

บทที่ 5

สรุปผลการศึกษาและข้อเสนอแนะ

การวิจัยเรื่อง การสร้างและพัฒนาแบบทดสอบความสามารถทั่วไปทางสมองที่วัดทั้งความสามารถและความรวดเร็วในการตอบ สำหรับข้าราชการพลเรือนไทยระดับแรกบรรจุ มีวัตถุประสงค์ 3 ประการ ประการแรก คือ เพื่อสร้างแบบทดสอบความสามารถทั่วไปทางสมองที่สามารถเปรียบเทียบความสามารถของผู้สอบได้ ในเวลาที่รวดเร็ว เพื่อการคัดเลือกบุคคลเข้ารับราชการในระดับแรกบรรจุ ประการที่สอง คือ เพื่อตรวจสอบคุณภาพของแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นในด้านความเชื่อมั่น ความเที่ยงตรงตามเนื้อหา ความเที่ยงตรงตามโครงสร้าง และความเที่ยงตรงตามสภาพในการทำงานผลการปฏิบัติงาน และประการสุดท้าย คือ เพื่อสร้างเกณฑ์ปกติของแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นทั้งในภาพรวมและจำแนกตามกลุ่มสายงาน

กลุ่มประชากรในการวิจัย ได้แก่ กลุ่มข้าราชการพลเรือนไทยในสายงานระดับแรกบรรจุ ระดับ 3 หรือ 4 ซึ่งยังดำรงตำแหน่งอยู่ในระดับ 3-5 จากส่วนราชการระดับกรม ในเขตกรุงเทพมหานครและปริมณฑล จำนวน 1,300 คน โดยผู้วิจัยกำหนดกลุ่มตัวอย่างตามขั้นตอนในการสร้างและพัฒนาแบบทดสอบ รวมจำนวนทั้งสิ้น 680 คน ส่วนเครื่องมือที่ใช้ในการวิจัย ประกอบด้วย แบบทดสอบความสามารถทั่วไปทางสมองที่ผู้วิจัยสร้างขึ้นเพื่อวิเคราะห์คุณภาพข้อสอบรายข้อ ซึ่งมุ่งวัดองค์ประกอบหลักของความสามารถทั่วไปทางสมองด้านภาษา-การศึกษา ตามทฤษฎีกลุ่มองค์ประกอบแบบลำดับขั้นของเวอร์นอน และแบบประเมินผลการปฏิบัติงานที่ศูนย์สรรหาและเลือกสรร สำนักงาน ก.พ. เป็นผู้สร้างขึ้นเพื่อใช้ในการศึกษาวิจัยเกี่ยวกับเครื่องมือประเมินบุคคล โดยใช้ชื่อว่าแบบสำรวจเพื่อการพัฒนากระบวนการสรรหาและเลือกสรรบุคคลเข้ารับราชการ

ในส่วนของขั้นตอนการวิจัย ผู้วิจัยกำหนดผังโครงสร้างข้อสอบ และสร้างแบบทดสอบความสามารถทั่วไปทางสมองชุดแรก ประกอบด้วยข้อสอบแบบเลือกตอบ 4 ตัวเลือก จำนวน 120 ข้อ และขอความอนุเคราะห์ให้ผู้เชี่ยวชาญตรวจสอบเพื่อวิเคราะห์ความเที่ยงตรงตามเนื้อหา แล้วนำไปทดสอบกับกลุ่มตัวอย่างกลุ่มแรก ซึ่งกำหนดไว้จำนวน 200 คน เพื่อวิเคราะห์คุณภาพรายข้อ คัดเลือกข้อสอบและจัดชุดแบบทดสอบ จากนั้นจึงวิเคราะห์คุณภาพด้านความเชื่อมั่นแบบแบ่งครึ่ง ข้อคู่และข้อคู่ และคุณภาพด้านความเที่ยงตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน แล้วนำไปทดสอบกับกลุ่มตัวอย่างกลุ่มที่สอง ซึ่งกำหนดตัวอย่างไว้จำนวน 80 คน เพื่อ

กำหนดเวลาตอบ จากนั้นจึงนำแบบทดสอบไปทดสอบกับกลุ่มตัวอย่างกลุ่มที่สาม ซึ่งกำหนดตัวอย่างไว้จำนวน 400 คน เพื่อวิเคราะห์ความเที่ยงตรงตามสภาพในการทำนายผลการปฏิบัติงานร่วมกับคะแนนที่ได้จากแบบประเมินผลการปฏิบัติงาน และสร้างเกณฑ์ปกติของแบบทดสอบ สำหรับการคัดเลือกบุคคลเข้ารับราชการ

