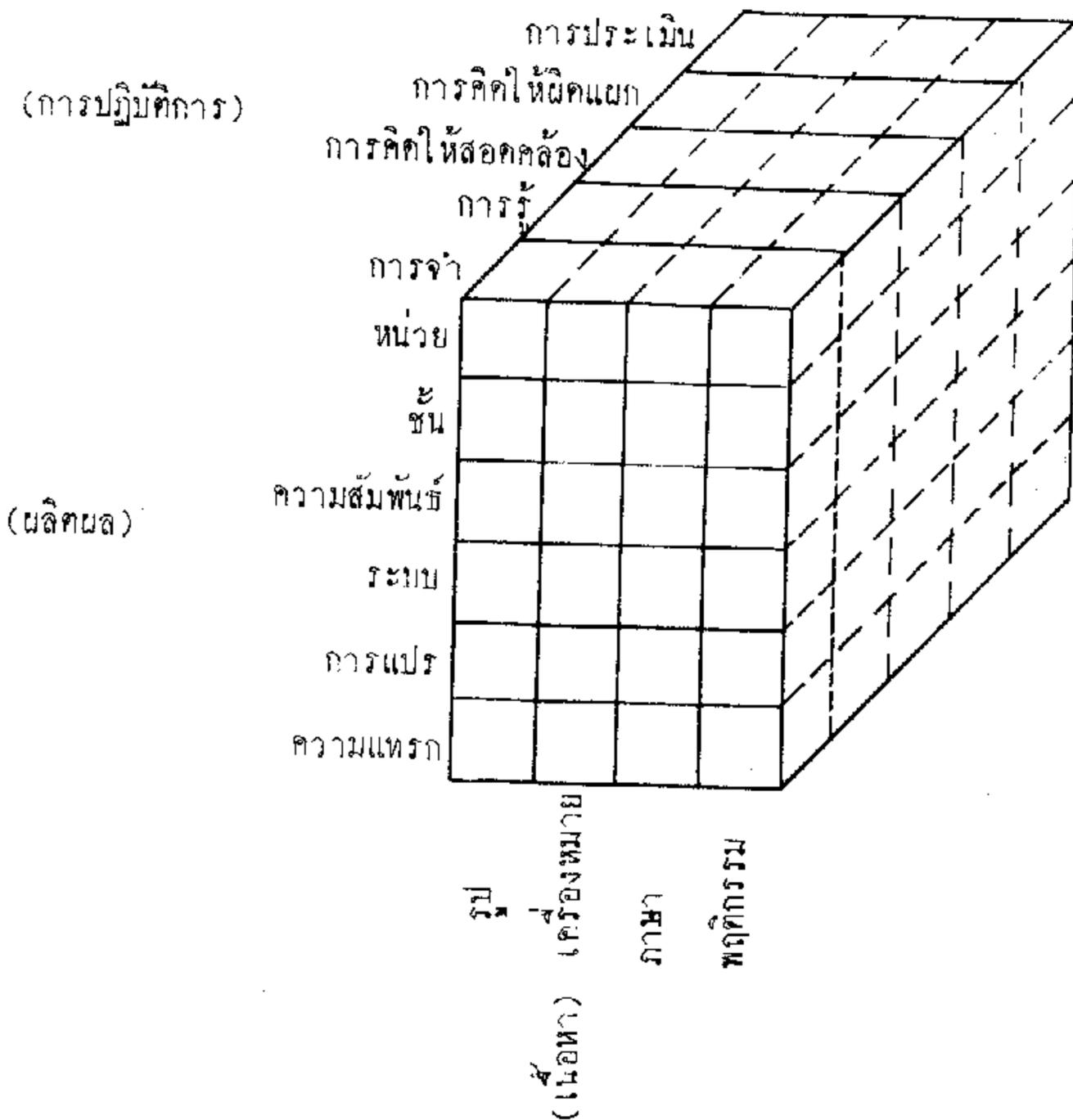
Guilford ได้เสนอทฤษฎีโครงสร้างของเขาวินโดยได้ไปปาฐกถาเรื่องนี้ที่มหาวิทยาลัย Stanford University เมื่อวันที่ 13 เมษายน ค.ศ. 1959 และได้พิมพ์เผยแพร่ทฤษฎีนี้ในวารสาร American Psychologist ในปีเดียวกัน และได้อธิบายความละเอียดของทฤษฎีอีกในวารสาร Psychological Review ในปี 1961 สารสำคัญของทฤษฎีนี้มีดังต่อไปนี้⁵

เขาวินมีสามมิติ ในการวิเคราะห์เขาวิน อาจพิจารณาได้เป็น 3 แง่คือ ในแง่ของสิ่งมาแรงเร้าให้เกิดปฏิบัติการได้ตอบซึ่งเรียกได้ว่าเป็นเนื้อหา (content) ของสิ่งต่างๆ ในสังคมเป็น

⁴ *Ibid.*, p.442.

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

ประการแรก ประการที่สองในแง่ลักษณะของขอบเขตการการแสดงผลออกซึ่งเรียกว่า การปฏิบัติการ (operations) ของผู้มีเชาวน์ และประการสุดท้ายในแง่ของผลที่เกิดจากการแสดงผลออก ซึ่ง เรียกว่า ผลิตภัณฑ์ (product) ดังแสดงไว้ในแผนภาพที่ 1



แผนภาพที่ 1 โครงสร้างของเชาวน์ที่มีสามมิติ
แหล่งที่มา : J.P. Guilford, "Three Faces of Intellect," in Douglas N. Jackson and Samuel Messick, eds., *Problems in Human Rssessment*, p.424.

เมื่อพิจารณาเชาวน์ในแง่เนื้อหา ก็อาจนำสิ่งต่าง ๆ ที่จะมีผลกระทบต่อความรู้สึกของบุคคลมาแยกประเภท ทั้งนี้แยกได้เป็น 4 ประเภทคือ เนื้อหาที่เป็นรูป (figural content) หมายถึงวัตถุสิ่งของต่าง ๆ ที่ประสาทจะรับรู้ได้ ได้แก่สิ่งที่มองเห็นหรือได้ยินหรือรู้สึกสัมผัสได้ เนื้อหาที่เป็นเครื่องหมาย (symbolic content) หมายถึงเครื่องหมายต่าง ๆ ที่ไม่มีความหมายในตัวเอง

แต่มีความหมายแทนสิ่งต่าง ๆ เช่น ตัวเลข ตัวอักษร ซึ่งจัดไว้เป็นระบบแล้ว เนื้อหาที่เป็นภาษา (semantic content) หมายถึง ข้อความหรือข่าวสารที่แสดงแนวคิดให้เป็นที่เข้าใจระหว่างผู้พูดหรือผู้เขียนกับผู้ฟังหรือผู้อ่าน เนื้อหาที่เป็นพฤติกรรม (behavioral content) ได้แก่ สภาพการณ์ทางสังคมซึ่งบุคคลได้รับข่าวสารจากพฤติกรรมของผู้อื่น เช่น การเข้าใจ หรือการแปลความหมายหรือแนวคิดจากการประพฤติปฏิบัติของบุคคลต่าง ๆ ในเชิงปฏิบัติ เนื้อหาต่าง ๆ เหล่านี้สะท้อนอยู่ในสาระสำคัญของเครื่องมีวัดหรือแบบทดสอบที่ใช้

การพิจารณาหาว่าในแง่การปฏิบัติการ หมายความว่าบุคคลแต่ละคนได้มีการดำเนินการอย่างไรในการแสดงความสามารถทางสมอง ขบวนการที่เห็นได้ชัดก็คือ การจำ (memory) และการคิด (thinking) การจำหมายถึง การเก็บรักษาข่าวสารไว้ในตัว ส่วนการคิดเป็นการนำข่าวสารมาใช้หรือปฏิบัติการอย่างหนึ่งอย่างใด ในการทำ การคิดนั้น ยังได้แยกเป็นอาการ 3 อย่างใหญ่ ๆ คือ การรู้ (cognition) การผลิต (production) และการประเมิน (evaluation) การรู้ คือ อาการที่บุคคลรู้จักหรือรับรู้แนวความคิดของสิ่งต่าง ๆ การรู้ทำให้บุคคลเข้าใจสภาพการณ์ต่าง ๆ ของสังคม การผลิต เป็นอาการที่บุคคลนำเอาสิ่งที่รู้และเข้าใจแล้วมาติดต่อเพื่อให้เกิดความรู้แตกแขนงออกไป อาการนี้แยกได้อีกเป็น 2 อย่างคือ การคิดให้สอดคล้อง (convergent production of ideas) และการคิดให้ผิตแผก (divergent production of ideas) การคิดให้สอดคล้องเป็นการคิดตามข่าวสารที่มีเพื่อให้หาผลสรุปคำตอบที่ดีที่สุดที่สอดคล้องตามแนวคิดที่เชื่อถือกันอยู่ในวงกรหนึ่ง ๆ หรือสังคมหนึ่ง ๆ ส่วนการคิดให้ผิตแผกเป็นการคิดในทิศทางที่ต่างไปจากแนวคิดที่ยึดถือกันอยู่ เพื่อให้สามารถค้นพบสิ่งแปลก ๆ ไปจากเดิม ส่วนการประเมินนั้นเป็นอาการที่แต่ละบุคคลใช้ดุลยพินิจของตนลงความเห็นว่าสิ่งหนึ่ง ๆ หรือเรื่องหนึ่ง ๆ มีความเพียงพอ หรือถูกต้อง หรือดี แล้วอย่างไร

ในการพิจารณาหาว่าในแง่ผลิตผล เป็นการพิจารณาว่า เมื่อมีเนื้อหาอย่างหนึ่ง และมีการปฏิบัติการอย่างหนึ่งแล้ว จะเกิดผลอย่างไร ไม่ว่าจะใช้เนื้อหาประเภทใด และมีการปฏิบัติการอย่างใดก็ตาม ผลที่เกิดจะอยู่ใน 6 ประเภทดังต่อไปนี้ คือ ในประเภทขั้นต่ำที่สุด บุคคลจะเข้าใจหรือจำได้ว่ามีแนวความคิดตามจุดต่าง ๆ อย่างไร ผลิตผลที่ในขั้นนี้เรียกว่าเป็นหน่วย (units) ถ้าบุคคลมีความสามารถทางสมองสูงกว่าระดับนี้ ก็จะนำเอาความคิดที่เข้าใจหรือจำได้ในประเด็นต่าง ๆ นั้นมาจัดเป็นกลุ่มตามลักษณะที่คล้ายกันหรือต่างกัน เรียกผลิตผลนี้ได้ว่าขั้น (classer) ถ้ามีความสามารถสูงขึ้นไป ก็อาจจัดได้ว่า ความคิดเหล่านั้นมีความสัมพันธ์ (relations) กันอย่างไรโดยพิจารณาจากลักษณะของตัวแปรที่มีอิทธิพลต่อแนวความคิดเหล่านั้น ยิ่งถ้าหากมีความสามารถสูงขึ้นไปอีก ก็อาจแยกแยะความสลับซับซ้อนของความคิดทั้งหมดให้เป็นระบบ (Systems) โครงสร้างของความคิดเหล่านั้น ผู้มีความสามารถสูงกว่านี้ก็จะนำระบบทั้งหมดมา

เปลี่ยนแปลงให้มีลักษณะที่จะเป็นประโยชน์ในการพิจารณาหรือวิเคราะห์เรื่องอื่น ๆ ได้อีก อา- การในขั้นนี้เรียกว่า การแปร (transformations) ในขั้นสุดท้าย บุคคลที่มีความสามารถทาง สมอง อาจพิจารณาไปจนเห็นความแทรก (implications) ของแนวความคิดที่ได้รับจากข่าวสาร ที่มีอยู่ โดยพิจารณาตัวแปรทั้งหมดในพิสัยที่ทราบ แล้วทำการคาดคะเน หรือหยั่งถึงผลที่อาจ เกิดขึ้นได้ (extrapolation)

ตารางองค์ประกอบของเชาวัน เมื่อพิจารณาเทียบลักษณะต่าง ๆ ของเชาวันตามแผน ภาพที่ 1 จะเห็นได้ว่า การวัดเชาวันตามทฤษฎีนี้จะสามารถแยกองค์ประกอบออกไปได้ถึง 120 ตัว แต่ละช่องสี่เหลี่ยมเล็ก ๆ ในแผนภาพนี้แทนองค์ประกอบหนึ่งตัวซึ่งค่ารวมของเนื้อหา การ ปฏิบัติการ และผลผลิตของเชาวัน

ตัวอย่างเช่น หากมีการปฏิบัติกรในเรื่องการรู้ (cognition) จากเนื้อหาประเภท รูป เครื่องหมาย และภาษา โดยมีผลผลิตในทุกประเภท จะได้องค์ประกอบที่แสดงเชาวัน ดังแสดง ไว้ในตารางที่ 1

ตารางที่ 1
องค์ประกอบที่เกี่ยวกับการรู้ของเชาวัน

ผลผลิต	เนื้อหา		
	รูป	เครื่องหมาย	ภาษา
หน่วย ชั้น	การรู้ทางประสาท การแยกประเภทของ รูป	การรู้เครื่องหมาย การรู้จักกลุ่มต่าง ๆ ของเครื่องหมาย	การเข้าใจภาษา การแยกประเภทแนว คิด
ความสัมพันธ์	การสรุปความสัมพันธ์ ของรูป	การสรุปความสัมพันธ์ ของเครื่องหมาย	การสรุปหาเหตุผล ทั่วไป
ระบบ	การเข้าใจช่วง	การสรุประบบของ เครื่องหมาย	การคิดหาเหตุผล ทั่วไป
การแปร ความแทรก	การวาดภาพในใจ การวาดภาพใน ระยะต่อไป	— การรู้ความแทรกของ ของเครื่องหมาย	ความหยั่งรู้ การมีมโนภาพทาง แนวคิด

แหล่งที่มา : Robert M. Guion, *Personnel Testing*, (New York: McGraw-Hill, Inc., 1965), p. 70

ในตารางที่ 1 แสดงองค์ประกอบไว้ 17 ตัว ความจริงควรจะตั้งมีตัวที่ 18 คือ องค์ประกอบอันเป็นคำรวมระหว่างเครื่องหมายกับ การแปร การที่เว้นว่างไว้ก็เพราะยังไม่สามารถตั้งศัพท์เรียกชื่อองค์ประกอบที่ควรจะมีส่วนได้

องค์ประกอบในตารางนี้มีความหมายดังต่อไปนี้ คือ

1. การรู้ทางประสาท (visual cognition and auditory cognition) ความสามารถที่จะรู้จักรูปที่เคยเห็นในสภาพการณ์ที่ลำบากต่อจักขุประสาท และการจำเสียงแบบต่าง ๆ ได้
2. การรู้เครื่องหมาย (symbolic cognition) ความสามารถที่จะรู้จักหน่วยของเครื่องหมาย เช่น คำแต่ละคำ
3. การเข้าใจภาษา (verbal comprehension) ความรู้ในความหมายของคำ
4. การแยกประเภทของรูป (figural classification) ความสามารถที่จะชี้ว่ารูปแต่ละรูปมีลักษณะทำให้ต่างประเภทกันอย่างไร
5. การรู้จักกลุ่มต่าง ๆ ของเครื่องหมาย (cognition of symbolic classes) ความสามารถที่จะชี้ว่าเครื่องหมายแต่ละอันมีลักษณะทำให้ต่างประเภทกันอย่างไร
6. การแยกประเภทแนวคิด (conceptual classification) ความสามารถที่จะชี้ว่าแนวคิดที่ใช้ในภาษาแต่ละตอนมีลักษณะทำให้ต่างประเภทกันอย่างไร
7. การสรุปความสัมพันธ์ของรูป (eduction of figural relations) ความสามารถที่ทำให้ค้นพบว่ารูปต่าง ๆ เกี่ยวโยงกันอย่างไร
8. การสรุปความสัมพันธ์ของเครื่องหมาย (eduction symbolic relations) ความสามารถที่ทำให้ค้นพบว่าเครื่องหมายต่าง ๆ เกี่ยวโยงกันอย่างไร
9. การสรุปความสัมพันธ์ของภาษา (eduction of semantic relations) ความสามารถที่ทำให้ค้นพบว่าแนวคิดต่าง ๆ ในภาษาเกี่ยวโยงกันอย่างไร
10. การเข้าใจช่วง (spatial orientation) การรู้จักและเข้าใจช่วงหรือมิติต่าง ๆ โดยเอาตนเองเป็นจุดตั้งต้น
11. การสรุประบบของเครื่องหมาย (eduction of symbolic patterns) ความสามารถที่จะค้นพบระบบหรือแบบแผนทางเรขาคณิต

12. การคิดหาเหตุผลทั่วไป (general reasoning) ความสามารถเข้าใจปัญหาหนึ่ง ๆ ก่อนที่จะแก้ปัญหาต่อไป

13. การวาดภาพในใจ (visualization) ความสามารถในรูปภาพของสิ่งต่าง ๆ

14. ความหยั่งรู้ (penetration) ความสามารถเข้าใจสาระสำคัญของแนวความคิดแต่ละอันได้อย่างลึกซึ้ง

15. การวาดภาพในระยะต่อไป (perceptual foresight) ความสามารถที่จะนึกเห็นภาพของการปฏิบัติในแนวต่าง ๆ หากได้มีการกระทำขึ้นในระยะต่อไป