จากการดำเนินการวิจัยตามที่กำหนด สามารถสรุปผลการวิจัยได้ดังนี้

สรุปผลการวิจัย

1. ลักษณะทั่วไปของกลุ่มตัวอย่าง

1.1 กลุ่มตัวอย่างในการวิเคราะห์คุณภาพข้อสอบรายข้อ การจัดชุดแบบทดสอบ การวิเคราะห์ความเชื่อมั่นของแบบทดสอบ และการวิเคราะห์ความเที่ยงตรงตามโครงสร้างของแบบทดสอบ มีจำนวน 201 คน ส่วนใหญ่เป็นเพศหญิง โดยมากเป็น กลุ่มที่มีอายุ 31-40 ปี และมีกลุ่มสายงานครบทั้ง 8 กลุ่ม โดยกลุ่มสายงานที่มีจำนวนคนใกล้เคียงกันและมากกว่ากลุ่มสายงานอื่นๆ 3 กลุ่ม คือ กลุ่มสายงานการศึกษา ศิลปะ สังคม และการพัฒนา กลุ่มสายงานบริหารและธุรการ งานสถิติ งานนิติการ งานการทูตและต่างประเทศ และกลุ่มสายงานการคลัง การเศรษฐกิจ การพาณิชย์ และอุตสาหกรรม

1.2 กลุ่มตัวอย่างในการกำหนดเวลาตอบ มีจำนวน 80 คน ส่วนใหญ่เป็นเพศหญิง โดยมากเป็นกลุ่มที่มีอายุ 31- 40 ปี และ มีกลุ่มสายงานครบทั้ง 8 กลุ่ม โดยกลุ่มสายงานที่มีจำนวนคนใกล้เคียงกันและมากกว่ากลุ่มสายงานอื่นๆ 2 กลุ่ม คือ กลุ่มสายงานการคลัง การเศรษฐกิจ การพาณิชย์และอุตสาหกรรม และกลุ่มสายงานเกษตรกรรม

1.3 กลุ่มตัวอย่างในการวิเคราะห์ความเที่ยงตรงตามสภาพ มีจำนวน 225 คน จากที่กำหนดไว้ 400 คน โดยเป็นกลุ่มตัวอย่างที่มีคะแนนผลการปฏิบัติงานจากผู้บังคับบัญชาระดับต้นเท่านั้น เพศชายและเพศหญิงมีจำนวนในสัดส่วนที่ใกล้เคียงกัน ส่วนใหญ่มีอายุ 31-40 ปี และมีกลุ่มสายงานครบทั้ง 4 กลุ่มสายงานตามที่กำหนดในวิธีการวิจัย โดยกลุ่มสายงานที่มีจำนวนคนมากที่สุด คือ กลุ่มสายงานการคลัง การเศรษฐกิจ การพาณิชย์และอุตสาหกรรม รองลงมาเป็นกลุ่มสายงานบริหารและธุรการ งานสถิติ งานนิติการ งานการทูตและต่างประเทศ กลุ่มสายงานวิศวกรรม สถาปัตยกรรม และช่างเทคนิคต่างๆ และกลุ่มสายงานการศึกษา ศิลปะ สังคม และการพัฒนาชุมชน นอกจากนี้ ยังประกอบด้วยกลุ่มสายงานอื่นๆ จำนวนเล็กน้อย

1.4 กลุ่มตัวอย่างในการสร้างเกณฑ์ปกติ มีจำนวน 363 คน โดยเป็นกลุ่มตัวอย่างเดียวกันกับกลุ่มตัวอย่างในการวิเคราะห์ความเที่ยงตรงตามสภาพ แต่รวมผู้ที่ไม่มีความเห็นผลการปฏิบัติงานจากผู้บังคับบัญชาระดับต้น และกลุ่มตัวอย่างในการกำหนดเวลาตอบด้วย ส่วนใหญ่เป็นเพศหญิง โดยมากเป็นกลุ่มที่มีอายุ 31-40 ปี และมีกลุ่มสายงานครบทั้ง 4 กลุ่มสายงานตามที่กำหนดตามที่กำหนดในวิธีการวิจัย โดยกลุ่มสายงานที่มีจำนวนคนมากที่สุด คือ กลุ่มสายงานการคลัง การเศรษฐกิจ การพาณิชย์และอุตสาหกรรม รองลงมาเป็นกลุ่มสายงานบริหารและธุรการ งานสถิติ งานนิติกร งานการทูตและต่างประเทศ กลุ่มสายงานวิศวกรรม สถาปัตยกรรม และช่างเทคนิคต่างๆ และกลุ่มสายงานการศึกษา ศิลปะ สังคม และการพัฒนาชุมชน นอกจากนี้ ยังประกอบด้วยกลุ่มสายงานอื่นๆ จำนวนเล็กน้อย

2. ความเที่ยงตรงตามเนื้อหาของแบบทดสอบ

แบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นชุดแรก จำนวน 120 ข้อ เพื่อการวิเคราะห์รายข้อและคัดเลือกข้อสอบ ผ่านการพิจารณาตัดสินความสอดคล้องตามเนื้อหาจากผู้เชี่ยวชาญด้านแบบทดสอบความสามารถทางสมองจำนวน 6 ท่าน จำนวน 113 ข้อ โดยมีค่าอัตราส่วนความเที่ยงตรงตามเนื้อหาตามวิธีการของลอร์วี เท่ากับ 1.00 และเมื่อทำการคัดเลือกข้อสอบที่มีค่าอัตราส่วนความเที่ยงตรงตามเนื้อหาผ่านเกณฑ์เพื่อจัดชุดแบบทดสอบจำนวน 48 ข้อ แล้ว แบบทดสอบที่จัดชุด มีค่าอัตราส่วนความเที่ยงตรงตามเนื้อหาตามรายข้อ เท่ากับ 1.00

3. คุณภาพข้อสอบรายข้อ

3.1 ข้อสอบที่สร้างขึ้นจำนวน 120 ข้อในแบบทดสอบชุดแรก มีข้อที่มีค่าความยากง่ายและค่าอำนาจจำแนกผ่านเกณฑ์จำนวน 66 ข้อ โดยในองค์ประกอบของความสามารถด้านภาษา มีข้อสอบที่มีค่าความยากง่ายและค่าอำนาจจำแนกผ่านเกณฑ์จำนวน 27 ข้อ ส่วนในองค์ประกอบของความสามารถด้านตัวเลข มีข้อสอบที่มีค่าความยากง่ายและค่าอำนาจจำแนกผ่านเกณฑ์ จำนวน 39 ข้อ ทั้งนี้ ข้อสอบที่มีค่าความยากง่ายและค่าอำนาจจำแนกมากที่สุด ได้แก่ข้อสอบด้านความสัมพันธ์เชิงปริมาณ จำนวน 17 ข้อ รองลงมาเป็นข้อสอบด้านการแก้ไขโจทย์เลข จำนวน 16 ข้อ การสรุปความ 12 ข้อ อุปมาอุปไมย จำนวน 12 ข้อ และอนุกรมตัวเลข

จำนวน 6 ข้อ ส่วนด้านการเข้าใจภาษา มีค่าความยากง่ายและค่าอำนาจจำแนกผ่านเกณฑ์น้อยที่สุด จำนวน 3 ข้อ

3.2 ข้อสอบที่นำมาจัดชุดแบบทดสอบฉบับจริงจำนวน 48 ข้อ มีค่าความยากง่ายอยู่ในช่วง .29 - .88 และมีค่าอำนาจจำแนกอยู่ในช่วง .16 - .64 โดยเป็นข้อสอบที่มีระดับค่อนข้างง่าย จำนวนมากที่สุด 28 ข้อ รองลงมาเป็นข้อสอบที่มีระดับง่ายมาก 13 ข้อ ระดับค่อนข้างยาก 5 ข้อ และระดับยากปานกลาง 2 ข้อ และเป็นข้อสอบที่มีค่าอำนาจจำแนกในระดับพอใช้ได้ จำนวนมากที่สุด 25 ข้อ ระดับดี 17 ข้อ ระดับดีมาก 2 ข้อ และระดับค่อนข้างต่ำ 4 ข้อ ทั้งนี้ มีข้อที่มีค่าความยากง่ายและค่าอำนาจจำแนกผ่านเกณฑ์อย่างแท้จริงทั้งหมด จำนวน 35 ข้อ

4. การจัดชุดแบบทดสอบ

ข้อสอบด้านการแก้โจทย์เลขมีจำนวนข้อกระจายตามช่วงค่าความยากง่ายที่กำหนดไว้ตามผังโครงสร้างข้อสอบมากที่สุด รองลงมาเป็นข้อสอบด้านการแก้โจทย์เลขและข้อสอบด้านการสรุปความ ส่วนข้อสอบด้านการเข้าใจภาษา ข้อสอบด้านอุปมาอุปไมย และข้อสอบด้านอนุกรมตัวเลข มีจำนวนข้อกระจายตามช่วงค่าความยากง่ายที่กำหนดไว้ค่อนข้างน้อย ไม่เพียงพอสำหรับการจัดเรียงชุดข้อสอบแบบเกลียวตามที่กำหนดไว้ในผังโครงสร้างข้อสอบ จึงจำเป็นต้องนำข้อสอบบางข้อที่มีคุณภาพไม่ผ่านเกณฑ์มาจัดชุด และตัดข้อสอบบางข้อที่แม้จะมีคุณภาพดีก็ตามออก เพื่อให้คงคุณลักษณะของการจัดชุดข้อสอบแบบเกลียวมากที่สุด ทั้งนี้ เนื่องจากข้อสอบที่ผ่านเกณฑ์ส่วนใหญ่มีลักษณะค่อนข้างง่าย แบบทดสอบจึงมีลักษณะค่อนข้างง่ายกว่าที่กำหนดไว้ในผังโครงสร้างข้อสอบ การจัดเรียงชุดข้อสอบแสดงในตารางที่ 4.10 และภาพที่ 4.1

5. ความเชื่อมั่นของแบบทดสอบ

ผลการวิเคราะห์ค่าความเชื่อมั่นแบบแบ่งครึ่งข้อคือและข้อคู่ของแบบทดสอบความสามารถทั่วไปทางสมองฉบับจริง ด้วยวิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์ตามสูตรของสเปียร์แมนบราวน์ พบว่าแบบทดสอบมีค่าความเชื่อมั่นแบบแบ่งครึ่งข้อคือและข้อคู่ประมาณ .86 โดยองค์ประกอบของความสามารถด้านภาษามีค่าความเชื่อมั่นประมาณ .74 และองค์ประกอบของความสามารถด้านตัวเลขมีค่าความเชื่อมั่นประมาณ .84 ทั้งนี้ ข้อสอบด้านการเข้าใจภาษา ด้านอุปมาอุปไมย ด้านการสรุปความ ด้านอนุกรมตัวเลข ด้านความสัมพันธ์เชิงปริมาณ และด้าน