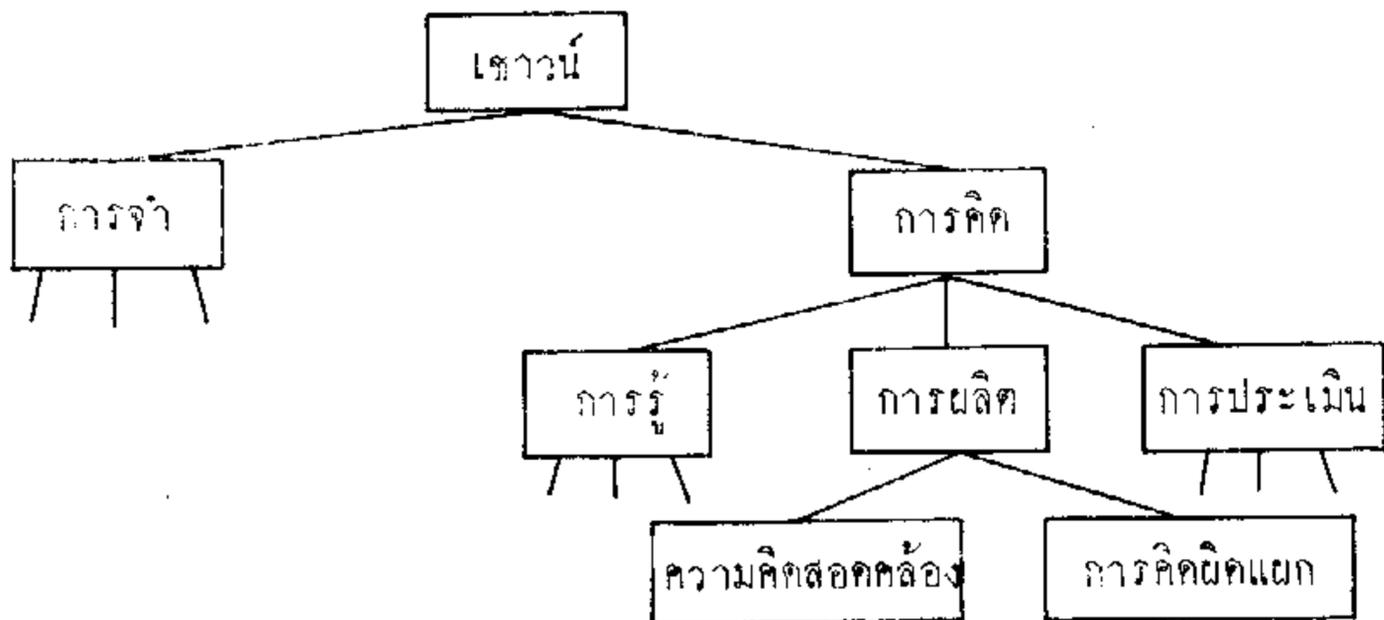
16. การรู้ความแทรกของเครื่องหมาย (cognition of symbolic implications) ความสามารถในการวางโครงสร้างของเครื่องหมายต่าง ๆ

17. การมีมโนภาพทางแนวคิด (conceptual foresight) ความสามารถมองเห็นล่วงหน้าได้ว่าแนวความคิดต่าง ๆ นั้นมีความหมายซ้อนแทรกอยู่อย่างไร

องค์ประกอบอื่น ๆ ก็อาจพิจารณาได้ในทำนองเดียวกัน เช่น องค์ประกอบที่เป็นค่ารวมระหว่าง การจำ กับความสัมพันธ์และเครื่องหมายคือ ความจำแบบนกแก้วนกขุนทอง (rote memory) การจำกับความแทรกและเครื่องหมาย คือ ความคล่องแคล่วในเรื่องตัวเลข (numerical fluency) การคิดให้สอดคล้องกับความแทรก และภาษา คือความสามารถในการอนุมาน (deduction) ส่วนการคิดให้ผิดแผกกับความแทรกและภาษาคือ ความสามารถขยายความ (semantic elaboration) เป็นต้น

ลำดับลดหลั่นขององค์ประกอบของเชาว์นี้ เนื่องจากองค์ประกอบที่แยกตามทฤษฎีนี้มี นับเป็นร้อยตัว จึงทำให้การที่จะสร้างเครื่องมือวัดตามทฤษฎีนี้เป็นไปด้วยความลำบาก ฉะนั้นจึงได้มีการกำหนดลำดับลดหลั่น (hierarchy) ขององค์ประกอบที่ทดลองได้มาเพื่อให้ทราบความสัมพันธ์ที่แท้จริงระหว่างองค์ประกอบแต่ละตัว ในแผนภาพที่ 2 แสดงลำดับลดหลั่นขององค์ประกอบในด้านการผลิต

แผนภาพที่ 2 แสดงว่าเมื่อพิจารณาเชาว์ในแง่การปฏิบัติการ องค์ประกอบที่อยู่ลำดับขั้นแรกคือ ความจำ และ การคิด ดังได้เคยอธิบายแล้ว ส่วนการรู้ การผลิต และการประเมิน

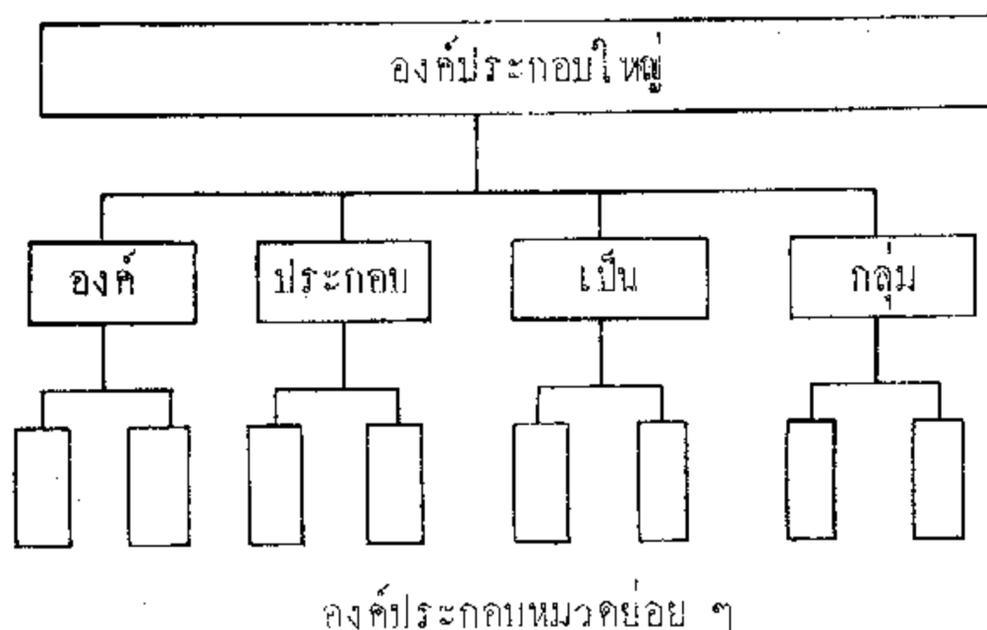


แผนภาพที่ ๒ ลำดับลดหลั่นขององค์ประกอบของเชาว์

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

เป็นองค์ประกอบในลำดับที่ลดลงไป ในแผนภาพนี้จะเห็นได้ว่าลดหลั่นลำดับขององค์ประกอบในขั้นการผลิตได้แสดงไว้โดยสมบูรณ์ที่สุด เพราะได้แยกแยะไปอีกว่า ถัดจากองค์ประกอบในขั้นการผลิตไปแล้วยังมีองค์ประกอบอะไรอีกบ้าง จะเห็นได้ว่าหากพิจารณาแต่เพียงแผนภาพที่ 1 จะไม่ทราบลดหลั่นลำดับขององค์ประกอบต่าง ๆ เลย

การที่การกำหนดลำดับลดหลั่นองค์ประกอบช่วยให้ทราบความสัมพันธ์ระหว่างองค์ประกอบนั้นเป็นการอำนวยความสะดวกในการสร้างเครื่องมือวัด เพราะการที่สร้างเครื่องมือวัดให้ครอบคลุมองค์ประกอบทุกตัวโดยไม่มีขอบเขตจำกัดว่าจะเน้นหนักองค์ประกอบตัวใดตัวหนึ่งเป็นหลักย่อมทำได้ยาก นักจิตบุคคลชาวอังกฤษเช่น Cyril Burt และ Philip E. Vernon ได้สนับสนุนการกำหนดลำดับลดหลั่นขององค์ประกอบไว้ในที่หลายแห่ง อย่างไรก็ตามก็ถือการค้นคว้าเรื่องลำดับลดหลั่นขององค์ประกอบนี้ยังไม่เป็นที่ยุติ เพราะผลการวิจัยในเรื่องความสามารถของบุคคลนี้ยังไม่เป็นที่สิ้นสุด ดังนั้นในการนี้จึงจะเสนอแต่เพียงโครงแบบ (Model) ของลำดับลดหลั่นที่ควรจะเป็นตามแนวคิดของ Humphreys



แผนภาพที่ 8 โครงแบบของลำดับผลหลั้องค์ประกอบสามชั้น

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

วิธีการวัดเชาว์

นับตั้งแต่ Binet ได้สร้างแบบทดสอบเชาว์ ก็ได้มีผู้สร้างแบบทดสอบเชาว์ขึ้นอีกมากมาย แบบทดสอบบางแบบใช้เวลาทำเพียง 15 นาที บางแบบก็ต้องใช้เวลาถึงกว่าชั่วโมง แต่แบบทดสอบทั่ว ๆ ไปก็ได้วัดเชาว์อย่างแท้จริง⁶ เพราะสิ่งที่ได้อธิบายไว้ในทฤษฎีโครงสร้างของเชาว์ การวัดเชาว์จริง ๆ นั้นทำได้ด้วยความลำบากเพราะเชาว์มีองค์ประกอบมากมายจนไม่สามารถทราบลำดับผลหลั้องค์ประกอบ แม้แบบทดสอบของ Binet ที่วัดผลมาได้เป็น I.Q. นั้นก็เป็นเพียงการวัดผลของการที่บุคคลที่มีความสามารถทางสมองมิใช่เป็นการวัดความสามารถทางสมองของบุคคลโดยตรง ในการวัดผลของการมีความสามารถทางสมองนี้ก็วัดได้เพียงว่ามีความสามารถในด้านต่าง ๆ เช่น ความสามารถในการเข้าใจภาษา ความสามารถในการหาเหตุผล ความสามารถในการจำ ความคล่องในเรียงตัวเลข เป็นต้น ยังไม่สามารถวัดผลในระดับสูงของความสามารถทางสมองที่ได้แก่ ความคิดริเริ่ม ความสามารถในการวางแผน หรือการมีคุณยพินิจอันสุขุมได้

⁶ Robert M. Guion, *Personnel Testing* (New York: McGraw-Hill, Inc., 1965), p. 238

C.H. Lawshe and Michael J. Balma, *Principles of Personnel Testing* (New York: McGraw-Hill, Inc., 1966) pp. 83-84

ตัวอย่างสิ่งที่น่าสนใจใช้วัดเชาว์ บุคคลต่าง ๆ ที่ได้กล่าวมาแล้วในเรื่องนี้ต่างพยายามสร้างแบบทดสอบมาตรฐานเพื่อวัดเชาว์ตามแนวความคิดของตน ในบรรดาแบบทดสอบเรื่องนี้นั้นอาจนำมายกตัวอย่างเพื่อให้เห็นว่า ถ้าจะวัดเชาว์แล้วจะวัดกันอย่างไรดังต่อไปนี้⁷

1. ในเรื่องศัพท์ (vocabulary)

จงเลือกคำที่มีความหมายใกล้เคียงที่สุดกับคำว่า ล้ำลิ้น (คำอุทานที่ถูกคือคำที่ซัดเส้นใต้)

- | | |
|------------|-------------------|
| ก. ร่าเริง | ข. <u>แข็งแรง</u> |
| ค. อ้วน | ง. เล็ก |
| จ. รวย | |

2. ในเรื่องความหมายในภาษา (verbal analogies)

กิ่งไม้มีความสัมพันธ์กับต้นไม้ เหมือนกับลำธารมีความสัมพันธ์กับ

- | | |
|-----------------------|------------------|
| ก. น้ำ | ข. ราก |
| ค. <u>ริมฝั่งน้ำ</u> | ง. <u>แม่น้ำ</u> |
| จ. การแตกกระจายของน้ำ | |

3. ในเรื่องการทำเหตุผลทางคณิตศาสตร์ (arithmetic reasoning)

เด็กชายคนหนึ่งซื้ลูกอมมากล่องละ 90 สตางค์ ในกล่องมีลูกอม 24 เม็ด เขานำลูกอมไปขายเม็ดละ 5 สตางค์ ถ้าว่าเขาได้กำไรเม็ดละกี่สตางค์

- | | |
|------------------------|-----------------|
| ก. 30 สตางค์ | ข. 3 3/4 สตางค์ |
| ค. <u>1 1/4 สตางค์</u> | ง. 4/5 สตางค์ |
| จ. ไม่มีข้อใดถูก | |

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

4. ในเรื่องการเรียงลำดับตัวเลข (number series)

เลขอะไรควรต่อจากลำดับเลขที่จัดไว้แล้ว ดังนี้ คือ

1 2 4 7 11

ก. 14

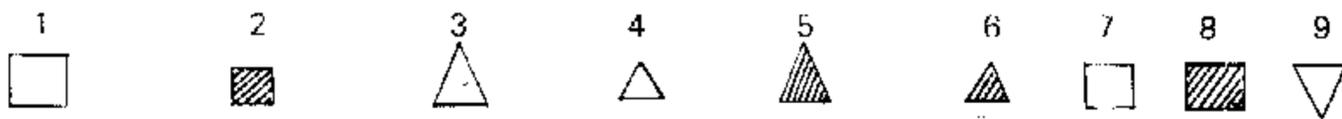
ข. 15

ค. 16

ง. 18

จ. 22

5. ความหมายของรูป



รูปที่ 1 เทียบกับรูปที่ 2 ได้เหมือนกับรูปที่ 3 เทียบกับ รูปที่.....

4, 5, 6, 7, 8, 9

แนวคิดเรื่องลักษณะต่าง ๆ ของแบบทดสอบ แบบทดสอบที่ได้สร้างขึ้นเป็นเครื่องมือวัดเชาวน์มีลักษณะที่ต่าง ๆ กันไป อันควรนำมากล่าวไว้ในที่นี้คือ ประการแรก แบบทดสอบนี้มีทั้งใช้ภาษา (verbal tests) และที่ไม่ใช้ภาษา (non-verbal tests) ประการที่สองมีแบบทดสอบบางชุดที่ปลอดจากอิทธิพลของวัฒนธรรม (culture-free tests) และประการที่สาม แบบทดสอบเชาวน์อาจแยกประเภทตามลักษณะของการควบคุมการทดสอบก็ได้ คือ แยกประเภทเป็นแบบทดสอบกลุ่ม (group tests) และแบบทดสอบรายตัว (individual tests)⁸

โดยทั่ว ๆ ไปแบบทดสอบก็สร้างขึ้นเป็นข้อเขียนโดยประสงค์ให้ผู้เข้าสอบตอบคำถามที่ปรากฏบนกระดาษ แต่เนื่องจากมีบุคคลบางกลุ่มที่จะเสียเปรียบในการตอบแบบทดสอบเพราะมีความชำนาญทางภาษาดีกว่าบุคคลทั่ว ๆ ไป เช่น ชาวต่างประเทศที่ไม่ได้ใช้ภาษาเดียวกับที่ใช้ในแบบทดสอบอีกประการหนึ่งในบรรดาบุคคลที่ยังเยาว์วัยอยู่ต่างก็มีอัตราความเร็วในการ

⁸ Guion, *op. cit.*, pp. 491-493.

Thorndike and Hagen, *op. cit.*, pp. 214-218, 221-226.