การแก้โจทย์เลข มีค่าความเชื่อมั่นในระดับ .40, .68, .40, .66, .74 และ .68 ตามลำดับ ทั้งนี้ การวิเคราะห์ค่าความเชื่อมั่นแบบแบ่งครึ่งข้อดีและข้อคู่ของแบบทดสอบความสามารถทั่วไปทาง สมรรถนะจริง ในช่วงกำหนดเวลาตอบที่เหมาะสมซึ่งได้จากขั้นตอนการกำหนดเวลาตอบ พบว่า แบบทดสอบในช่วงเวลาตอบที่เหมาะสม (ช่วงเวลาตอบที่ 30 นาที) มีค่าความเชื่อมั่นแบบแบ่งครึ่ง ข้อดีและข้อคู่ประมาณ .90

6. ความเที่ยงตรงตามโครงสร้างของแบบทดสอบ

ผลการวิเคราะห์ความเที่ยงตรงตามโครงสร้างของ โดยใช้วิธีการวิเคราะห์ องค์ประกอบแบบยืนยัน (Confirmatory Factor Analysis: CFA) พบว่า ลักษณะของข้อมูลไม่ สอดคล้องกับโมเดลความสามารถทั่วไปทางสมรรถนะด้านภาษา-การศึกษาตามทฤษฎีกลุ่มองค์ประกอบ ลำดับขั้นของเวอร์นอน โดยค่า goodness of fit index มีค่าต่ำมาก และค่าความคลาดเคลื่อนสูงมาก ผู้วิจัยจึงใช้การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis: EFA) อีกครั้ง

ผลจากการวิเคราะห์องค์ประกอบเชิงสำรวจในขั้นที่หนึ่ง พบว่า ข้อสอบด้านการแก้ โจทย์เลข ด้านความสัมพันธ์เชิงปริมาณ ด้านอนุกรมตัวเลข และด้านการสรุปความ มีค่าน้ำหนัก ขึ้นกับองค์ประกอบด้านตัวเลข โดยมีค่าน้ำหนักประมาณ .84, .79, .66 และ .55 ตามลำดับ ส่วน ข้อสอบด้านการเข้าใจภาษา และ อุปมาอุปไมย มีค่าน้ำหนักขึ้นกับองค์ประกอบด้านภาษา โดยมี ค่าน้ำหนักประมาณ .87 และ .74 ตามลำดับ

ส่วนผลจากการวิเคราะห์องค์ประกอบเชิงสำรวจในขั้นที่สอง พบว่า องค์ประกอบด้าน ภาษาและด้านตัวเลข มีค่าน้ำหนักกับองค์ประกอบความสามารถด้านการภาษา-การศึกษา โดย องค์ประกอบด้านภาษา และองค์ประกอบด้านตัวเลขมีค่าน้ำหนักกับองค์ประกอบความสามารถ ด้านภาษา-การศึกษา ประมาณ .78 และ .62 ตามลำดับ

7. การกำหนดเวลาตอบของแบบทดสอบ

จากการพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สันระหว่างคะแนนรวมของ ช่วงที่จับเวลาตอบที่ 15, 20, 25, 30 และ 35 นาที กับช่วงที่ไม่จับเวลา พบว่า คะแนนรวมของช่วง ที่จับเวลาตอบทุกช่วงเวลามีความสัมพันธ์กับคะแนนรวมในช่วงที่ไม่จับเวลา แสดงว่า การกำหนด ความเร็วในการทดสอบทั้ง 5 ช่วงเวลาไม่ทำให้ผลของการวัดความสามารถของผู้สอบแตกต่างกันไป

จากช่วงที่ไม่จับเวลา ดังนั้น การกำหนดเวลาตอบจึงพิจารณาจากค่าเบี่ยงเบนมาตรฐานเป็นสำคัญ ซึ่งพบว่าช่วงเวลาตอบที่ 35 นาที มีค่าเบี่ยงเบนมาตรฐานของคะแนนสูงสุด รองลงมาเป็นเวลาตอบที่ 30 นาที แต่เนื่องจากจุดมุ่งหมายของการสร้างแบบทดสอบฉบับนี้ คือ เพื่อให้วัดความสามารถของผู้สอบได้ในเวลาที่รวดเร็วด้วย ประกอบกับช่วงเวลาตอบที่ 30 นาที มีค่าเบี่ยงเบนมาตรฐานใกล้เคียงมากกับช่วงเวลาตอบที่ 35 นาที ผู้วิจัยจึงกำหนดเวลาตอบที่ 30 นาที เป็นช่วงเวลาตอบที่เหมาะสมของแบบทดสอบ

8. ความเที่ยงตรงตามสภาพของแบบทดสอบ

ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ระหว่างคะแนนรวมจากแบบทดสอบความสามารถทั่วไปทางสมอง และผลการปฏิบัติงานของทุกกลุ่มสายงาน ในภาพรวม และจำแนกตามกลุ่มสายงาน ไม่พบว่า คะแนนรวมของแบบทดสอบความสามารถทั่วไปทางสมองมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับผลการปฏิบัติงาน ทั้งในด้านผลงาน ด้านทักษะ ความสัมพันธ์ระหว่างบุคคล และด้านการคิด

9. เกณฑ์ปกติของแบบทดสอบ

การเปรียบเทียบคะแนนเฉลี่ยจากแบบทดสอบความสามารถทั่วไปทางสมอง ระหว่างเพศ และระหว่างกลุ่มอายุโดยใช้สถิติทดสอบที และการเปรียบเทียบคะแนนเฉลี่ยจากแบบทดสอบความสามารถทั่วไปทางสมองระหว่างกลุ่มสายงานโดยใช้การวิเคราะห์ความแปรปรวนทางเดียว เพื่อใช้ในการพิจารณาจำแนกเกณฑ์ปกติ พบว่า คะแนนเฉลี่ยจากแบบทดสอบความสามารถทั่วไปทางสมองมีความแตกต่างกันตามเพศ กลุ่มอายุ และกลุ่มสายงาน โดยในส่วนของกลุ่มสายงานพบว่ากลุ่มสายงานวิศวกรรม มีคะแนนเฉลี่ยจากแบบทดสอบความสามารถทั่วไปทางสมองสูงกว่ากลุ่มงานอื่นๆ ดังนั้น ผู้วิจัยจึงสร้างเกณฑ์ปกติแบบเปอร์เซ็นต์ไทล์ ทั้งโดยภาพรวม และจำแนกตามเพศ กลุ่มอายุและกลุ่มสายงาน โดยในส่วนของกลุ่มสายงาน มีการจำแนกเกณฑ์ปกติ โดยรวมกลุ่มสายงาน 3 กลุ่มเข้าด้วยกัน คือ กลุ่มสายงานการคลัง การเศรษฐกิจ การพาณิชย์ และอุตสาหกรรม กลุ่มสายงานบริหารและธุรการ งานสถิติ งานนิติการ งานการทูตและต่างประเทศ และกลุ่มสายงานการศึกษา ศิลปะ สังคม และการพัฒนาชุมชน ส่วนกลุ่มสายงานวิศวกรรม สถาปัตยกรรม และช่างเทคนิคต่างๆ ควรมีการสร้างเกณฑ์ปกติจำแนกสำหรับกลุ่มสายงานโดยเฉพาะ แต่เนื่องจากกลุ่มสายงานวิศวกรรม สถาปัตยกรรม และช่างเทคนิคต่างๆ มีจำนวนตัวอย่างไม่เพียงพอ จึงยังไม่มี

การสร้างเกณฑ์ปกติสำหรับกลุ่มสายงานในการวิจัยครั้งนี้

อภิปรายผล

จากวัตถุประสงค์การวิจัย การดำเนินการตามวิธีการวิจัย และสรุปผลการวิจัยข้างต้น ผู้วิจัยขอเสนอประเด็นในการอภิปรายผลดังนี้

1. ข้อมูลจากกลุ่มตัวอย่าง

กลุ่มตัวอย่างที่เข้ารับการทดสอบในการวิจัยครั้งนี้มีจำนวนมาก ข้อมูลที่ได้มีความครอบคลุมทั้งเพศ กลุ่มอายุ สายงาน และกลุ่มสายงานที่หลากหลาย สภาพการดำเนินการทดสอบจัดในห้องประชุมขนาดใหญ่ มีการชี้แจงวิธีการทดสอบ และวัตถุประสงค์ของการทดสอบอย่างชัดเจน ทำให้กลุ่มตัวอย่างใช้ความสามารถอย่างเต็มที่ในการทดสอบ ส่งผลให้ข้อมูลจากการทดสอบสามารถเชื่อถือได้ในระดับหนึ่ง

2. ความเที่ยงตรงตามเนื้อหาของแบบทดสอบ

แบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นครั้งแรก จำนวน 120 ข้อ ผ่านการพิจารณาตัดสินความสอดคล้องโดยผู้เชี่ยวชาญด้านแบบทดสอบความสามารถทางสมอง โดยมีค่าอัตราส่วนความเที่ยงตรงตามเนื้อหาตามวิธีการของลอร์วีร์ผ่านเกณฑ์ จำนวน 113 ข้อ โดยทุกข้อที่นำมาจัดชุด มีค่าอัตราส่วนความเที่ยงตรงตามเนื้อหาเท่ากับ 1 ทำให้เชื่อได้ว่าแบบทดสอบที่สร้างขึ้นสามารถวัดความสามารถทั่วไปทางสมอง ทั้งนี้ การที่ข้อสอบมีความเที่ยงตรงตามเนื้อหาอย่างชัดเจน อาจเป็นเพราะรูปแบบของข้อสอบเป็นที่คุ้นเคยและนิยมวัดโดยทั่วไปในแบบทดสอบที่มีการวัดความสามารถทางสมองด้านภาษาและด้านตัวเลข เช่น แบบทดสอบ Wonderlic Personnel Test แบบทดสอบ Graduate Record Examinations และแบบทดสอบ Scholastic Aptitude Test เป็นต้น