เรียนภาษาไม่เท่ากัน แบบทดสอบที่เป็นข้อเขียนจึงทำให้บุคคลเหล่านี้บางคนต้องเสียเปรียบในการตอบ ทำให้ค่าคะแนนต่ำกว่าที่ควรได้ตามระดับของเขาวงกตที่มีอยู่จริง จึงได้มีการสร้างแบบทดสอบเขาวงกตที่ไม่ต้องใช้ความรู้ในเรื่องภาษา (non-language tests) เช่น *Large-Thorndike Intelligence Test Non-Verbal Series* และ *Pintner Non-Language Test* คำถามในแบบทดสอบประเภทนี้เป็นปัญหาเกี่ยวกับความหมายของรูปหรือลำดับตัวเลข

นอกจากปัญหาในการที่ความสามารถทางภาษาทำให้ไม่ได้ค่าคะแนนที่แท้จริงของเขาวงกตแล้วยังมีปัญหาในเรื่องเกี่ยวกับอิทธิพลของวัฒนธรรมที่ทำให้ผู้ตอบแบบทดสอบบางกลุ่มต้องได้คะแนนต่ำกว่าที่ควรได้ตามระดับของเขาวงกตที่มีอยู่จริง กล่าวคือเนื้อหาของแบบทดสอบตามธรรมเนียมที่ยอมรับกันสภาพการณ์ทางวัฒนธรรมซึ่งทำให้บุคคลทั่วไปมีความคุ้นเคย เพราะมีประสบการณ์จากการครองชีวิตในสังคมหนึ่ง ๆ แต่ทำให้บุคคลที่ไม่เคยใช้ชีวิตในสังคมเดียวกันนี้มีความลำบากในการที่จะเข้าใจ เช่น คำถามที่เกี่ยวกับหิมะย่อมทำให้บุคคลในประเทศไทยทั่ว ๆ ไปตอบได้ยากเพราะไม่เคยเห็น หรือคำถามที่เกี่ยวกับต้นหมากก็อาจทำให้ชาวอเมริกันไม่มีมโนภาพในการนึกตอบ เพราะไม่เคยเห็นเช่นเดียวกัน ภาษานั้นเป็นส่วนหนึ่งของวัฒนธรรมแบบทดสอบที่เป็นข้อเขียนก็มีส่วนที่ทำให้ปัญหาในเรื่องวัฒนธรรม แบบทดสอบที่เป็นข้อเขียนก็มีส่วนที่ทำให้ปัญหาในเรื่องวัฒนธรรมมาบังค้ำคะแนนของผู้ตอบแบบทดสอบ ฉะนั้นจึงได้มีการพยายามสร้างแบบทดสอบที่ปลอดจากอิทธิพลของวัฒนธรรม (culture-free tests) แต่การสร้างแบบทดสอบที่ปลอดจากอิทธิพลของวัฒนธรรมจริง ๆ ทำได้ยาก จึงมีชื่อเรียกแบบทดสอบที่พยายามขจัด แต่ไม่ปลอดจากอิทธิพลของวัฒนธรรมเสียทีเดียวว่า แบบทดสอบที่ไม่ลำเอียงในแง่วัฒนธรรม (culture-fair tests) เช่น *Cattell Culture Free Intelligence Test* และ *Miller Analogies Test* คำถามในแบบทดสอบประเภทนี้มักเป็นรูปภาพ

แบบทดสอบกลุ่ม (group tests) ต่างกับแบบทดสอบรายตัว (individual tests) ๓ ประการใหญ่ ๆ คือ ประการแรก คำถามในแบบทดสอบกลุ่มเป็นข้อเขียนเหมือนข้อสอบทั่วไป ผู้ตอบแบบทดสอบต้องอ่านคำถามเอง แต่แบบทดสอบรายตัวนั้น ผู้ควบคุมการสอบอ่านให้ผู้ตอบแบบทดสอบฟัง ประการที่สอง ปัญหาทั้งหมดในแบบทดสอบกลุ่มกำหนดเวลารวม ๆ ที่จะต้องตอบไว้ให้โดยไม่จำกัดว่าผู้ตอบแบบทดสอบคนใดจะใช้เวลาตอบปัญหาแต่ละข้อเป็นปริมาณเท่า ๆ กันหรือไม่ ส่วนแบบทดสอบรายตัว ผู้ตอบมีเวลาตอบปัญหาเรียงไปที่ละข้อโดยไม่จำกัดเวลา และประการสุดท้ายแบบทดสอบกลุ่มมักจะกำหนดคำตอบไว้ให้ผู้ตอบเลือก ส่วนผู้ตอบแบบทดสอบรายตัวนั้นจะคิดคำตอบอย่างใดก็ได้ ความแตกต่างเช่นนี้หมายความว่า

ผู้ตอบแบบทดสอบรายตัวไม่จำเป็นต้องใช้ความสามารถในการอ่าน แต่ผู้ควบคุมการทดสอบแบบนี้ก็ต้องใช้เวลามากกว่าการทดสอบแบบกลุ่ม อย่างไรก็ตาม การที่ผู้ตอบสามารถเลือกคำตอบอย่างใดก็ได้ทำให้การทดสอบช่วยให้ทราบลักษณะของพฤติกรรมต่าง ๆ ได้กว้างขวางกว่าการทดสอบแบบกลุ่ม แบบทดสอบรายตัวจึงมีประโยชน์ในการทดสอบบุคคลรุ่นเยาว์มากกว่าบุคคลที่มีวัยสูงแล้ว

Wechsler Adult Intelligence Scale (WAIS)⁹ แบบทดสอบนี้เป็นแบบทดสอบเชาวน์ที่มีชื่อเสียงมากจึงนำมากล่าวถึงเป็นตัวอย่างไว้ในที่นี้ แบบทดสอบฉบับนี้ปรับปรุงแก้ไขเสร็จเมื่อปี 1955 มี 11 ภาค จากผลการวิจัยโดยแบบทดสอบนี้ปรากฏว่า สามารถวัดองค์ประกอบได้ 3 ตัวคือ ความเข้าใจภาษา (verbal comprehension) ซึ่งได้แก่ศัพท์ความเข้าใจข่าวสาร และเรื่องต่างๆ ในภาษา องค์ประกอบตัวที่สองคือ ผลรวมของสัญญาณ (perceptual organization) ซึ่งได้แก่ความสามารถในการออกแบบ การประมวลสิ่งของและจัดรูปภาพ องค์ประกอบตัวที่สาม คือ การจำ (memory) ซึ่งได้แก่ เลขคณิต และตัวเลขต่างๆ แต่ถึงแม้ว่าแบบทดสอบนี้เป็นที่ยกย่องกันมากกว่าสามารถวัดเชาวน์ได้ดี แต่ก็ไม่ค่อยเหมาะในการวัดผลสำหรับงานบุคคล เพราะแบบทดสอบฉบับนี้เป็น แบบทดสอบรายตัวจึงต้องเปลืองเวลามากในการควบคุมการสอบ ในปี 1956 ได้มีการทดลองใช้แบบทดสอบนี้ในงานบุคคล ปรากฏว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าคะแนนของผู้ตอบแบบทดสอบนี้กับค่าคะแนนผลการปฏิบัติงานของบุคคลกลุ่มเดียวกันอยู่ในระดับค่อนข้างต่ำ สาระสำคัญของแบบทดสอบฉบับนี้ในแต่ละภาคมีดังต่อไปนี้

1. ข่าวสารทั่วไป (general information) คำถามที่ให้ผู้ตอบเติมคำในช่องว่างเกี่ยวกับเรื่องราวทั่ว ๆ ไปที่ทุกคนอาจพบเห็นได้
2. ความเข้าใจ (comprehension) มีปัญหา 10 รายการ ถามถึงเหตุผลที่จำเป็นต้องมีกฎเกณฑ์ต่าง ๆ ทางสังคมเพื่อแก้ไขปัญหาทั่ว ๆ ไปในสังคม

⁹ H.A. Witkin, et. al., *Psychological Differentiation* (New York: John Wiley & Sons, Inc., 1962), pp. 73-74.

Guion, *op. cit.*, pp. 215-219.

3. การหาเหตุผลทางคณิตศาสตร์ (arithmetic reasoning) เป็นคำถามปากเปล่า ผู้ตอบจะต้องตอบให้ทันเวลา
4. ตัวเลขต่างๆ (digit span) ผู้เข้าสอบฟังตัวเลขหลาย ๆ ตัว และว่าตามให้ถูกต้อง มีตัวเลขตั้งแต่สามตัวถึงเก้าตัว
5. ความคล้าย (similarities) แต่ละคำถามมีคำเป็นคู่ ผู้เข้าสอบต้องอธิบายลักษณะที่คล้ายกันของแต่ละคู่
6. ศัพท์ (vocabulary) ผู้เข้าสอบต้องนิยามศัพท์ 42 ตัว มีทั้งศัพท์ที่ใช้ทั่วไปประจำวัน และศัพท์ที่ยาก
7. การเติมรูปให้สมบูรณ์ (picture completion) ผู้เข้าสอบต้องชี้ว่า แต่ละภาพในปัญหาที่ยังขาดส่วนใดบ้าง
8. การจัดรูป (picture arrangement) มีรูปภาพ 8 ชุด ที่ต้องเรียงให้ได้ความเป็นเรื่องราว
9. การจัดสิ่งของ (object assembly) มีรูปของสิ่งของ 4 อย่างที่ตัดไว้เป็นชิ้น ๆ เพื่อให้ต่อให้สมบูรณ์ภายในเวลาที่กำหนด
10. การออกแบบ (block design) มีแท่งสีขาวและสีแดงที่จะต้องนำมาต่อให้เป็นรูปแบบ 10 อย่าง
11. เครื่องหมายตัวเลข (digit symbol) กำหนดเครื่องหมาย 9 อย่างสำหรับตัวเลขแต่ละตัว ผู้เข้าสอบจะต้องนำเอาโจทย์ซึ่งมีเครื่องหมายต่างๆ มาแปลให้เป็นตัวเลขตามเกณฑ์ที่กำหนดไว้

Miller Analogies Test (MAT)¹⁰ แบบทดสอบนี้เชื่อกันว่าใช้ได้ผลในการวัดบุคคล ได้มีการวิจัยแล้วว่าค่าคะแนนของผู้เข้าสอบมีสัมประสิทธิ์สหสัมพันธ์กับผลการปฏิบัติงานในหน้าที่ผู้จัดการฝ่ายขาย และเจ้าหน้าที่เทคนิคในห้องปฏิบัติการวิทยาศาสตร์ แบบทดสอบฉบับนี้แบ่งเป็น 4 ชุด แต่ละชุดมีคำถาม 100 ข้อ เป็นปัญหาเกี่ยวกับความรู้ต่างๆ อย่างกว้างขวาง

¹⁰ *Ibid.*, pp. 227-228.

เนื้อหาของแบบทดสอบมีสัมประสิทธิ์ความเชื่อมั่นในระดับสูง จากผลการวิจัยปรากฏว่าสัมประสิทธิ์ความเชื่อมั่นตามวิธีเครื่องมือวัดนี้กับเครื่องมือวัดอื่น ๆ ได้ค่าระหว่าง 0.85 ถึง 0.89 แต่การที่จะใช้แบบทดสอบนี้วัดบุคคล หน่วยงานที่ประสงค์จะต้องส่งบุคคลไปรับการทดสอบ ที่ทำการของหน่วยการทดสอบที่ได้รับอนุญาตเฉพาะ (licensed testing centers) เท่านั้น ผลการทดสอบจะแจ้งให้ทราบแต่เฉพาะบริษัทธุรกิจ หรือหน่วยงานที่ได้รับการพิจารณาแล้ว เท่านั้น

เชาวน์กับการปฏิบัติงาน

หากจะวัดเชาวน์ให้ได้ประโยชน์เต็มที่ในการบริหารงานบุคคล ก็จะต้องมีเครื่องมือวัดที่ช่วยให้ได้ค่าคะแนนที่สอดคล้องกับความสำเร็จในการปฏิบัติงานตามทฤษฎีของ Guilford สิ่งที่แสดงว่าบุคคลหนึ่ง ๆ มีเชาวน์ในระดับสูงที่จะเป็นประโยชน์ต่อการปฏิบัติงานแยกได้เป็น 3 ประการคือ ความสร้างสรรค์ (creativity) ความสามารถแก้ปัญหา (problem solving) และความสามารถในการวางแผน (planning)¹¹ ความสร้างสรรค์หมายถึง ความสามารถในการคิดเรื่องต่าง ๆ ให้แตกแขนงออกไป (divergent thinking ability) ความสามารถแก้ปัญหาคือ ส่วนผสมของความสามารถในการคิดหยั่งรู้เหตุการณ์ล่วงหน้า และความสามารถที่จะรู้จักปัญหา ประกอบกับความเป็นตัวของตัวเอง ส่วนความสามารถในการวางแผนเป็นองค์ประกอบของความสามารถในการคิดหยั่งรู้เหตุการณ์ล่วงหน้า ความสามารถในการคิดหยั่งรู้เหตุการณ์ล่วงหน้า ความสามารถที่จะจัดลำดับ และความสามารถที่จะอธิบายความหมายต่าง ๆ ได้

หากพิจารณาตามแนวที่กล่าวมานี้ เครื่องมือวัดที่มีอยู่ในปัจจุบันยังไม่สามารถที่จะวัดเชาวน์ในลักษณะที่ครบถ้วนได้ กล่าวได้ว่า เครื่องมือวัดเชาวน์ต่าง ๆ วัดองค์ประกอบบางตัวของเชาวน์ได้ เช่นความเข้าใจในเรื่องภาษา การแยกแนวคิดในภาษา การสรุปความสัมพันธ์ของภาษา องค์ประกอบในเรื่องความจำบางตัว ส่วนที่จะวัดว่าบุคคลมีความคิดริเริ่ม หรือมีความสามารถในการวางแผนอย่างไคนั้นยังทำไม่ได้

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

Quion. *op. cit.*, pp. 77-79.

ถึงแม้เครื่องมือวัดเชาว์จะยังขาดความสมบูรณ์ตามแนวคิดกล่าว ลักษณะของเชาว์ที่วัดให้ตามเครื่องมือวัดที่มีอยู่ในปัจจุบันก็มีความสัมพันธ์กับความสำเร็จในการปฏิบัติงานในระดับที่น่าพอใจ โดยเฉพาะในเรื่องการคัดเลือกคนเพื่อรับการฝึกอบรมก่อนการปฏิบัติงาน (trainability in job)

เมื่อเทียบบุคคลในอาชีพต่าง ๆ แล้ว ปรากฏว่าบุคคลที่มีอาชีพต่างกันมีระดับของเชาว์ต่างกัน จากผลการวิจัย ทหารเกณฑ์ในสหรัฐอเมริกาในสมัยสงครามโลกครั้งที่ 2 ปรากฏว่าอาชีพต่าง ๆ มีส่วนสัมพันธ์กับเชาว์ ดังแสดงไว้ในตารางที่ 2

ตารางที่ 2

ความสัมพันธ์ระหว่างเชาว์กับอาชีพ

อาชีพ	เปอร์เซ็นต์ไทล์				
	10	25	50	70	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
ช่างแก้วิทยุ	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
คนชำแหละเนื้อ	80	94	108	117	126

อาชีพ	เปอร์เซ็นต์ไทล์				
	10	25	50	75	90
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
ช่างตัดผม	66	79	93	109	120
กรรมกรทั่วไป	65	76	93	108	119
กรรมกรเหมืองแร่	67	75	87	103	119
ลูกจ้างทำนา	61	70	86	103	115
คนตัดไม้ประจำโรงเลื่อย	60	70	85	100	116

แหล่งที่มา : Robert L. Thorndike, and Elizabeth Hagen, *Measurement and Evaluation in Psychology and Education*, p. 232.

ผลการวิจัยในตารางที่ 2 เป็นค่าเปอร์เซ็นต์ไทล์ที่ได้จากการทดลองใช้แบบทดสอบ Army General Classification Test เมื่อเทียบระหว่างอาชีพกับบัญชีหรือครูกับ อาชีพกรรมกร หรือ ลูกจ้างทำนา จะเห็นว่ามีความแตกต่างกันอย่างเห็นได้ชัด อย่างไรก็ตามบุคคลที่ดีที่สุดของอาชีพที่ รong ๆ ลงมากก็มีระดับเขาวนที่ไม่แพ้บุคคลในอาชีพสูง ๆ ขึ้นไป หรืออาจจะดีกว่าด้วยซ้ำ เช่น ช่างไม้ หรือ คนขับรถบรรทุกขนาดใหญ่ที่ดีที่สุดใน 10% แรก มีเขาวนในระดับเดียวกับ พนักงานบัญชี หรือช่างเขียนแบบที่อยู่ในระดับปานกลาง และมีเขาวนสูงกว่าทนายความหรือครูที่อยู่ใน ประเภท 10% ท้ายสุดเสียอีก

Ghiselli และ Brown ได้ทำการค้นคว้าความเที่ยงตรงของเครื่องมือวัดเขาวนชุดต่างๆ ในแง่การคาดคะเน (predictive validity) ปรากฏผลว่า เครื่องมือวัดเขาวนสามารถคาดคะเนใน เรื่องการคัดเลือกบุคคลเพื่อรับการฝึกอบรมก่อนการปฏิบัติงานดีกว่าในเรื่องผลสำเร็จในการปฏิบัติ งาน สรุปผลการค้นคว้าในเรื่องนี้มีแสดงไว้ในตารางที่ 3

ตารางที่ 3

เกณฑ์เฉลี่ยของความสามารถในการกะเนของแบบทดสอบเซาวัน

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

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

จากตารางที่ 3 จะเห็นว่าค่าในช่องการฝึกอบรมมีแนวโน้มสูงกว่าในช่องความสำเร็จของงาน ค่านี้คือสัมประสิทธิ์ความเที่ยงตรงในแง่ภาคคะเน คำนวณได้โดยนำค่าคะเนของการวัดเซาวันโดยแบบทดสอบต่างๆ ไปเทียบกับค่าที่ประเมินเพื่อรับการฝึกอบรม และค่าที่ประเมินผลงานตามลำดับ แล้วนำมาเฉลี่ย เฉพาะในตำแหน่งงานในฝ่ายจัดการและพนักงานขายนั้นจะเห็นว่ายังลงความเห็นเป็นที่ยุติไม่ได้ว่า เซาวันที่วัดได้โดยแบบทดสอบต่างๆ นั้นจะช่วยในการคาดคะเนความเหมาะสมที่จะรับเลือกเข้าฝึกอบรมก่อนปฏิบัติงาน อย่างไรก็ตามก็ตีในตำแหน่งงาน

ทั้งสองประเภทนี้ แบบทดสอบเชาวน์ช่วยในการคาดคะเนความสำเร็จในการปฏิบัติงานได้ตามสมควร ทั้งนี้จะสังเกตได้จากสัมประสิทธิ์ .37 และ .32 ตามลำดับ

จากข้อเท็จจริงดังกล่าวนี้ Guion¹² จึงสรุปว่า เนื่องจากการที่จะปฏิบัติงานในหน้าที่ต่าง ๆ ได้นั้น บุคคลจะต้องมีความสามารถทางสมองที่จะเรียนรู้สิ่งต่าง ๆ ที่เกี่ยวกับงาน ฉะนั้น แม้ว่า แบบทดสอบต่าง ๆ จะยังไม่สามารถวัด "ตัวเชาวน์" อย่างแท้จริง แบบทดสอบเหล่านี้ก็ช่วยในการคาดคะเนในงานบุคคลอย่างน่าพอใจ กล่าวคือ นอกจากจะช่วยในการพิจารณาว่าควรรับบุคคลใดเข้าฝึกอบรมเพื่อปฏิบัติงานต่อไป บุคคลใดจะปฏิบัติงานได้สำเร็จ บุคคลใดจะปฏิบัติงานได้นานเท่าใดและบุคคลใดสมควรได้รับการเลื่อนตำแหน่ง เพราะการที่บุคคลจะฝึกอบรมด้วยดี หรือปฏิบัติงานด้วยดีหรือจะมีกำลังใจปฏิบัติงานต่อไป หรือจะสามารถปฏิบัติงานในตำแหน่งสูงขึ้นไป จะต้องอาศัยความสามารถทางสมอง แบบทดสอบต่าง ๆ ที่มีอยู่ในปัจจุบันก็สามารถวัดผลของการใช้ความสามารถทางสมองของบุคคลได้กว้างขวางพอสมควร

แนวคิดในการสร้างมาตรวัดความสำเร็จของงาน

การวัดในการบริหารงานบุคคลที่มีลักษณะพิเศษกว่าการวัดทั่วไปในทางจิตวิทยาในแง่ที่ว่า การวัดทางจิตวิทยาเป็นเพียงการวัดลักษณะที่ต้องการวัดเท่านั้น แต่การวัดในการบริหารงานบุคคลเป็นการวัดเพื่อให้ได้ประโยชน์ในการคาดคะเนความสามารถในการปฏิบัติงานของบุคคลต่าง ๆ อีกด้วย โดยเฉพาะในการวัดสติปัญญาที่กล่าวมาแล้วในหัวข้อก่อน นักวิชาการในการวัดเพื่อการบริหารงานบุคคลจะไม่ประสงค์แค่เพียงที่จะวัดเชาวน์โดยทั่ว ๆ ไป แต่จะมุ่งที่จะทราบว่าการปฏิบัติงานหนึ่ง ๆ จำเป็นต้องมีสติปัญญาในด้านใดบ้าง และอย่างไร การที่จะทำได้เช่นนั้น นักวิชาการในการวัดจะต้องวิเคราะห์พฤติกรรมในการปฏิบัติงาน (job behavior) เพื่อให้ทราบว่าบุคคลที่จะปฏิบัติงานในตำแหน่งหนึ่ง ๆ อย่างไรก็ดีผลสำเร็จจะต้องแสดงพฤติกรรมที่เกี่ยวข้องโดยตรงกับงานนั้น ๆ อย่างไรก็ตาม พฤติกรรมนี้จะเป็นตัวแปรที่นำมาเทียบค่ากับผลการทดลองใช้เครื่องมือวัดเพื่อให้ทราบว่าเครื่องมือวัดหนึ่ง ๆ มีคุณค่าในการคาดคะเนว่าบุคคลจะปฏิบัติงานตำแหน่งต่าง ๆ ได้ผลสำเร็จประการใด การที่นำพฤติกรรมในการปฏิบัติงาน มาวิเคราะห์ค่าเทียบคุณค่าของเครื่องมือวัดเช่นนี้จึงได้มีการเรียกพฤติกรรมเฉพาะในเรื่องนี้ว่า มาตรวัดความสำเร็จของงาน (criterion measure)

¹² *Ibid.*, pp. 237-238.

เครื่องมือวัดบุคคลจะมีค่าในการบริหารงานบุคคลที่ต่อเมื่อเครื่องมือวัดนั้น ๆ ช่วยให้ผู้บริหารงานคาดคะเนได้ว่า ผู้ถูกวัดจะแสดงพฤติกรรมเช่นใดเมื่อได้รับการแต่งตั้งให้ทำงานหนึ่งๆ การที่จะทำได้เช่นนี้ก็ต้องเมื่อนักวิชาการในการวัดได้กำหนดมิติของมาตรวัดความสำเร็จของงานไว้เรียบร้อยแล้ว และเครื่องมือวัดที่ใช้กันนั้นๆ สามารถจัดลำดับหรือแบ่งกลุ่มผู้ถูกวัดในแง่พฤติกรรมการทำงานให้สอดคล้องกับมิติที่ได้อ้างไว้

การที่กล่าวว่า ให้ใช้พฤติกรรมในการปฏิบัติงานเป็นมาตรวัดความสำเร็จของงานเพื่อเทียบคุณค่าของเครื่องมือวัดนั้น พิจารณาดูเผินๆ ก็ไม่น่าจะมีปัญหาอะไร แต่โดยความเป็นจริงแล้วเป็นปัญหาที่ยุ่งยากมาก เพราะพฤติกรรมการทำงานที่จะนำมาใช้เป็นตัวแทนของความสำเร็จผลในการปฏิบัติงานจริงๆ มีลักษณะยากต่อการจำกัดความ ในการที่จะให้ได้ค่าจำนวนที่เทียบได้กับเครื่องมือวัด ยิ่งไปกว่านั้นยังมีปัญหาอื่น ๆ ตามมาอีกหลายปัญหา การวิเคราะห์ค่าพฤติกรรมของงานหนึ่ง ๆ ควรวิเคราะห์เพื่อให้มีมาตรเพียงอันเดียว ซึ่งจะมีลักษณะแสดงการปฏิบัติงานทุก ๆ อย่าง หรือจะต้องมีมาตรหลาย ๆ อันที่เป็นอิสระแก่กัน แต่เป็นมาตรที่วัดพฤติกรรมต่าง ๆ ของงานตำแหน่งเดียวกัน ปัญหาต่อไปก็คือว่า มีตัวแปรพื้นฐานอะไรอยู่บ้างที่เกี่ยวข้องกับมาตรฐานการประเมินค่าความสำเร็จในการปฏิบัติงาน ปัญหานี้เกี่ยวข้องกับเรื่องที่ว่า ควรพิจารณาตัวแปรในด้านเศรษฐกิจที่สัมพันธ์กับผลการปฏิบัติงานหรือควรถือตัวแปรในเรื่องความประพฤตินในการทำงาน เป็นหลักในการกำหนดมาตร และอีกเรื่องหนึ่งคือว่า ตำแหน่งงานที่วัดนั้นมีลักษณะสอดคล้องกับนโยบายของหน่วยงานอย่างไร เมื่อทราบวิธีแก้ปัญหเหล่านี้มาตรที่วิเคราะห์มาได้จึงจะมีคุณค่าเป็นมาตรที่ดี

มาตรวัดความสำเร็จของงานขั้นสุดท้าย ในทางปฏิบัติผู้สร้างมาตรวัดความสำเร็จของงานย่อมไม่สามารถสร้างมาตรฯ ที่มีคุณสมบัติครบถ้วนดีทุกประการโดยแก้ไขปัญหาต่างๆ ที่กล่าวมาแล้วได้ทั้งหมด อย่างไรก็ตามการสร้างมาตรฯ ตามหลักวิชาก็ต้องตั้งความมุ่งหมายให้ได้มาตรฯ ที่มีความเที่ยงตรงและสมบูรณ์ที่สุด ฉะนั้น Thorndike¹³ ซึ่งเป็นนักวิชาการในการวัดผลงานบุคคลชั้นนำของสหรัฐฯ จึงได้เสนอให้แบ่งมาตรฯ ออกเป็น 3 ชั้น คือ ชั้นสุดท้าย ชั้นกลาง และขั้นต้น มาตรฯ ชั้นสุดท้ายคือ มาตรฯ ในอุดมคติในการคัดเลือกบรรจุแต่งตั้ง หรือในการคัดเลือกผู้ผ่านการฝึกอบรม เพื่อให้ได้บุคคลที่ดีที่สุดได้ปฏิบัติงาน มาตรฯ ชั้นกลาง คือมาตรฯ ที่

¹³ Robert L. Thorndike, *Personnel Selection* (New York: John Wiley & Sons, Inc., 1964), pp. 121-124.

หามาทดแทน (substitute criterion) ซึ่งได้มาโดยการวิเคราะห์ให้ค่าการปฏิบัติงานโดยมีหลักการ ส่วนมาตรฐานขั้นต้นหมายถึงข้อเท็จจริงเกี่ยวกับการปฏิบัติงานเป็นผลสำเร็จหรือไม่สำเร็จที่อาจมาพิจารณาได้ในทันที

มาตรฐานสุดท้ายมีลักษณะเป็นนามธรรมที่ประมวลทุกสิ่งทุกอย่างซึ่งอธิบายว่าการปฏิบัติงานได้ผลสำเร็จอย่างใดนั้นมีขอบเขตและความลึกซึ้งแค่ไหน¹⁴ Thorndike ยกตัวอย่างว่า มาตรฐานสุดท้ายของการขายประกันชีวิตก็คือปริมาณกรรมธรรม์จำนวนมากที่สุดจำนวนหนึ่งที่ผู้ขายประกันจะต้องขายได้ในช่วงระยะเวลาที่กำหนดไว้ช่วงหนึ่ง การกำหนดมาตรฐานนี้จะต้องคำนึงถึงปริมาณกรรมธรรม์ที่ขายได้ อัตราการต่ออายุสัญญากรรมธรรม์ ความคงเส้นคงวาของการปฏิบัติงาน ผลกระทบที่มีต่อการขายกรรมธรรม์ของผู้ขายประกันคนอื่น ๆ ผลการปฏิบัติงานเกี่ยวกับเอกสาร และหลักฐานสัญญาต่าง ๆ ความร่วมมือในการช่วยฝึกอบรมพนักงานใหม่ และลักษณะการประชาสัมพันธ์พนักงาน การที่จะรวบรวมข้อมูลทั้งหมดนี้ได้ก็ต้องดำเนินการตลอดชั่วอายุการปฏิบัติงานของผู้ขายประกันซึ่งเป็นการเสียเวลามาก และยังถ้าเป็นงานที่วัดปริมาณได้โดยยากกว่าการขายประกันชีวิต แม้จะรอจนชั่วอายุการปฏิบัติงานของบุคคลหนึ่ง ๆ ก็ยังอาจไม่ได้ข้อมูลครบถ้วน ถึงแม้จะสามารถรวบรวมข้อมูลได้ในชั่วอายุการปฏิบัติงาน ช่วงระยะเวลาที่ผ่านไปก็ทำให้ข้อมูลหย่อนคุณภาพลงในแง่ความเป็นจริง

ถึงแม้มาตรฐานสุดท้ายจะเป็นสิ่งที่วิเคราะห์ออกมาได้โดยยากก็ตาม ผู้สร้างมาตรฐานก็มีหน้าที่สร้างมาตรฐานที่จะมีลักษณะเข้าข่ายเป็นมาตรฐานสุดท้ายให้มากที่สุดที่จะทำได้ ในการนี้อาจให้หลักสถิติมาประกอบในการพิจารณาความสำเร็จของงาน จึงทำให้เกิดมาตรฐานที่มีลักษณะทดแทนซึ่งเรียกว่า มาตรฐานชั้นกลาง เช่น ถ้าหากผู้รู้จักงานที่จะวัดหนึ่ง ๆ ได้ก็ตกลงเห็นพ้องกันว่า การที่จะปฏิบัติงานให้สำเร็จด้วยดี ผู้ปฏิบัติงานจะต้องมีพฤติกรรมอย่างหนึ่งอย่างใดที่เห็นได้ชัดในช่วงเวลาหนึ่ง ข้อตกลงเห็นพ้องกันนี้ก็จะเป็นข้อมูลที่ใกล้เคียงกับมาตรฐานสุดท้ายมากที่สุด หลังจากที่ได้มีการเทียบค่าพฤติกรรมที่กำหนดนี้ออกมาเป็นตัวเลขที่นับได้ และพิสูจน์ค่าสหสัมพันธ์ระหว่างค่าพฤติกรรมดังกล่าวกับการประเมินการปฏิบัติงานจริง ๆ ว่ามีความเกี่ยวข้องกันเชื่อถือได้ ผู้สร้างมาตรฐานก็อาจนำเอาผลการประเมินการปฏิบัติงานจริง ๆ มาถือเป็นมาตรฐานชั้นกลางเพื่อใช้เทียบกับเครื่องมือวัดของตนได้ อย่างไรก็ตาม การพิสูจน์ค่าสหสัมพันธ์ระหว่างผลการประเมินการปฏิบัติงานกับมาตรฐานสุดท้ายเป็นงานที่ยุ่งยากและเปลืองเวลา

14 *Ibid.*

เพราะจะต้องมีการติดตามผลการประเมินเป็นระยะ ๆ เป็นเวลานาน ๆ และยังมีปัญหาในเรื่อง การกำหนดมาตรฐานขั้นสุดท้าย ข้อมูลที่นำมาใช้สร้างมาตรฐานชั้นกลางจึงมักจะต้องอาศัยหลักการ วิเคราะห์หาเหตุผล (rational basis) เป็นแนวในการพิจารณาว่าจะเชื่อถือได้หรือไม่ การ วิเคราะห์หาเหตุผลนี้ทำได้โดยพิจารณาข้อมูลทั้งหมดอย่างรอบคอบว่า เป็นเหตุเป็นผลกัน ตาม หลักการวิจัยหรือไม่ และมีความสอดคล้องกับเป้าหมายในการคัดเลือกบุคคลเพื่อปฏิบัติงานใน หน่วยงานหนึ่ง ๆ เพียงใด นอกจากนี้ยังอาจใช้วิธีการทางสถิติหาค่าสหสัมพันธ์ระหว่างข้อมูล ต่าง ๆ ที่จะนำมาพิจารณาเป็นมาตรฐานชั้นกลาง แม้ว่าข้อมูลเหล่านี้มิได้เป็นมาตรฐานขั้นสุดท้าย และค่าสหสัมพันธ์ที่ได้ก็ไม่ใช่ค่าสหสัมพันธ์ของข้อมูลที่มีอยู่กับค่าในมาตรฐานอันใดก็ตาม ค่า สหสัมพันธ์ระหว่างกันของข้อมูลก็ยิ่งอาจช่วยเสริมให้เห็นจริงว่าเหตุและผลที่ได้ใช้เป็นหลักใน การพิจารณาข้อมูลในขั้นต้นก่อนหน้านั้นมีหลักฐานสนับสนุนพอที่จะเชื่อได้ว่าข้อมูลต่าง ๆ มีความ สอดคล้องกันทำให้ใช้ประกอบการพิจารณาสร้างมาตรฐานได้

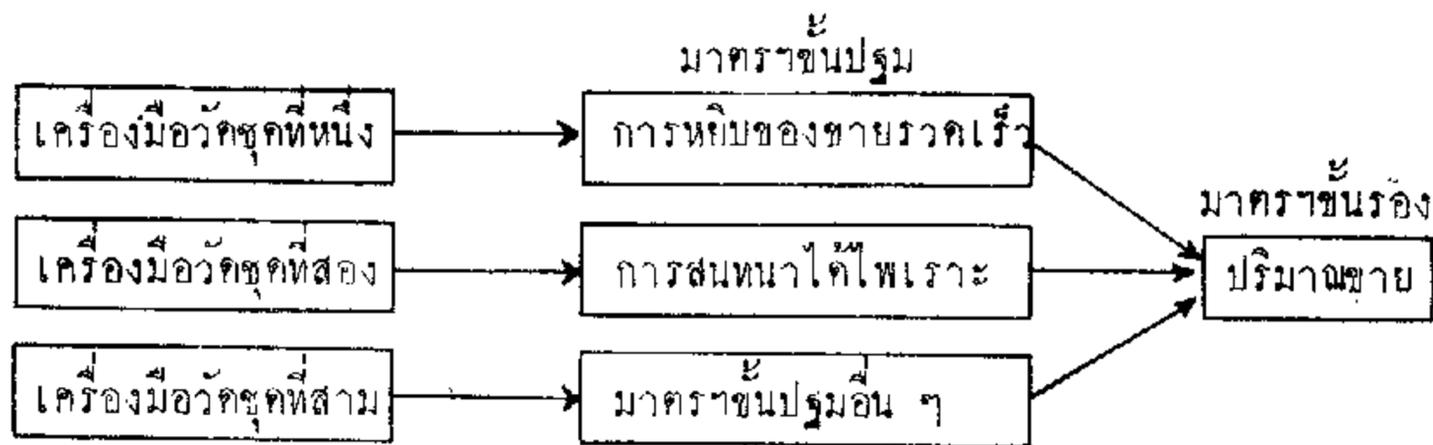
ข้อมูลเกี่ยวกับผู้ปฏิบัติงานที่สามารถรวบรวมได้ในทันทีถือว่าเป็น มาตรฐานชั้นต้น เช่น ความเห็นของผู้บังคับบัญชาที่ได้จากการประเมินในทันทีที่ผู้ปฏิบัติงานเริ่มรับตำแหน่งหน้าที่หรือ ผลการทดลองปฏิบัติงานในระยะเวลาสั้น ๆ ข้อมูลเหล่านี้เป็นเพียงส่วนหนึ่งของข้อเท็จจริงที่ว่า ผู้ปฏิบัติงานได้ทำงานในหน้าที่ของตนได้ดีเพียงไร ต่อเมื่อได้มีการวิเคราะห์เปรียบเทียบข้อมูล เหล่านี้กับผลการปฏิบัติงานในช่วงระยะยาว จึงจะสามารถสร้างมาตรฐานที่มีคุณภาพดีขึ้น เพื่อทำ ให้เป็นมาตรฐานชั้นกลาง ต่อไปดังกล่าวมาแล้ว

มาตรวัดความสำเร็จของงานขั้นปฐมและขั้นรอง (primary and secondary criteria) Guion¹⁵ ได้อ้างผลการค้นคว้าวิจัยของนักวิชาการว่า 10 คน ที่ยืนยันว่ามาตรวัดความสำเร็จของงาน พิจารณาได้เป็นหลายมิติ (dimensions) ฉะนั้นการที่ Thorndike กำหนดให้มีมาตรฐานขั้นสุดท้ายจึง เป็นการขัดต่อความเป็นจริงในเรื่องมิติ อย่างไรก็ตามก็ตีความคิดในเรื่องมาตรฐานขั้นสุดท้ายมิได้จำกัด ว่าผู้สร้างมาตรฐานจะต้องมองมาตรฐานขั้นสุดท้ายนี้ให้เป็นมิติเดียว ฉะนั้นจึงกล่าวในที่นี้ได้ว่า ความ เห็นเรื่องมาตรฐานขั้นสุดท้ายนี้ยังลบล้างไม่ได้ ส่วนความเห็นในเรื่องมิติเป็นเรื่องที่น่าสนใจและ ควรนำมาพิจารณาประกอบกันไปใน การสร้างมาตรฐานด้วย ในแง่มิติ มาตรฐานประกอบด้วยหลายมิติ ทั้งในแง่เวลาที่ปฏิบัติงานและในแง่ตัวบุคคลแต่ละคนที่รองรับตำแหน่งหน้าที่ ในแง่เวลาอาจแบ่ง เป็นสองระยะ เช่น ระยะเริ่มปฏิบัติงานและระยะเมื่อปฏิบัติงานไปแล้วช่วงเวลาหนึ่ง ผลิผล

¹⁵ Guion, *op.cit.*, pp. 114-115.