3. คุณภาพของข้อสอบรายข้อ

3.1.1 ข้อสอบที่วัดความสามารถด้านตัวเลข มีค่าความยากง่ายและค่าอำนาจ

จำแนกดีกว่าข้อสอบที่วัดความสามารถด้านภาษา ยกเว้นข้อสอบด้านการสรุปความ ทั้งนี้ อาจเป็นเพราะข้อสอบที่วัดความสามารถด้านตัวเลขมีความเป็นปรนัย คือสามารถตัดสินข้อถูกและข้อผิดได้ชัดเจนมากกว่าข้อสอบที่วัดความสามารถด้านภาษา ในขณะที่ข้อสอบด้านการสรุปความ ก็มีลักษณะของความถูกต้องชัดเจนตามกฎทางตรรกศาสตร์เช่นกัน

3.1.2 ข้อสอบด้านการเข้าใจภาษาส่วนใหญ่เป็นข้อสอบที่ง่าย เนื่องจากการสร้างข้อสอบด้านการเข้าใจภาษาให้ยาก มักจะต้องเลือกใช้ข้อความที่มีลักษณะคลุมเครือ ซึ่งมีความเสี่ยงต่อการโต้แย้งเกี่ยวกับข้อเฉลย ว่าขึ้นอยู่กับความตีความของแต่ละบุคคลหรือไม่ ในการสร้างข้อสอบด้านการเข้าใจภาษา จึงจำเป็นต้องหลีกเลี่ยงข้อความที่มีลักษณะคลุมเครือ และสร้างข้อสอบให้มีลักษณะถูกต้องอย่างชัดเจนมากที่สุด ซึ่งอาจเป็นสาเหตุหนึ่งที่ทำให้ผู้ตอบสามารถพิจารณาคำตอบได้โดยง่าย ซึ่งปัญหานี้ อาจมีทางแก้ไขได้โดยการมีผู้เชี่ยวชาญจำนวนหนึ่งที่มีประสบการณ์ในการสร้างข้อสอบด้านนี้กับกลุ่มผู้สมัครเข้ารับราชการพลเรือน เป็นผู้ตรวจสอบความถูกต้องของข้อสอบในขั้นตอนของการสร้างข้อสอบ

3.1.3 จากระดับค่าความยากง่ายและระดับค่าอำนาจแจกของข้อสอบ ในแบบทดสอบชุดจริง แบบทดสอบมีลักษณะโดยรวมค่อนข้างง่าย แต่สามารถจำแนกคนเก่งและคนไม่เก่งได้ อย่างไรก็ตาม ในการคัดเลือกบุคคลเข้ารับราชการในสายงานแรกบรรจุที่เริ่มต้นจากรดับ 3 หรือระดับ 4 ซึ่งเป็นสายงานทางด้านวิชาการ ลักษณะงานมีแนวโน้มที่จะมีระดับความซับซ้อนสูง การวัดความสามารถทั่วไปทางสมองด้วยข้อสอบที่ง่าย อาจไม่สะท้อนถึงความสามารถที่แท้จริงในการที่จะปฏิบัติงานที่มีลักษณะความซับซ้อนของงานสูง ให้ได้ผลดี ดังนั้นจึงควรมีการปรับปรุงแบบทดสอบให้มีจำนวนข้อสอบระดับที่ยากเพิ่มขึ้น หรืออาจนำไปใช้กับสายงานแรกบรรจุที่มีลักษณะงานมีความซับซ้อนน้อยกว่า ซึ่งต้องมีการศึกษาความเที่ยงตรงในการทำนายผลการปฏิบัติงานก่อนนำไปใช้จริง

4. การจัดชุดแบบทดสอบ

แบบทดสอบที่ผ่านการจัดชุด มีลักษณะการจัดเรียงข้อสอบแบบชุดเกลียวพอสสมควร คือในแต่ละชั้น เริ่มจากข้อที่ง่ายกว่าไปสู่ข้อที่ยากกว่า และในแต่ละชั้นมีองค์ประกอบครบตามผังโครงสร้างข้อสอบ แต่ข้อสอบจะมีลักษณะที่ค่อนข้างง่ายกว่าผังโครงสร้างข้อสอบที่กำหนด เนื่องจากข้อสอบด้านการเข้าใจภาษา ด้านอุปมาอุปไมย และด้านอนุกรมตัวเลข ที่มีคุณภาพรายชื่อผ่านเกณฑ์ มีจำนวนไม่สอดคล้องกับผังโครงสร้างข้อสอบ และเป็นข้อที่ส่วนใหญ่มี