ของงานในสองระยะนี้ต้องต่างกัน ในแง่ตัวบุคคลแต่ละคนบ้างเพื่อบุคคลต่างกัน เช่น พนักงานขายของในตำแหน่งเหมือนกัน คนหนึ่งหยิบของให้ลูกค้าได้รวดเร็ว แต่อีกคนหนึ่งทำให้ลูกค้าพอใจด้วยการสนทนากับลูกค้าได้ไพเราะ มิตินี้เหล่านี้ทำให้ต้องมีเครื่องมือวัดหลาย ๆ อย่างเพื่อให้สามารถคาดคะเนผลการปฏิบัติงานในแง่ต่าง ๆ ได้

การที่ความสำเร็จของงานที่มีมิติหลายมิติเช่นนี้ กล่าวได้อีกอย่างหนึ่งว่าผู้ปฏิบัติงานแต่ละคนอาจมีพฤติกรรมไม่เหมือนกัน แต่ก็สามารถปฏิบัติงานให้ได้ผลสำเร็จเหมือนกัน ในตัวอย่างพนักงานขายในย่อหน้าก่อน ถ้าพนักงานทั้งสองคนจะทำให้บริษัทได้รับประโยชน์ในทางธุรกิจ จำนวนเป็นปริมาณเงินได้เท่า ๆ กัน พฤติกรรมของพนักงานขายที่ต่างกันเช่นนี้ก็ถือได้ว่าเป็นพฤติกรรมที่ทำให้ปฏิบัติงานสำเร็จได้ผลดี จึงแยกได้ว่าพฤติกรรมที่แตกต่างกันแต่ละบุคคลในระหว่างผู้ปฏิบัติงานเป็นมาตรฐานชั้นปฐม และประโยชน์ในทางธุรกิจของบริษัทเป็นมาตรฐานชั้นรอง มาตรฐานชั้นปฐมอาจมีได้ตั้งแต่สองขึ้นไป ทำให้ต้องมีเครื่องมือวัดตามจำนวนของมาตรฐานที่มีอยู่ ลักษณะของความสัมพันธ์ระหว่างมาตรฐานทั้งสองประเภทกับเครื่องมือวัดมีแสดงไว้ในแผนภาพที่ 4



แผนภาพที่ 4 มาตรฐานชั้นปฐมและชั้นรอง

ไม่ว่าจะใช้แนวคิดของ Thorndike หรือของ Guion ประเด็นสำคัญก็อยู่ตรงที่ว่า จะประเมินอะไรจึงจะได้ค่าเป็นตัวแทนความสำเร็จของงานอันแท้จริง สิ่งที่จะต้องปฏิบัติก็คือ ต้องมีการวิเคราะห์งาน (job analysis) ซึ่งทำให้ทราบสาระสำคัญของงานในตำแหน่งหน้าที่หนึ่ง ๆ เมื่อได้วิเคราะห์งานแล้ว จึงจะเก็บข้อเท็จจริง ในการปฏิบัติงานมาเทียบค่าประกอบการสร้างมาตรฯ ต่อไป ตามที่ได้ถือปฏิบัติกันมา การเก็บข้อเท็จจริงนี้อาจทำได้ในวิธีต่าง ๆ คือ เก็บข้อมูลโดยวิธีปรนัย ทั้งนี้อาจทดสอบความรู้และความเข้าใจงาน หรือหาหลักฐานการปฏิบัติงาน

หรือสังเกตตัวอย่างงาน มิฉะนั้นก็อาจถือเอาข้อวินิจฉัยส่วนบุคคลเป็นข้อมูลในการสร้างมาตร ฯ

มาตรวัดความสำเร็จของงานที่ได้โดยวิธีปรนัย

มาตรที่นำมาใช้ควรจะต้องเป็นมาตรที่ได้มาโดยวิธีปรนัย เพื่อให้ค่าของมาตรนั้นไม่บิดเบือนหรือผันผวนตามอารมณ์หรือความรู้สึกของผู้ประเมินค่าพฤติกรรมการทำงาน หรือกล่าวอีกนัยหนึ่งว่า หากได้มาตร ฯ มาโดยวิธีปรนัยย่อมมีโอกาสที่จะได้มาตรที่มีความเชื่อมั่นสูงกว่ามาตร ฯ ที่ได้โดยวิธีอื่น

แต่เป็นที่น่าเสียดายว่ามาตร ฯ ที่จะได้มาโดยวิธีปรนัยนั้นมีลักษณะจำกัดที่ไม่สามารถเป็นตัวแทนลักษณะทุกอย่างของพฤติกรรมการทำงาน ข้อเสียในเรื่องความจำกัดนี้จะพิจารณาเห็นได้จากลักษณะของข้อมูลที่จะได้มาโดยวิธีนี้ ข้อมูลที่จะนำมาพิจารณาประกอบเป็นมาตร ฯ ประเภทนี้แยกได้เป็น 3 ประการคือ ๑. คะแนนจากการทดสอบความรู้และความเข้าใจงาน (examination of knowledge and understanding) หลักฐานการปฏิบัติงานที่เป็นจริง (objective performance record) และผลจากการสังเกตตัวอย่างงาน (observer-scored job performance)

คะแนนการทดสอบความรู้และความเข้าใจงาน ทำเป็นส่วนใหญ่มิได้ใช้ข้อมูลประเภทนี้ไว้ในจำพวกที่จะใช้พิจารณาหามาตรวัดความสำเร็จของงาน ทั้งนี้อาจเป็นเพราะความรู้และความเข้าใจงานนั้นมีลักษณะห่างจากการปฏิบัติงานให้สำเร็จ จึงไม่อาจถือว่ามีสัมพันธ์กับมาตร ฯ ขั้นสุดท้ายได้อย่างแท้จริง แต่ Thorndike¹⁶ ก็ยังอ้างว่าข้อมูลประเภทนี้อาจนำมาใช้สร้างมาตร ฯ ได้ และมาตร ฯ นี้จะมีความเชื่อถือได้ว่าไม่บิดเบือนเพราะอคติของผู้ประเมิน และโดยเฉพาะงานในตำแหน่งหน้าที่ที่เป็นธรรม (abstract) ควรจะพิจารณาเทียบหามาตร ฯ ด้วยข้อมูลประเภทนี้อย่างยิ่ง เพราะเป็นข้อมูลที่ให้ค่าความสำเร็จของงานอย่างใกล้เคียงที่สุด การทดสอบความรู้และความเข้าใจงานอาศัยข้อสอบปรนัยเป็นหลัก ลักษณะและวิธีการสร้างข้อสอบอาจนำเอาหลักการของการทดสอบความสัมฤทธิ์มาใช้ได้ (achievement test) ในการนี้ ผู้สร้างข้อสอบวัดความรู้และความเข้าใจงานจะต้องแน่ใจว่าเนื้อหาของข้อสอบครอบคลุมความรู้ทั้งหมดที่ผู้ปฏิบัติงานที่รู้งานดีจะต้องมีตามสัดส่วนที่แท้จริง คำถามแต่ละข้อจะต้องถามประเด็นที่สำคัญ และเรารู้ว่าประเด็นไหนสำคัญนี้จะต้องเป็นที่ตกลงกันระหว่างผู้ออกข้อสอบ ฯ และหัวหน้างาน หรือผู้เชี่ยวชาญในตำแหน่งงานนั้น ๆ การที่จะทำข้อสอบนี้ให้สมบูรณ์ได้จึงต้องอาศัยความร่วมมือของบุคคลสามฝ่าย คือ นักจิตวิทยาในการวัดงานบุคคล นักวิเคราะห์งาน หรือผู้เชี่ยวชาญในตำแหน่ง

¹⁶ Thorndike, *op. cit.*, pp. 145-149.

ตำแหน่งงาน และหัวหน้างาน หรือผู้บังคับบัญชาโดยตรง นักจิตวิทยาในการวัดงานบุคคลมีหน้าที่นำความเห็นและข้อวิพากษ์วิจารณ์ของนักวิเคราะห์งาน และหัวหน้างานมาดำเนินการตั้งหัวข้อคำถามตามแบบของการทดสอบที่ดี

หลักฐานการปฏิบัติงานที่เป็นจริง ข้อมูลประเภทนี้ได้แก่บันทึกข้อเท็จจริงเกี่ยวกับผลการปฏิบัติงานที่เก็บไว้เป็นหลักฐานประจำตัวผู้ปฏิบัติงาน ข้อเท็จจริงนี้ต้องเป็นข้อเท็จจริงที่นับได้ หรือสังเกตได้ (countable or observable)¹⁷ ฉะนั้น พฤติกรรมในการปฏิบัติงานที่รวบรวมมาเป็นข้อมูลประเภทนี้จึงเป็นเพียงพฤติกรรมบางด้านเท่านั้น ไม่ใช่พฤติกรรมทั้งหมด และเป็นพฤติกรรมของงานบางตำแหน่ง ไม่ใช่พฤติกรรมของงานทุก ๆ ตำแหน่ง พฤติกรรมที่ได้มาเป็นข้อมูลประเภทนี้อาจแยกออกได้เป็น 2 ประเภทคือ ข้อเท็จจริงเกี่ยวกับการผลิตผลงาน และข้อเท็จจริงในก้านงานบุคคล

ข้อเท็จจริงในเรื่องผลิตผลงานหมายถึง ปริมาณและคุณภาพของงานที่ทำได้ รวมทั้งความสามารถในการรับการฝึกอบรม (trainability) ข้อเท็จจริงในเรื่องนี้ต้องเป็นข้อเท็จจริงที่นับได้และแสดงออกมาเป็นตัวเลขได้ ตำแหน่งปริมาณของผลงาน ผู้สร้างมาตรฯ อาจนับจำนวนชิ้นของงานผู้ที่ปฏิบัติงานทำสำเร็จในหนึ่งชั่วโมงหรือ วัน หรือสัปดาห์ เช่น ปริมาณปอที่คัดเลือกในหนึ่งชั่วโมง ปริมาณกระดาษที่พิมพ์ดีดในหนึ่งวัน หรือปริมาณการขายของพนักงานขายในหนึ่งปี หรืออาจพิจารณาปริมาณของผลงานในแง่เวลาก็ได้ กล่าวคือพิจารณาปริมาณเวลาที่ต้องใช้ในการปฏิบัติงานชิ้นหนึ่ง ๆ เช่น ปริมาณเวลาที่ต้องใช้ในการดำเนินการโครงการวิจัยหนึ่ง ๆ หรืออาจพิจารณาในด้านคุณภาพซึ่งหมายถึงว่าผู้ปฏิบัติงานคนหนึ่ง ๆ ปฏิบัติงานได้ดีเพียงไร แทนที่จะพิจารณาในแง่ที่ว่าแต่ละคนปฏิบัติงานได้มากและได้เร็วเพียงใด แต่การวัดว่าบุคคลปฏิบัติงานได้ดีเพียงไรเป็นเรื่องทำได้ยาก ฉะนั้น ในการลงบันทึกข้อเท็จจริงเกี่ยวกับคุณภาพของงานจึงเลี่ยงพิจารณาไปในแง่ที่ว่า การปฏิบัติงานหนึ่ง ๆ สอดคล้องกับอุปสงค์ของงานเพียงใด (job demands)¹⁸ กล่าวคือ วัดว่ามีชิ้นงานส่วนใดบ้างที่ใช้ไม่ได้ หรือต้องปฏิบัติใหม่ ค่าใช้จ่ายที่ต้องเสียไปในการทำงานที่ใช้ไม่ได้เสียใหม่เป็นปริมาณเท่าใด เช่นคุณภาพงานของพนักงานพิมพ์ดีดก็หมายถึงจำนวนคำที่พิมพ์ผิด คุณภาพของสมุหบัญชีหมายถึง ปริมาณข้อผิดพลาดที่ตรวจพบในระยะเวลาหนึ่ง ๆ ส่วนความสามารถในการรับการฝึกอบรมนั้นเป็นผลการ

¹⁷ Lawshe and Balma, *op. cit.*, pp. 35-37.

Guion *op. cit.*, p. 91.

¹⁸ *Ibid.*, p. 92.

บันทึกความเป็นจริงว่าผู้ปฏิบัติงานมีความเร็วในการเรียนรู้งานเพียงไร ความสามารถนี้คือ ปริมาณเวลาที่ผู้ปฏิบัติงานต้องใช้ในการที่จะสร้างความรู้ ความสามารถ (proficiency) ให้แก่ตนเองในการปฏิบัติงานหนึ่ง หรือกล่าวอีกนัยหนึ่งก็คือ ระดับความรู้ความสามารถของผู้ปฏิบัติงาน ในระยะเวลาที่กำหนดให้ช่วงหนึ่ง ๆ ในงานที่อุปสงค์และหน้าที่ความรับผิดชอบค่อนข้างตายตัว (static) ระยะเวลาที่ผู้ปฏิบัติงานต้องใช้ในการสร้างความรู้ความสามารถมีปริมาณค่อนข้างน้อย งานประเภทนี้ได้แก่งานง่าย ๆ ที่ต้องทำซ้ำ ๆ ซาก ๆ เช่น งานขับรถรับจ้างใช้เวลาเพียงประมาณ 18 สัปดาห์ ก็สามารถมีความรู้ความสามารถในการปฏิบัติงานเป็นอย่างดี (หลังจากขับรถเป็นแล้ว) แต่งานที่มีความสลับซับซ้อนรวมทั้งงานที่ใช้สมองจะต้องใช้เวลานานกว่า จะมีความรู้ความสามารถเต็มที่ เช่น ผู้ปฏิบัติหน้าที่พนักงานขายเพื่อการลงทุนธุรกิจ (investment salesman) สามารถเพิ่มความรู้ความสามารถในการปฏิบัติงานได้เรื่อย ๆ ไปจนถึงสิบปี¹⁹ ข้อมูลในตำนานนี้จะช่วยให้ทราบว่าผู้ปฏิบัติงานที่สำเร็จด้วยดีใช้เวลาเท่าใดจึงจะมีความรู้ความสามารถเต็มที่ ทำให้ผู้สร้างมาตรฐาน นำข้อเท็จจริงนี้มาเป็นบรรทัดฐานสำหรับทดสอบเครื่องมือวัดของตน อย่างไรก็ตาม มีข้อควรระวังที่ว่า ข้อมูลในเรื่องความสามารถในการรับการฝึกอบรมมิใช่เป็นข้อมูลประเภทเดียวกับความรู้ความสามารถในการผลิตผลงาน

ข้อเท็จจริงในตำนานบุคคล หมายถึงพฤติกรรมในการปฏิบัติงานในส่วนที่ไม่เกี่ยวกับเนื้อหาของงานโดยตรง แต่พฤติกรรมประเภทนี้ก็สามารถแสดงได้ว่า ผู้ปฏิบัติงานที่ถือได้ว่าดี มีคุณลักษณะเช่นใด พฤติกรรมนี้ได้แก่การขาด การลา ระยะเวลาตำรงตำแหน่ง อัตราความเร็วในการเลื่อนขั้น เลื่อนตำแหน่ง และอุบัติเหตุในระหว่างการทำงาน ในเรื่องการขาด การลา อาจลงบันทึกหลักฐานได้ว่าผู้ใดมาปฏิบัติงานมากที่สุด ผู้ใดมาน้อยที่สุด หรืออาจลงบันทึกหลักฐานในแง่จำนวนชั่วโมงและวันที่ขาดลา หรือจำนวนครั้งที่ขาดลาก็ได้ หลักฐานการขาด การลา นี้ มีประโยชน์ในแง่ที่ว่า *ceteris paribus* ผู้ที่มาปฏิบัติงานประจำโดยสม่ำเสมอมีประโยชน์ต่อหน่วยงานมากกว่าผู้ที่ขาด ๆ ลา ๆ ข้อเท็จจริงเรื่องระยะเวลาตำรงตำแหน่งทำให้ทราบว่าผู้ใดบ้างปฏิบัติงานกับหน่วยงานหนึ่ง ๆ ได้ยั่งยืนหรือไม่ยั่งยืนเพียงใด ข้อมูลนี้จะช่วยให้ผู้สร้างมาตรฐาน มีบรรทัดฐานที่จะกล่าวได้ว่าผู้ใดค่าคะแนนตามเครื่องมือวัดจะปฏิบัติงานกับหน่วยงานอย่างน้อยก็ช่วงเวลาหนึ่งที่กำหนดไว้หรือไม่ อัตราความเร็วในเรื่องการเลื่อนขั้นเลื่อนตำแหน่งช่วยให้ทราบว่าบุคคลต่าง ๆ ที่ปฏิบัติงานหนึ่งๆ ประสบความมีโอกาสที่จะได้เลื่อนขั้น

¹⁹ *Ibid.*, p. 93.

เลื่อนเงินเดือนและตำแหน่งอย่างใด (promotability) ความมีโอกาสนี้เป็นข้อเท็จจริงในประเด็นที่ว่าแต่ละคนต้องเสียเวลาเท่าใดจึงจะได้รับการเลื่อน และสามารถนำเอาไปเทียบกับเครื่องมือวัด ทำให้แยกแยะได้ว่า ผู้ถูกวัดจะมีโอกาสได้รับการเลื่อนในระยะเวลาที่กำหนดไว้หรือไม่ ส่วนอุบัติเหตุในระหว่างการปฏิบัติงานนั้นเป็นหลักฐานที่อาจเทียบกับเครื่องมือวัดที่สร้างขึ้นเพื่อให้แยกแยะได้ว่าผู้ใดมีโอกาสประสบอุบัติเหตุมากน้อยเพียงใด (accident proneness)

โดยหลักการ หลักฐานการปฏิบัติงานที่เป็นจริงย่อมใช้เป็นข้อมูลในการสร้างมาตราวัดที่ดี เพราะมาตราวัดขั้นสุดท้ายก็คือ ผลการที่ปฏิบัติงานที่ตนเอง ข้อเท็จจริงที่บันทึกไว้ย่อมมีความเป็นปรนัย (objectivity) มากกว่า ผลการปฏิบัติที่ประเมินได้ด้วยวิธีอื่น ๆ เพราะได้เก็บประมวลไว้เป็นระยะยาวและรวบรวมจากความเป็นจริง แต่ข้อมูลนี้มีข้อจำกัดอยู่ 2 ประการคือ ในเรื่องผลิตผลงานโดยตรงนั้น ข้อเท็จจริงที่บันทึกเป็นข้อมูลเฉพาะที่นับได้เท่านั้น แต่ในงานต่าง ๆ นั้นก็มีงานที่มีผลงานที่นับไม่ได้อยู่หลายอย่าง อีกประการหนึ่งข้อเท็จจริงนี้มีตัวแปรหลายอย่าง (variables) ที่ทำให้นำมาแปลความหมายเปรียบเทียบกันเพื่อถือเป็นบรรทัดฐานได้ยากตัวแปรเหล่านี้ได้แก่สภาพการณ์ต่าง ๆ ที่เกี่ยวข้องกับกรปฏิบัติงานหนึ่ง ๆ กล่าวคือ ผลิตผลการปฏิบัติงานของบุคคลหนึ่ง ๆ นั้นไม่ขึ้นอยู่กับความสามารถและความตั้งใจพยายามของผู้ปฏิบัติงานแต่ฝ่ายเดียว เช่น ผู้ปฏิบัติงานที่ต้องใช้เครื่องจักร คุณภาพและความเร็วของเครื่องจักรมีส่วนในการกำหนดปริมาณงานที่บุคคลหนึ่ง ๆ จะทำสำเร็จ ในกรณีที่ปฏิบัติงานกันเป็นคณะหรืองานที่ปฏิบัติต้องมีความเกี่ยวข้องกันหลายฝ่าย การปฏิบัติของผู้อื่นก็มีผลกระทบต่อปริมาณและคุณภาพของผลิตผลงานของผู้ปฏิบัติแต่ละคน

ผลจากการสังเกตตัวอย่างงาน ในโอกาสที่ผู้สร้างมาตราวัด ไม่สามารถค้นหาหลักฐานการปฏิบัติงานที่เป็นจริงได้ และเห็นว่าข้อตัดสินและความเห็นของผู้บังคับบัญชาขาดความเป็นปรนัยเชื่อถือไม่ได้ จึงจะกล่าวรายละเอียดถึงในหัวข้อต่อไป ผู้สร้างมาตราวัด ก็อาจสุ่มตัวอย่างของงานที่ปฏิบัติในตำแหน่งหนึ่ง ๆ และไปทำการสังเกตด้วยตนเอง ในระหว่างสังเกตผู้สร้างมาตราวัด ก็จะให้ค่าคะแนนของการปฏิบัติงานไปด้วย หลักการที่สำคัญในเรื่องนี้คือ การที่ผู้สร้างมาตราวัด ได้ใช้ความระมัดระวังในการเลือกตัวอย่างที่จะไปสังเกตให้เป็นส่วนหนึ่งของงานที่บุคคลทุกคนที่ปฏิบัติงานในหน้าที่นี้ต้องทำเสมอเหมือนกัน และเมื่อได้ทำเช่นนั้นแล้ว ในระหว่างการสังเกตก็จะต้องพยายามกำหนดมาตรฐานในการให้ค่าคะแนน การปฏิบัติงานเพื่อให้ได้ผลที่มีความเป็นปรนัยมากที่สุด²⁰ ในการปฏิบัติจริง ๆ นั้น ผู้สร้างมาตราวัด ย่อมไม่สามารถหลีกเลี่ยง

²⁰ Thorndike, *op. cit.*, pp. 141-145.

Lawshe and Balma, *op. cit.*, pp. 51-52.

เสียงอุปสรรคทั้งสองประการนี้ได้เป็นอย่างดี ในกรณีที่เป็นงานง่าย ๆ ปฏิบัติไปตามคัลลอง (routine) ก็อาจเลือกตัวอย่างงานที่เป็นตัวแทนจริง ๆ ของการปฏิบัติงานได้โดยสะดวก และยังสามารถกำหนดคุณค่าความสำคัญของการปฏิบัติงานส่วนนั้นในขั้นตอนต่าง ๆ ได้โดยมีอคติของผู้สร้างมาตร ฯ ไปปะปนเป็นส่วนน้อย แต่ในกรณีที่เป็นการสลับซับซ้อน นอกจากจะมีความลำบากในการกำหนดมาตรฐานการปฏิบัติงานแล้ว ผู้สังเกตยังไม่สามารถทำให้เป็นเสมือนเครื่องจักรที่บันทึกผลการปฏิบัติงานส่วนนั้น ๆ ในขั้นตอนต่าง ๆ โดยไม่มีความลำเอียง เพราะในการให้ค่าของการปฏิบัติ ผู้สร้างมาตร ฯ จำต้องแปลความหมายของการกระทำต่าง ๆ ในการปฏิบัติงานโดยอาศัยดุลยพินิจของตน ทำให้ค่าที่วัดออกเป็นผลงานมีน้ำหนักต่าง ๆ กันไปตามแต่ผู้สร้างมาตร ฯ แต่ละคน เช่น ถ้าจะสังเกตตัวอย่างงานของคนขับรถโดยสารสาธารณะ ก็อาจทำได้โดยสะดวก ผู้สร้างมาตร ฯ อาจกำหนดตัวอย่างงานไว้ว่ามีการขับถอยหลัง การเลี้ยว การติดเครื่อง และการเปลี่ยนเกียร์บนทางชัน ความนุ่มนวลและการปฏิบัติในเรื่องเหล่านี้โดยไม่บกพร่องในระดับต่าง ๆ ก็จะทำให้ผู้สร้างมาตร ฯ สามารถกำหนดค่าเป็นตัวเลขได้ทันที แต่การปฏิบัติงานในหน้าที่ผู้ตรวจการหรืองานอื่นที่มีความซับซ้อนมากกว่าย่อมทำให้ข้อมูลประเภทนี้มีความเที่ยงตรงและความเชื่อมั่นลดลงไปตามส่วน

มาตรวัดความสำเร็จของงานที่ได้จากการวินิจฉัยส่วนบุคคล

เพื่อที่จะให้ได้ทราบพฤติกรรมทุกด้านที่ปรากฏในการปฏิบัติงาน ผู้สร้างเครื่องมือวัดจำเป็นต้องอาศัย ภาพประทับใจ (impression) ที่บุคคลที่สามมีต่อผู้ปฏิบัติงาน บุคคลที่สามนี้อาจจะเป็นผู้บังคับบัญชาของผู้ปฏิบัติงานโดยตรง หรืออาจเป็นผู้ที่มีหน้าที่ประเมินความประพฤติและการปฏิบัติของผู้ดำรงตำแหน่งหน้าที่อื่น ๆ ก็ได้ ภาพประทับใจนี้จำเป็นต้องเป็นสิ่งที่มีค่าออกมาเป็นปริมาณได้ ในการนี้จึงได้มีวิธีการให้ค่า (rating scale) ในรูปต่าง ๆ กัน วิธีการให้ค่านี้นกำหนดลักษณะนิสัยหนึ่ง ๆ ของผู้ปฏิบัติงานเป็นตัวหลักที่ผู้ปฏิบัติงานคนหนึ่ง ๆ จะได้รับคะแนนสูงหรือต่ำตามแต่ลักษณะการประพฤติและปฏิบัติตนของผู้ปฏิบัติงานนั้น ๆ วิธีการให้ค่านี้นทำได้ในหลายลักษณะ ลักษณะวิธีการให้ค่าที่ทำได้สะดวกและแพร่หลายมากคือ วิธีการให้ค่าแบบกราฟ (graphic rating scale)²¹ ดังตัวอย่างในแผนภาพที่ 5

²¹ Thorndike and Hagen, *op. cit.*, pp. 350-351.
Guion, *op. cit.*, pp. 97-98.

ลักษณะพฤติกรรม	ค่าพฤติกรรม			
	ดีกว่าคนอื่นเสมอ	ดีกว่าคนอื่นบางครั้ง	เฉลี่ยแล้วปานกลาง	เฉลี่ยแล้วใช้ไม่ได้
ความรับผิดชอบ				
ในการ				
ทำงานให้เสร็จ
ความแม่นยำ
ความขยันหมั่นเพียร
ความปราณีต

แผนภาพที่ 5 ตัวอย่างวิธีการให้ค่าแบบกราฟ

ข้อบกพร่องของวิธีการให้ค่า เนื่องจากวิธีการให้ค่าเป็นวิธีที่ทราบกันแพร่หลายในการบริหารงานบุคคลและมีการกล่าวไว้ในที่ต่าง ๆ มากมายหลายแห่งแล้ว เพื่อไม่ให้เป็นการซ้ำซากในที่นี้จึงจะไม่กล่าวถึงรายละเอียดของวิธีการนี้ แต่จะพิจารณาเฉพาะข้อเสียของการใช้วิธีนี้ในการประเมินบุคคล เมื่อทราบข้อบกพร่องของวิธีนี้แล้ว จึงจะได้เปรียบเทียบผลการค้นคว้าวิจัยในการที่จะแก้ไขวิธีนี้ให้ได้ผลในการที่จะได้พฤติกรรมที่แท้จริงในการปฏิบัติงาน

ข้อเสียของวิธีการให้ค่าอยู่ตรงที่ความบกพร่องของผู้ให้ค่า (rater) ความบกพร่องนี้แยกได้เป็น 2 นัย คือ ผู้ให้ค่าอาจขาดความตั้งใจที่จะให้ค่าอย่างตรงไปตรงมา หรือผู้ให้ค่าอาจขาดความสามารถที่จะให้ค่าได้อย่างเที่ยงตรง²²

ในเรื่องการขาดความตั้งใจที่จะให้ค่าอย่างตรงไปตรงมา เป็นปัญหาที่พิจารณาแล้วดูจะขัดความเป็นจริง เพราะย่อมเป็นที่เข้าใจกันว่า ผู้ให้ค่าแต่ละคนย่อมพยายามอย่างที่สุดที่จะปฏิบัติตามระเบียบข้อบังคับในการให้คะแนน และพยายามที่จะให้ความเป็นธรรมแก่ผู้ปฏิบัติงานในตำแหน่งที่ถูกประเมินทุกคน ถ้าหากจะมีการผิดพลาดในการประเมินงานก็ต้องเป็นเพราะความบกพร่องส่วนบุคคลไม่ใช่ด้วยความตั้งใจที่จะไม่ให้ความยุติธรรม แต่ตามความเป็นจริงทางวิชาการนั้นผู้ให้ค่าก็อาจจะให้คะแนนงานที่ประเมินอย่างมีอคติด้วยความตั้งใจได้ การ

22 H. H. Rummers, et al., *A Practical Introduction to Measurement and Evaluation* (New York: Harper and Row, 1965), pp. 359-360.

Thorndike and Hagen, *op. cit.*, pp. 338-344

ที่มีการบิดเบือนในการให้ค่าคะแนนนี้ อาจเป็นเพราะเหตุหนึ่งเหตุใดจากสาเหตุสองประการคือ ประการแรก ผู้ให้ค่าไม่ยอมอุทิศตนที่จะลำบากในการดำเนินการอย่างละเอียดถี่ถ้วนในการให้ค่าคะแนน หรือประการที่สองผู้ให้ค่าคะแนนมีความสัมพันธ์กับผู้ปฏิบัติงานที่ถูกประเมินในลักษณะที่ทำให้ผู้ให้ค่าไม่ประสงค์ให้ผู้ถูกประเมินต้องเสียประโยชน์ เป็นความจริงที่ว่า การให้คะแนนในเรื่องใด ๆ ก็ตาม หากจะให้เป็นที่เที่ยงตรงจะต้องเสียเวลามาก เพราะจะต้องมีการวิเคราะห์พฤติกรรมในเรื่องนั้นออกมาเป็นประเด็นที่สำคัญแล้วกำหนดประเด็นเหล่านี้เป็นเกณฑ์ในการที่จะนำผู้ถูกประเมินมาเทียบกัน จะเห็นได้ว่าแบบกรอกรายการที่ใช้ประเมินข้าราชการนั้นมีค่าเป็นเพียงแบบฟอร์มเท่านั้น ผู้บังคับบัญชาที่มีหน้าที่ตอบแบบกรอกรายการนี้มิได้เสียเวลาจริง ๆ ในการพิจารณาเทียบค่าของพฤติกรรมในการปฏิบัติงาน ผลการประเมินจึงเป็นเพียงภาพประทับใจที่เป็นอัตนัยของผู้บังคับบัญชา มิได้วัดค่าการปฏิบัติงานจริง ๆ ตลอดปีของผู้ที่อยู่ใต้บังคับบัญชา เมื่อถึงขั้นตอนที่จะพิจารณาให้ความดีความชอบ จึงอาศัยการอภิปรายในที่ประชุมของผู้บังคับบัญชาเป็นหลัก มิได้ถือข้อเท็จจริงของแบบกรอกรายการนี้เป็นหลัก การที่เป็นเช่นนี้จึงทำให้มีผู้กล่าวว่า เป็นเพราะการพิจารณาความดีความชอบของข้าราชการอาศัยการประชุมดังกล่าว จึงไม่จำเป็นต้องเสียเวลารอกแบบรายการประเมินงานให้ละเอียดถี่ถ้วนเที่ยงตรง อย่างไรก็ตาม การประเมินพฤติกรรมของงานเป็นเรื่องที่จะต้องยอมอุทิศเวลาหากประสงค์จะให้ค่าประเมินเที่ยงตรง โดยผู้บังคับบัญชาจะต้องหมั่นประเมินผู้ที่อยู่ใต้บังคับบัญชาเป็นระยะ ๆ อยู่เสมอ และเมื่อถึงวาระที่จะต้องลงบันทึกให้ค่าประเมินก็สามารถทำได้โดยมิได้อาศัยความรู้สึกชั่วขณะหรือความลำเอียง แต่อาศัยหลักฐานการประเมินครั้งต่างๆ เปรียบเทียบกันส่วนใหญ่ผู้มีหน้าที่ให้ค่าพฤติกรรมการปฏิบัติงานก็ได้แก่ผู้บังคับบัญชาชั้นต้น ทั้งนี้เพราะเจ้าหน้าที่ฝ่ายงานบุคคลโดยตรงเช่น สำนักงาน ก.พ. หรือหัวหน้าฝ่ายการเจ้าหน้าที่ของบริษัทธุรกิจย่อมเป็นผู้ที่อยู่ห่างไกลผู้ปฏิบัติงานมากกว่าหัวหน้าแผนกหรือหัวหน้ากองของผู้ปฏิบัติงานนั้น ๆ ในกรณีเช่นนี้จึงเป็นเหตุให้เกิดผลเสียอีกประการหนึ่งคือ ในบางครั้ง ผู้บังคับบัญชาจะพยายามให้ค่าคะแนนแก่บุคคลในบังคับบัญชาสูงเพื่อไม่ให้ลดหน้าบุคคลที่อยู่ใต้บังคับบัญชาของผู้อื่น แทนที่จะพยายามให้ค่าคะแนนตรงกับพฤติกรรมของงาน และโดยเฉพาะในการที่การประเมินค่าพฤติกรรมของงานนี้เป็นเรื่องเกี่ยวข้องกับการเลื่อนขั้นอันดับเงินเดือนโดยตรง ก็ยังทำให้ผู้บังคับบัญชาต้องพยายามรักษาผลประโยชน์ของบุคคลในกลุ่มของตนในหน่วยงานทั่ว ๆ ไป ค่าคะแนนของพฤติกรรมในการปฏิบัติงานตำแหน่งต่างๆ จึงมักมารวมกลุ่มกันอยู่ในระดับใดระดับหนึ่ง ทำให้ไม่สามารถแยกแยะ

ออกไปได้ว่า พฤติกรรมในการปฏิบัติงานตำแหน่งใดควรจะได้รับการประเมินสูงกว่าตำแหน่งอื่น ๆ เท่าใด

ในค่านความสามารถของผู้ให้ค่าที่กล่าวว่ายังมีทางบกพร่องนั้น เป็นเพราะเหตุหนึ่ง เหตุใดในสาเหตุ 5 ประการดังต่อไปนี้ คือความกำกวม (ambiguity) ของลักษณะพฤติกรรมที่จะต้องประเมิน ลักษณะนิสัยที่จะให้ค่านั้นเห็นไม่ได้เด่นชัดจากการสังเกตภายนอก ผู้ให้ค่าไม่มีโอกาสหรือไม่ได้ใช้โอกาสเต็มที่ในการสังเกตพฤติกรรม เกณฑ์ที่จะให้ค่าไม่ตรงกันระหว่างผู้ให้ค่าและลักษณะพิเศษประจำตัว (idiosyncrasies) ของผู้ให้ค่าแต่ละคน²³