ค่าความยากง่ายในระดับง่ายถึงง่ายมาก ดังนั้น ในการจัดชุดแบบทดสอบให้มีการจัดเรียงข้อสอบในลักษณะเกลียวให้มากที่สุด จึงจำเป็นต้องเริ่มจากข้อสอบด้านการสรุปความซึ่งมีค่าความยากง่ายในระดับง่ายมากก่อน (ค่าความยากง่ายในช่วง .81-1.0) อย่างไรก็ตาม เมื่อเปรียบเทียบกับแบบทดสอบ Wonderlic Personnel Test พบว่าข้อสอบของ Wonderlic Personnel Test ก็มีลักษณะง่ายมากจนถึงยากปานกลาง (Berry, 2003, p. 221) ซึ่งอาจเป็นไปได้ว่าแบบทดสอบ Wonderlic Personnel Test ได้รับการออกแบบให้สามารถวัดความสามารถทั่วไปทางสมองของบุคคลทุกระดับ ขอบเขตความยากของข้อสอบจึงต้องมีข้อที่ง่ายมากเพื่อวัดความสามารถของคนได้ทุกกลุ่มด้วย ดังนั้น หากมีการนำแบบทดสอบที่สร้างขึ้น ไปทดสอบกับบุคคลทุกระดับ ค่าความยากง่ายของแบบทดสอบอาจมีการเปลี่ยนแปลง ประกอบกับกลุ่มตัวอย่างที่เข้ารับการทดสอบในการวิจัยครั้งนี้ ส่วนใหญ่เป็นผู้ที่เคยสอบผ่านแบบทดสอบความสามารถทั่วไปที่ใช้ในการคัดเลือกบุคคลเข้ารับราชการมาแล้ว ดังนั้น กลุ่มตัวอย่างนี้จึงอาจเป็นกลุ่มที่มีระดับความสามารถในการทำข้อสอบสูงกว่ากลุ่มผู้สมัครเข้ารับราชการจริง หากมีการนำแบบทดสอบไปใช้กับผู้สมัครจริง ข้อสอบอาจยากขึ้น

สำหรับในการวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อสร้างเครื่องมือสำหรับการคัดเลือกบุคคลเข้ารับราชการในสายงานแรกบรรจุที่เริ่มต้นจากระดับ 3 หรือระดับ 4 การนำข้อสอบไปใช้ในการคัดเลือกจึงต้องพิจารณาถึงระดับความสามารถทั่วไปทางสมองที่เหมาะสมสำหรับสายงานแรกบรรจุที่เริ่มต้นจากระดับ 3 หรือระดับ 4 ที่มีระดับความซับซ้อนของงานสูงด้วย ดังนั้นในการพัฒนาแบบทดสอบฉบับนี้ ให้มีประสิทธิภาพในการเปรียบเทียบความสามารถของผู้สอบมากขึ้น อาจใช้วิธีการเพิ่มจำนวนข้อสอบหรือเพิ่มระดับความยากของข้อสอบ (Anastasi and Urbina, 1997, p. 102) ซึ่งผู้วิจัยเห็นว่าควรใช้ทั้งสองวิธีร่วมกัน โดยการเพิ่มขึ้นของข้อสอบที่ยากขึ้นอีกหนึ่งชั้น และกำหนดเวลาสอบเท่าเดิม

5. ความเชื่อมั่นของข้อสอบ

ผลการวิเคราะห์ค่าความเชื่อมั่นแบบแบ่งครึ่งข้อคือและข้อคู่ของแบบทดสอบความสามารถทั่วไปทางสมองฉบับจริง ด้วยวิธีการหาค่าสัมประสิทธิ์สหสัมพันธ์ตามสูตรของ สเปียร์แมนบราวน์ แบบทดสอบที่จัดชุดมีค่าความเชื่อมั่นแบบแบ่งครึ่งข้อคือและข้อคู่ประมาณ .86 แสดงว่าแบบทดสอบมีค่าความเชื่อมั่นอยู่ในระดับดี แต่ถ้าพิจารณาค่าความเชื่อมั่นแยกตามข้อสอบ ข้อสอบด้านการเข้าใจภาษาและข้อสอบด้านการสรุปความ มีค่าความเชื่อมั่นประมาณ.40

ซึ่งอยู่ในระดับที่อาจนำไปใช้ได้อย่างจำกัด ทั้งนี้ อาจเป็นเพราะในการจัดชุดข้อสอบแบบเกลียว นั้น ชุดข้อสอบด้านการเข้าใจภาษาและข้อสอบด้านการสรุปความในบางชั้นมีค่าความยากง่าย แตกต่างกันมาก ซึ่งจากตารางที่ 4.10 หน้า 64-66 และภาพที่ 4.1 หน้า 67 จะเห็นได้ว่าข้อสอบ ด้านการเข้าใจภาษาในชั้นที่ 2 (ข้อ 15 และข้อ 16) และในชั้นที่ 4 (ข้อ 43 และข้อ 45) และ ข้อสอบด้านการสรุปความในชั้นที่ 2 (ข้อ 19 และข้อ 20) ในแต่ละชั้นข้อสอบเหล่านี้มีค่าความยากง่าย แตกต่างกันมาก จึงอาจทำให้ค่าสหสัมพันธ์ระหว่างข้อคือและข้อคู่ที่คำนวณได้มีค่าต่ำ ประกอบกับ ข้อสอบด้านการเข้าใจภาษาและข้อสอบด้านการสรุปความ เป็นข้อสอบที่มีความเกี่ยวข้องกับ ภาษา ซึ่งผู้ตอบมักใช้การตีความส่วนบุคคล ความเชื่อมั่นของข้อสอบด้านนี้จึงอาจมีค่าต่ำ เนื่องจากลักษณะของการตอบข้อสอบแต่ละข้อขึ้นอยู่กับความคิดเห็นส่วนบุคคล อย่างไรก็ตาม ในช่วงกำหนดเวลาตอบที่เหมาะสม (ช่วงเวลาตอบที่ 30 นาที) แบบทดสอบมีค่าความเชื่อมั่นแบบ แปรครึ่งข้อคือและข้อคู่ประมาณ .90 แสดงว่าแบบทดสอบมีค่าความเชื่อมั่นในระดับที่ดีมาก ("Testing and assessment: An Employer's guide to good practices," online, 2000)