การที่กล่าวว่ายักษณะพฤติกรรมที่ประเมิน มีความกำกวมก็ เพราะว่าการประเมินโดยทั่ว ๆ ไป ผู้ให้ค่าจะต้องให้คะแนนรายการพฤติกรรมที่มีความหมายกว้าง ๆ เช่นที่แสดงไว้ในแผนภาพที่ 5 คำที่ว่า ความขยันหมั่นเพียร หรือความรับผิดชอบในการทำงานให้เสร็จ มีความหมายต่อผู้ให้ค่าแต่ละคนต่าง ๆ กันไป ที่ว่าขยันนั้นหมายถึงว่าอยู่ปฏิบัติหน้าที่ในสำนักงานตลอดเวลาราชการ และเมื่อผู้บังคับบัญชาเรียกตัวเมื่อเวลาบ่ายสี่โมงครึ่งก็ยังไม่พร้อมที่จะไปรายงานตัวได้ หรือหมายความว่าปฏิบัติหน้าที่ที่รับมอบหมายเสร็จเรียบร้อยก่อนเวลาทุกครั้ง ในกรณีที่ผู้ให้ค่าคือ ความหมายแรกเป็นเกณฑ์ บุคคลบางคนก็จะให้ค่าคะแนนสูง แต่เมื่อผู้ให้ค่าคือความหมายประการหลังผู้ที่ให้ค่าคะแนนสูงจะกลายเป็นบุคคลต่างประเภทออกไป สภาพการณ์เช่นนี้ย่อมทำให้การให้ค่าพฤติกรรมขาดความคงเส้นคงวา ปัญหาที่ว่าลักษณะนิสัยที่จะให้ค่าเห็นไม่ได้เด่นชัดจากการสังเกตภายนอก หมายความว่าถึงลักษณะนิสัยหลายอย่างที่ใช้เป็นเครื่องอธิบายพฤติกรรมเป็นสิ่งที่ซ่อนอยู่ในตัวผู้ปฏิบัติงาน อาจปรากฏท่าทางที่แสดงออกภายนอกเป็นเพียงเครื่องสะท้อนอย่างผิวเผินจะถือเอาเป็นลักษณะที่แท้จริงไม่ได้ เช่น ผู้ให้ค่าลักษณะความเชื่อมั่นในตัวเองของผู้ปฏิบัติงานจะถือเอาเพียงน้ำเสียงในการพูด และความสามารถในการสนทนาของบุคคลหนึ่งเป็นหลักในการตัดสินย่อมไม่สมควร ผู้ให้ค่าจะประเมินลักษณะนี้ได้ก็ต่อเมื่อได้คลุกคลีสังเกตผู้ปฏิบัติงานเป็นเวลานาน ๆ และอาศัยความรู้ทางจิตวิทยาแยกแยะความประพฤติของผู้ทำงาน จึงจะสรุปได้ว่าผู้ปฏิบัติงานนั้น ๆ มีลักษณะนิสัยอย่างไร ปัญหานี้ก็เกี่ยวโยงไปถึงข้อเท็จจริงอีกประการหนึ่งที่ว่าผู้ให้ค่าไม่มีโอกาสคลุกคลีสังเกตผู้ปฏิบัติงานอย่างแท้จริง ผู้บังคับบัญชาที่มีหน้าที่ให้ค่าพฤติกรรมในวงกว้างก็มีผู้อยู่ได้บังคับบัญชาหลายคน ถึงแม้จะมีเวลาปฏิบัติงานร่วมกันเป็นเวลาหลาย ๆ ชั่วโมงในวันหนึ่ง ๆ ผู้บังคับบัญชาก็มีภาระกิจอื่น ๆ ที่สำคัญกว่าที่จะต้องทำให้สำเร็จ ปัญหาจึงเกิดขึ้นว่า ผู้ให้ค่าหรือผู้บังคับบัญชามีโอกาสคลุกคลี

23 Thorndike and Hagen, *op. cit.*, pp. 340-344.

สังเกตผู้ปฏิบัติงานในสภาพงานต่าง ๆ กันจนครบทุกแง่ทุกมุมจนเชื่อได้ว่าการให้ค่าลักษณะนิสัยต่าง ๆ มีคุณค่าเชื่อถือได้อย่างแท้จริงหรือไม่ ถึงหากจะแก้ปัญหานั้นตามประการนี้ได้ ก็ยังมีปัญหาใหญ่ ๆ อื่น ๆ อีก คือ ผู้ให้ค่าจะต้องที่ค่าพฤติกรรมออกมาเป็นปริมาณซึ่งไม่สามารถพิสูจน์ว่าวัดได้ว่าตรงกันระหว่างผู้ให้ค่าเอง เช่น การให้ค่าลักษณะกันหนึ่งอันใดของพฤติกรรมจะระบุว่า ผู้ปฏิบัติงานมีคุณสมบัติตามลักษณะนั้น ๆ อยู่ในระดับใด ดีมาก หรือปานกลาง หรือใช้ไม่ได้ การที่ว่าดีมากนั้นหมายความว่าอย่างไร หมายความว่า เป็นที่ดีที่สุดในระหว่างผู้ปฏิบัติงานทั้งหมด หรือเป็นคนหนึ่งในสองสามคนที่ดีกว่าบุคคลที่เหลือ และที่ว่าดีมากนี้เทียบกับผู้ปฏิบัติงานทั้งหมดของแผนกหรือกอง หรือทั่ว ๆ ไป และได้เทียบโดยพิจารณาข้อได้เปรียบหรือเสียเปรียบของการที่บุคคลผู้นั้นได้ปฏิบัติงานโดยระยะเวลาที่แตกต่างกันและในสภาพการณ์ที่ไม่เหมือนกันกับบุคคลอื่น ๆ หรือเปล่า ปัญหานี้เป็นปัญหาการขาดเกณฑ์ที่จะให้ค่าที่ตรงกันระหว่างผู้ให้ค่า ปัญหาสุดท้ายเป็นปัญหาประจำตัวผู้ให้ค่าแต่ละคน กล่าวคือ ผู้ให้ค่าแต่ละคนต่างก็มีประสบการณ์ในชีวิตมาในลักษณะต่าง ๆ กัน ซึ่งทำให้แต่ละคนชอบหรือรังเกียจพฤติกรรมและปรากฏการณ์ต่าง ๆ ในลักษณะที่ไม่เหมือนกัน ประสบการณ์นี้ทำให้ผู้ให้ค่าคนหนึ่งถือว่าพฤติกรรมของบุคคลคนหนึ่งแสดงความอ่อนแอหรือความใจต่ำ ในขณะที่ผู้ให้ค่าอีกคนหนึ่งมีได้ถือเช่นนั้น กล่าวอีกอย่างหนึ่งได้ว่าผู้ให้ค่าแต่ละคนมีแนวโน้มที่จะมีอคติ (bias) ซึ่งมีผลกระทบต่อการประเมินรายการละเอียดของลักษณะนิสัยต่าง ๆ ที่ระบุไว้ในพฤติกรรมกรรมการปฏิบัติงาน

ข้อบกพร่องต่าง ๆ เหล่านี้ทำให้ผลการประเมินพฤติกรรมในการปฏิบัติด้วย การให้ค่าตามวิธีที่ใช้ทั่ว ๆ ไปเกิดผลเสียสามประการคือ ความบิดเบือนจากอคติ (halo) ความเบ้ในการให้คะแนน (leniency) และความโน้มสู่ศูนย์กลาง (central tendency) ความบิดเบือนจากอคติคือ การที่ผู้ให้ค่าถือความรู้สึกส่วนตนอันเกิดจากอคติที่มีต่อผู้ปฏิบัติงานเป็นแนวโน้มในการให้ค่าแก่ลักษณะนิสัยทุกอย่างของพฤติกรรม ความเบ้ในการให้คะแนนเป็นลักษณะการที่ผู้ให้ค่าอาจเป็น "คนกระตือรือร้น" หรือ "คนให้คะแนนง่าย" ทำให้ค่าที่ให้เกิดความเบ้ (skewed distributions of ratings) ส่วนความโน้มสู่ศูนย์กลางนั้นเป็นการที่ค่าคะแนนมักจะมารวมกระจุกกันอยู่ไม่แยกกระจาย รายละเอียดในเรื่องนี้มีกล่าวไว้ในตำราและเอกสารทางวิชาการหลายแห่งแล้วจึงจะไม่กล่าวซ้ำอีก

การประเมินด้วยการบังคับเปรียบเทียบบุคคล (employee comparisons) การที่วิธีการให้ค่าพฤติกรรมในการปฏิบัติงานตามวิธีทั่ว ๆ ไปมีความบกพร่องและผลเสียดังกล่าว ขั้วด้วยบุคคลหลายท่านเช่น S.N. Stevens, E.F. Wonderlic, R.K. Rosensteel, W.K. Kirchner, และ M.D. Dunette

เป็นต้น ได้พยายามค้นคว้าวิธีใหม่ เพื่อให้ได้มาตราฯ ที่ไม่มีข้อบกพร่องต่าง ๆ โดยกำหนดหลักการว่า เมื่อจะประเมินพฤติกรรมงาน ให้ประเมินทุก ๆ คนพร้อมกันที่ลักษณะนิสัย เพื่อที่จะให้ผู้ประเมินต้องมีการเปรียบเทียบบุคคลทั้งหมดจริง ๆ วิธีการใหม่นี้ก็เท่ากับเอาวิธีให้ค่าเดิมมาปรับปรุงใหม่นั้นเอง ตามหลักการใหม่นี้จะมีวิธีประเมินใหญ่ ๆ 2 วิธีคือ วิธีจัดลำดับที่ (ranking) และวิธีจัดบัญชีตรวจสอบ (checklist) ²⁴

วิธีจัดลำดับที่อาจทำได้ 3 อย่าง ประการแรก จัดลำดับที่สลับ (alternation ranking) อย่างที่สอง จัดลำดับที่โดยเทียบคู่ (paired comparisons) และ อย่างที่สาม จัดลำดับที่ตามกำหนดอัตราส่วน (forced distribution) การจัดลำดับที่สลับทำได้โดยผู้ประเมินทำบัญชีชื่อผู้ที่ถูกให้ค่าเป็นแบบสุ่ม (random order) แล้วนำชื่อเหล่านี้มาพิจารณาหาว่าผู้ใดดีที่สุด และผู้ใดเลวที่สุดในกลุ่ม พิจารณาไล่ไปที่ละคู่เช่นนี้ก็จะได้บุคคลที่สมควรได้รับการประเมินให้อยู่ในอันดับแรกไล่ลงไป และในขณะเดียวกันก็ได้บุคคลที่สมควรให้ได้รับการพิจารณาให้อยู่ในอันดับสุดท้ายไล่ขึ้นมา ในการจัดลำดับที่โดยเทียบคู่ทำได้โดยบุคคลแต่ละคนไปพิจารณาเทียบคู่กับบุคคลอื่น ๆ ทุกคนในกลุ่มทีละคน เช่น ถ้ามีบุคคลในกลุ่ม 10 คน ก็จะมีการจัดคู่ 45 คู่ หรือถ้ามีบุคคลในกลุ่ม 150 คนก็จะมี 11,175 คู่ การที่มีคู่มากเช่นนี้ทำให้วิธีการจัดลำดับที่ตามแบบนี้ต้องอาศัยเครื่องจักรกลคณนา (computer) ส่วนวิธีจัดลำดับที่แบบกำหนดอัตราส่วนนั้นอาศัยหลักการแจกแจงปกติ (normal distribution) กล่าวคือตามหลักสถิตินี้ทราบอยู่แล้วว่าปรากฏการณ์ของสิ่งต่างๆ จะเกิดขึ้นโดยมีอัตราส่วนสัมพันธ์กับค่าเบี่ยงเบนมาตรฐาน (standard deviation) จากเกณฑ์เฉลี่ย (mean) ออกไปทั้งข้างบวกและลบเป็นปริมาณ .5 ของค่าเบี่ยงเบนมาตรฐาน จะมีกรณีของปรากฏการณ์เกิดขึ้นเป็นอัตราประมาณร้อยละ 38.30 และเมื่อตัดค่าเบี่ยงเบนนี้ออกไปไม่ว่าจะเป็นทางข้างบวกหรือลบอีก .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) และประเภทที่ห้าร้อยละ 10 (เท่ากับประเภทที่หนึ่ง) ถ้าหากมีผู้ปฏิบัติงาน

²⁴ Guion, *op. cit.*, pp. 100-109.

50 คน ผู้ประเมินก็ต้องจัดให้มีบุคคลในประเภทต่าง ๆ จากสูงสุดมาต่ำสุด เป็นจำนวน 5, 10, 20, 10, 5 ตามลำดับ:

วิธีจัดลำดับที่เป็นวิธีประเมินที่ต้องใช้เวลามาก แต่ก็ได้ผลดีในการที่บังคับให้ผู้ประเมินต้องมีการวิเคราะห์ผู้ปฏิบัติงานที่ถูกประเมินอย่างแท้จริง ผู้ประเมินไม่สามารถจัดให้ผู้ถูกประเมินได้รับค่าคะแนนเท่า ๆ กันเป็นส่วนใหญ่ได้ ยิ่งไปกว่านั้นยังหลีกเลี่ยงความบิดเบือนจากอคติและความเบ้ในการให้คะแนนอีกด้วย กล่าวคือ ถึงแม้ผู้ประเมินจะเป็น “คนใจดี” ในการให้คะแนนอย่างไรก็ตามก็จำเป็นต้องจัดให้ผู้ถูกประเมินคนใดคนหนึ่งเป็นที่สุดท้าย และถึงแม้ว่าจะเป็น “คนกระตือรือร้น” ก็ต้องจัดให้มีคนได้อันดับที่หนึ่ง อย่างไรก็ตาม ปัญหาที่ยังเกิดขึ้นในแง่ที่ว่าลำดับที่แต่ละคนได้รับในแต่ละกลุ่มนั้นมีค่าเท่ากันหรือเหมือนกันระหว่างกลุ่มหรือไม่ ผู้ที่ได้ที่สามในแผนกที่มีผู้ปฏิบัติงานเพียง 3 คน ย่อมมีค่าต่างจากผู้ที่ได้สาม ในแผนกที่มีผู้ปฏิบัติงาน 20 คน นอกจากนี้ช่วงห่างระหว่างลำดับที่แต่ละที่มีได้มีความห่างเป็นปริมาณเท่า ๆ กัน ฉะนั้นในการประเมินด้วยวิธีจัดลำดับที่นี้จึงได้มีการเปรียบเทียบระหว่างกลุ่มที่ถูกประเมินด้วยการเปรียบเทียบนี้ทำได้โดยนำบุคคลจำนวนหนึ่งในกลุ่มมาจัดให้เป็นบุคคลหลัก (key men) บุคคลหลักนี้จะต้องเป็นผู้ที่ผู้ประเมินทุกคนของกลุ่มต่าง ๆ รู้จักเป็นอย่างดีด้วย ลำดับที่ของบุคคลต่าง ๆ ในกลุ่มจะต้องนำมาคำนวณเป็นค่ามาตรฐานปกติ (normalized standard scores) ค่าของบุคคลหลักนี้ก็จะได้เป็นค่าในการลากเส้นสหสัมพันธ์ถดถอย (regression line) เพื่อที่จะเทียบค่าของบุคคลทุก ๆ คนในกลุ่มกับคะแนนบรรทัดฐาน (master scale) ซึ่งกำหนดค่าคะแนนที่เทียบได้เหมือนกันหมดระหว่างบุคคลทุกคนในกลุ่มต่าง ๆ

วิธีจัดบัญชีตรวจสอบเป็นวิธีที่กำหนดให้มีรายการของลักษณะต่าง ๆ ที่ควรจะมีในพฤติกรรมกรรมการปฏิบัติงานทั้งหมด การสร้างรายการที่ดีว่าดีที่สุด คือ การสร้างโดยใช้เทคนิคเหตุการณ์วิกฤติ (critical incident technique) เทคนิคนี้มีหลักว่าผู้สร้างรายการจะต้องวิเคราะห์เหตุการณ์ที่มีความสำคัญที่สุดต่อการปฏิบัติงานให้ได้สำเร็จ หรือทำให้การปฏิบัติงานต้องล้มเหลว เช่นบัญชีตรวจสอบพฤติกรรมกรรมการปฏิบัติงานของพนักงานขายของบริษัท Mining and Manufacturing Company กำหนดรายการไว้ดังนี้²⁵

- ก. เกี่ยวข้องกับความลับของลูกค้า
- ข. ไม่มีความรู้เกี่ยวกับผลผลิตที่จะขาย

²⁵ *Ibid.*, p. 105.

- ค. สัญญากับลูกคำมากเกินไป
- ง. ให้ลูกคำติดหนึ่มากเกินไปจนลูกคำเองเกิดสงสัย
- จ. เมื่อลูกคำมีเรื่องไม่พอใจรีบคิดตามผลให้เป็นน่าพอใจ

วลว

การประเมินโดยวิธีตรวจสอบอาจทำได้หลายวิธีเช่น วิธีจัดบัญชีสรุป (summated ratings) วิธีจัดช่วงเสมอกัน (equal-appearing ratings) และวิธีบังคับตอบ (forced-choice ratings) เนื่องจากวิธีบังคับตอบเป็นวิธีที่มีชื่อเสียงกว่าวิธีอื่น จึงจะอธิบายเฉพาะวิธีนี้เท่านั้น ในที่นี้ วิธีนี้ถือหลักการเช่นเดียวกับวิธีจัดลำดับที่ในการพิจารณาเปรียบลักษณะนิสัยของพฤติกรรมของทุก ๆ คน พร้อมกัน ต่างกันแต่ว่าแทนที่จะพิจารณาลักษณะนิสัยทีละอย่างให้พิจารณาลักษณะนิสัยทีละชุดแทน (a set of attributes) โดยทั่วๆ ไปผู้ประเมินจะทำรายการที่สะท้อนให้เห็นภาพผู้ปฏิบัติงานที่ดีในตำแหน่งหนึ่ง ๆ เป็นชุดละ 4 รายการ (tetrad) โดยให้รายการในชุดหนึ่ง ๆ มีค่าเท่าเทียมกันในสายตาของผู้อธิบายคุณลักษณะของพฤติกรรมในการปฏิบัติงานทั่วๆ ไป แต่ในแต่ละชุดนี้จะมีอยู่ 2 รายการที่มีความสำคัญในการชี้ว่าผู้ปฏิบัติงานที่ดีนั้นต้องมีลักษณะนิสัยอย่างไร เช่นในบัญชีตรวจสอบซึ่งใช้ประเมินงานในตำแหน่งครูโรงเรียนการช่างของกองทัพอากาศสหรัฐ ฯ ได้กำหนดรายการชุดหนึ่งไว้ดังนี้

- ก. อุดหนุนที่จะสอนผู้เรียนซ้ำ
- ข. บรรยายด้วยความมั่นใจ
- ค. ทำให้นักเรียนสนใจและกระตือรือร้น
- ง. อธิบายให้นักเรียนเข้าใจความมุ่งหมายก่อนสอนทุกเรื่อง

แต่ละรายการในชุดจะเป็นคำอธิบายครูที่ดีในโรงเรียนทั้งสิ้น แต่รายการหนึ่ง ๆ จะมีค่าในการวัดว่าเป็นครูที่ดีในปริมาณต่างกัน โดยวิธีวิเคราะห์องค์ประกอบ (factor analysis) ซึ่งอาศัยหลักคณิตศาสตร์ทำให้ทราบว่า แต่ละรายการในชุดนี้มีค่าถรรชนีแยกคุณลักษณะ (discrimination index) เรียงลำดับดังต่อไปนี้ คือ +1.15, +.54, +1.39 และ +.79 ค่าถรรชนีทำให้สามารถให้ค่าแก่รายการแต่ละอันที่ผู้ประเมินเลือกกว่าคำอธิบายการปฏิบัติงานของผู้ถูกประเมิน เช่น อาจจะให้ค่า 2 แก่รายการข้อ ค. ในฐานะเป็นรายการที่มีค่าถรรชนีแยกคุณลักษณะสูงสุด และให้ค่าศูนย์แก่รายการข้อ ง. เนื่องจากมีค่าถรรชนี ๖ ค่าที่สุก ส่วนข้อ ก. และ ข. ได้ 1 เท่า ๆ กัน คะแนนรวมของรายการทุกชุดก็จะเป็นค่าประเมินของพฤติกรรมงานของผู้ถูกประเมินคนหนึ่ง ๆ การที่จัดรายการให้เป็นเช่นนี้มีเหตุผลที่ว่า ผู้ประเมินจะไม่เอนเอียง เพราะรายการทุกข้อก็ไม่มี

อะไรที่เลวร้ายเป็นพิเศษกว่ากัน ผู้ประเมินเองก็ไม่ทราบว่ารายการข้อใดจะเป็นประโยชน์หรือเป็นผลเสียต่อผู้ถูกประเมินอย่างไร อีกประการหนึ่งค่าคะแนนที่ได้จะมีลักษณะกระจายใกล้เคียงกับเกณฑ์เฉลี่ยตามปกติ (normal distribution) อย่างไรก็ตาม การจัดบัญชีตรวจสอบโดยวิธีบังคับตอบก็ยังมีปัญหาอยู่ 2 ประการ คือ ผู้ประเมินที่จำเป็นต้องใช้วิธีนี้ในการประเมินมักจะอ้างว่ามีความลำบากในการเลือกรายการ อีกประการหนึ่งการเปรียบเทียบลักษณะนิสัยซึ่งต่างๆ กันยังไม่มีวิธีการมาตรฐาน ทำให้ค่าที่ประเมินได้โดยวิธีนี้ขาดความหมายที่แท้จริงนำมาใช้อธิบายภาพผู้ปฏิบัติงานที่ถูกประเมิน

การประเมินมาตรฐานวัดความสำเร็จของงาน

ปัญหาสำคัญของเครื่องมือวัดในการบริหารงานบุคคลข้อหนึ่งคือ การที่ต้องวินิจฉัยว่าจะวัดอะไรจึงจะให้ข้อเท็จจริงที่แสดงพฤติกรรมในการปฏิบัติงานสำเร็จ ถ้าได้ข้อเท็จจริงนี้มากก็ทำให้มีมาตรฐานวัดความสำเร็จของงานที่ดี เป็นทางให้มาเทียบเท่ากับเครื่องมือวัดเพื่อสร้างเครื่องมือวัดที่ดีต่อไปได้ จากแนวคิดอันนี้ก็ย่อมหมาความว่ามาตร ๖ ควรเป็นปรนัยให้มาก และควรจะต้องสัมพันธ์สอดคล้องกับตัวแปรร่วมของข้อเท็จจริงเพื่อให้มีความเที่ยงตรง (validity) และความเชื่อมั่น (reliability) นอกจากนี้มาตร ๖ ที่ดียังต้องมีคุณภาพในการแยกแยะว่าผู้ที่ปฏิบัติงานสำเร็จด้วยดีนั้นดีเป็นระดับอย่างไร และมีความสะดวกและความประหยัดในการรวบรวมข้อมูล²⁶

ความเที่ยงตรง ศาสตร์หลักแนวคิดในการสร้างมาตร ๖ ดังได้กล่าวไว้ในตอนต้นบทความนี้ ผู้อ่านก็อาจจะเห็นได้ว่ามาตร ๖ จะมีความเที่ยงตรงได้ก็ต่อเมื่อมีค่าสหสัมพันธ์กับมาตร ๖ ชั้นสุดท้ายในอัตราสูง แต่ในทางปฏิบัติก็ได้กล่าวแล้วว่า เป็นการยากที่จะค้นหามาตร ๖ ชั้นสุดท้ายจะนั้นในการพิจารณาความเที่ยงตรงของมาตร ๖ ก็อาจจะทำได้โดยวิเคราะห์ว่า รายละเอียดที่จะนำมาสร้างมาตร ๖ เป็นประเด็นที่สำคัญต่อการปฏิบัติงานในตำแหน่งหน้าที่หนึ่ง ๆ หรือไม่ การวิเคราะห์นี้เป็นแบบความเที่ยงตรงในแง่แนวคิด (construct validity) ซึ่งหมายความว่า ผู้สร้างมาตร ๖ หลีกเลี่ยงการใช้ความรู้สึกส่วนตัวได้ลำบาก²⁷ อย่างไรก็ตาม หากผู้สร้างใช้คุณสมบัติวิเคราะห์เหตุผลเป็นอย่างดีก็อาจหลีกเลี่ยงความมีอคติได้ โดยที่ข้อมูลสำหรับบางอย่างเห็นได้ชัดว่า หากผู้สร้างมาตร ๖ ใช้ความระมัดระวังก็อาจคัดเลือกเฉพาะที่ตรงประเด็นได้ เช่น หากมี

²⁶ Thorndike, *op. cit.*, pp. 124 - 125.

²⁷ Lee J. Cronbach, and Paul E. Meehl, "Construct Validity in Psychological Tests," in Jackson and Messick, eds., *op. cit.*, pp. 60 - 62.

การวิเคราะห์ความสำเร็จของงานในตำแหน่งเจ้าหน้าที่ประจำบ้าน ความแม่นยำ และความรวดเร็วในการยิงไปถูกเป้าหมายก็คือ ข้อมูลที่มีความตรงประเด็นที่จะนำมาใช้สร้างมาตราฯ ที่เที่ยงตรงได้ และในกรณีที่ไม่สามารถที่ลงไปได้ชัดว่าข้อมูลนั้นตรงประเด็นกับความสำเร็จของงานหรือไม่ ก็อาจใช้หลักสถิติช่วยในการวิเคราะห์ได้²⁸

ความเชื่อมั่น สัมประสิทธิ์ความเชื่อมั่นของมาตราฯ ควรมีค่าสูงกว่าศูนย์ ทั้งนี้หมายความว่าผู้สร้างมาตราฯ ควรได้มีการทดลองคำนวณค่าทางสถิติของข้อมูลต่างๆ ที่นำมาประกอบการสร้างมาตราฯ²⁹ Thorndike ยืนยันว่าไม่มีความจำเป็นถึงขั้นสูงสุดที่จะต้องให้มาตราฯ มีค่าสัมประสิทธิ์ความเชื่อมั่นสูง ต่อเมื่อกำหนดค่ามากก็ควรจะแก้ไข การแก้ไขพิจารณาเป็น 2 นัย คือคุณค่าความเชื่อมั่นในแง่เนื้อหาแท้ (intrinsic reliability) และคุณค่าความเชื่อมั่นในแง่อาการภายนอก (extrinsic reliability) หากค่าสัมประสิทธิ์ความเชื่อมั่นต่ำตามนัยแรกก็เป็นเพราะเกิดจากความไม่คงเส้นคงวา ส่วนบุคคลที่ปฏิบัติงานนั้นๆ แก้ไขได้โดยเพิ่มจำนวนตัวอย่างที่ประเมินให้มากขึ้น แต่ถ้าหากต่ำตามนัยที่สองก็เป็นเพราะเกิดจากสภาพแวดล้อมที่เกี่ยวข้องกับการปฏิบัติงาน เช่น ดินฟ้าอากาศ เครื่องจักร ลักษณะการปฏิบัติงานของผู้ร่วมงานของผู้ถูกประเมิน หรือการไม่จำกัดความหมายของงานที่จะประเมินให้แน่นอนตายตัว ในกรณีเช่นนี้ต้องมีการควบคุมสภาพแวดล้อมเหล่านี้โดยอาศัยหลักวิธีการวิจัยอย่างแท้จริง รายละเอียดมีกล่าวไว้ในตำราของผู้วิจัยคนเดียวกันนี้แล้ว³⁰

ความสามารถในการแยกแยะ (discrimination) หมายความว่า มาตราฯ ที่สร้างนั้นจะต้องมีคุณภาพแยกได้ว่าพฤติกรรมในลักษณะใดจึงทำให้ปฏิบัติงานได้ผลสำเร็จอย่างดีที่สุด และเป็นอย่างไรๆ ลงมาประการใด หากมาตราฯ ที่สร้างขึ้นนั้นไม่สามารถแยกได้ว่าบุคคลที่ปฏิบัติงานอย่างไรจะมีผลสำเร็จของงานต่างกันเป็นระดับอย่างไร ก็ย่อมทำให้ไม่สามารถสร้างเครื่องมือวัดที่จะคาดคะเนได้ว่าควรคัดเลือกเฉพาะบุคคลประเภทใดจึงจะได้บุคคลที่มีคุณสมบัติตามที่ต้องการ

²⁸ Edwin E. Ghiselli, *Theory of Psychological Measurement* (New York : McGraw - Hill, Inc., 1964) pp. 354 - 357.

²⁹ Thorndike, *op.cit.*, pp. 125 - 127.

³⁰ สวัสดิ์ สุคนธรังษี และ สมศรี จุลละนันท์, *ระเบียบเบื้องต้นในการวิจัย (โรเนียว)*, สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2511, หน้า 112-115.

ความสะดวกและความประหยัด (practicality) การรวบรวมข้อมูลเพื่อการสร้างมาตราก็มีลักษณะเช่นเดียวกับการเก็บข้อมูลเพื่อการวิจัยทั่วไป คือเนื่องจากงานนี้เป็งานที่ต้องเสียค่าใช้จ่ายและเวลามาก ผู้สร้างมาตรฯ จึงต้องใช้ดุลยพินิจให้รอบคอบในการกำหนดข้อจำกัดเพื่อการวิจัย (limitation) เพื่อให้สอดคล้องกับทุนและเวลาที่มีอยู่⁸¹ ยิ่งไปกว่านั้นข้อมูลในการสร้างมาตรฯ ยังมีลักษณะพิเศษกว่าข้อมูลธรรมดาทั่วไปในการวิจัยในแง่ที่ว่า ข้อมูลส่วนใหญ่ที่จะใช้ในการสร้างมาตรฯ เป็นข้อมูลที่ต้องรบกวนเวลาและความสะดวกของผู้ปฏิบัติงานฝ่ายต่าง ๆ เพราะเป็นข้อมูลที่เป็นข้อเท็จจริงในการปฏิบัติงานจริง ๆ ฉะนั้นนักวิจัยงานบุคคลในท่านี้อย่อมหลีกเลี่ยงไม่ได้ที่จะสร้างความไม่พอใจให้แก่บุคคลฝ่ายต่าง ๆ ถ้าไม่ได้ทำให้บุคคลที่จะต้องสละเวลาและความสะดวกในการให้ข้อเท็จจริงเข้าใจเสียก่อนล่วงหน้าถึงประโยชน์สำคัญของมาตรฯ และถึงแม้ผู้ปฏิบัติงานฝ่ายต่าง ๆ จะเห็นความสำคัญในเรื่องนี้ แต่การที่ผู้วิจัยต้องใช้ระยะเวลาานกว่าจะสร้างมาตรฯสำเร็จก็อาจทำให้ผู้ให้ความร่วมมือขาดความอดทนเสียก่อนก็ได้ นักวิจัยที่มีหน้าที่สร้างมาตรฯ จึงควรเทียบปัญหาเรื่องนี้กับระยะเวลาที่ใช้ปฏิบัติงาน และเมื่อเห็นว่าถึงขั้นที่ควรจะให้ผู้ปฏิบัติงานที่เกี่ยวข้องได้ทราบหรือตระหนักประโยชน์ของการรวบรวมข้อมูลในตอนใดก็ควรดำเนินการให้สอดคล้องกับการขอความร่วมมือของทุก ๆ ฝ่าย

ความสังเกต

ความพยายามของผู้บังคับบัญชาที่จะจับบุคคลให้ปฏิบัติงานในตำแหน่งหน้าที่ที่เหมาะสมที่สุดนั้นแม้จะกระทำไปโดยบริสุทธิ์ใจโดยไม่มีจิตใจลำเอียงแม้แต่น้อยก็อาจมีคุณค่าไม่แตกต่างกันไปกว่า lip service เท่าใดนัก หากขาดความเข้าใจในหลักเกณฑ์ที่ถูกต้องในการจัดงานบุคคล ปัญหาเรื่องการวัดความสามารถที่ยกมาวิเคราะห์ในที่นี้เป็นเพียงความสามารถด้านหนึ่งของบุคคลเท่านั้น นอกจากนี้แล้ว ในการวัดบุคคลนั้นยังต้องคำนึงคุณสมบัติด้านอื่น ๆ นอกเหนือจากความสามารถอีกหลายประการ เช่น บุคลิกภาพ ความตั้งใจและเอาใจใส่ ทักษะ และทัศนคติที่มีต่องาน เป็นต้น การวัดคุณสมบัติเหล่านี้โดยเที่ยงธรรมก็มีหลักเกณฑ์ในตนเองเดียวกับกับการวัดความสามารถทางสมองที่กล่าวมาแล้วทั้งหมด จะเห็นได้ว่าการปฏิบัติโดยคำนึงถึงหลักที่กล่าวมานี้อย่างทำได้ลำบากมาก เพราะนอกจากจะต้องเสียแรงงานและค่าใช้จ่ายมากแล้วยังต้องการความอดทนอย่างสูงอีกด้วย แต่ความลำบากในเรื่องนี้ก็ไมควรที่จะทำให้เป็นข้อก้างในการที่จะไม่ถือปฏิบัติและเสี่ยงไปพิจารณาแต่ “ความเหมาะสม” ดังที่นิยมใช้กันทั่วไปในวงการของทางราชการปัจจุบัน เป็นที่น่ายินดีว่า ขณะนี้สำนักงานคณะกรรมการข้าราชการพลเรือนกำลังจัดให้มี

⁸¹ Thorndike, *op. cit.*, pp. 131 - 132.

การใช้ระบบการจำแนกตำแหน่ง (position classification) ในวงราชการชั้น การริเริ่มของสำนักงาน ก.พ. ในเรื่องนี้จะช่วยบีบบังคับให้ผู้บัญชาได้ตระหนักถึงลักษณะของงานที่อยู่ใต้บังคับบัญชาของตนจะต้องปฏิบัติโดยแท้จริง การตระหนักถึงรายละเอียดความเป็นจริงในเรื่องนี้เป็นก้าวอีกก้าวหนึ่งที่จะไปสู่การวัดบุคคลตามหลักวิชาเพื่อให้บุคคลได้ปฏิบัติงานในตำแหน่งที่เหมาะสมที่สุดทั้งคำขวัญที่ตีพิมพ์ปากของนักบริหารทั้งหลาย ต่อเมื่อนักบริหารสามารถสร้างแนวคิดของตนเองให้สอดคล้องกับหลักการและทฤษฎีที่ได้มีการพิสูจน์ไว้ดังกล่าวมาแล้วในเรื่องนี้ ไม่เพียงแต่ผู้ที่อยู่ใต้บังคับบัญชาจะได้รับประโยชน์ในแง่ความยุติธรรมอันพึงได้รับแล้ว งานราชการก็จะสำเร็จลุล่วงได้ผลดีโดยที่บุคคลทุกฝ่ายได้เกิดมีกำลังใจที่จะปฏิบัติงานร่วมกัน เป็นทางให้ข้อครหาต่าง ๆ ในเรื่องการปฏิบัติงานของข้าราชการไทยในปัจจุบันได้ค่อย ๆ หดสั้นไป

ABSTRACT

Intellectual Abilities and Criterion Measure

It is tragic to see executives and personnel managers pay lip service to merit rating while they hardly heed the importance of personnel measurement as a scientific discipline. This article introduces modern concepts and methodology of assessing intellectual abilities. Intelligence is viewed as several mental abilities which one cannot hastily judge by a crude interview or simple examination. Though this point is familiar among students of personnel measurement Thai executives and personnel managers, especially in the public service, seem not to mind it in the least as much as ever in their practice.

The emphasis was given on the concept of test validation. Criterion measure as a step toward developing a test which is capable of measuring intellectual abilities desired for various jobs was extensively discussed.