6. ความเที่ยงตรงตามโครงสร้างของแบบทดสอบ

การวิเคราะห์ความเที่ยงตรงตามโครงสร้างของแบบทดสอบความสามารถทั่วไปทาง สมองที่สร้างขึ้น โดยใช้วิธีการวิเคราะห์องค์ประกอบแบบยืนยัน พบว่า ลักษณะของข้อมูลไม่ สอดคล้องกับโมเดลความสามารถทั่วไปทางสมองด้านภาษา-การศึกษาตามทฤษฎีกลุ่มองค์ประกอบ ลำดับชั้นของเวอร์นอน โดยค่า goodness of fit index ซึ่งควรมีค่าเข้าใกล้ 1 มีค่าต่ำมาก และ ค่าความคลาดเคลื่อนซึ่งควรมีค่าเข้าใกล้ 0 กลับสูงมาก ทั้งนี้ อาจเป็นเพราะระหว่างตัวแปรมีความ สัมพันธ์กัน ไม่สอดคล้องกับเงื่อนไขของการวิเคราะห์องค์ประกอบเชิงยืนยันซึ่งกำหนดว่าแต่ ละตัวแปรควรเป็นอิสระจากกัน (Hair, Black, Babin, Anderson and Tatham, 2006, pp.781- 782) โดยจากผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ ระหว่างคะแนนของข้อสอบด้านการเข้าใจ ภาษา ด้านอุปมาอุปไมย และด้านการสรุปความ เพิ่มเติม (แสดงในภาคผนวก ซ) พบว่า ข้อสอบ ทุกด้านมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 โดยค่าสัมประสิทธิ์ สหสัมพันธ์ระหว่างตัวแปร มีค่าอยู่ในช่วง .24 - .57 อย่างไรก็ตาม ค่าสัมประสิทธิ์สหสัมพันธ์ ดังกล่าวอยู่ในระดับที่ไม่สูงมาก ดังนั้นสาเหตุที่ทำให้ผลการวิเคราะห์โมเดลมีค่า goodness of fit index ต่ำมาก และค่าความคลาดเคลื่อนสูงมาก จึงน่าจะเป็นผลจากลักษณะบางอย่างทางสถิติของ ข้อมูลที่ทำให้ผลการวิเคราะห์ไม่สมบูรณ์

ผลจากการวิเคราะห์องค์ประกอบเชิงสำรวจของแบบทดสอบ พบว่า องค์ประกอบ

ของแบบทดสอบในขั้นที่หนึ่ง มีโครงสร้างไม่สอดคล้องกับโมเดลองค์ประกอบที่ดัดแปลงมาจาก ทฤษฎีกลุ่มองค์ประกอบแบบลำดับขั้นของเวร์นอน โดยพบว่า ข้อสอบด้านการสรุปความ ซึ่งเป็น ข้อสอบความสามารถด้านภาษา มีค่าน้ำหนักขึ้นกับองค์ประกอบของข้อสอบความสามารถด้าน ตัวเลข ทั้งนี้ อาจเป็นเพราะข้อสอบด้านการสรุปความ เป็นข้อสอบที่ใช้หลักตรรกศาสตร์ในการค้นหา คำตอบ ซึ่งในหลักสูตรการศึกษาโดยทั่วไปกำหนดให้ตรรกศาสตร์ เป็นหัวข้อหนึ่งที่เรียนในวิชา คณิตศาสตร์ ดังนั้น จึงอาจเป็นสาเหตุหนึ่งที่ทำให้ข้อสอบด้านการสรุปความ วัดความสามารถด้าน ตัวเลขมากกว่าด้านภาษา

ส่วนผลจากการวิเคราะห์องค์ประกอบเชิงสำรวจ ของแบบทดสอบในขั้นที่สอง (second-order factor) พบว่า ทั้งองค์ประกอบของข้อสอบด้านภาษา และองค์ประกอบของ ข้อสอบด้านตัวเลข มีค่าน้ำหนักขึ้นกับองค์ประกอบความสามารถด้านภาษา-การศึกษา แสดงว่า องค์ประกอบของแบบทดสอบในลำดับขั้นนี้ สอดคล้องกับโมเดลองค์ประกอบที่ดัดแปลงมาจาก ทฤษฎีองค์ประกอบลำดับขั้นของเวร์นอน ทั้งนี้ องค์ประกอบด้านภาษามีค่าน้ำหนักกับ องค์ประกอบความสามารถด้านภาษา-การศึกษา มากกว่าองค์ประกอบด้านตัวเลข ซึ่งผล การวิเคราะห์ที่ได้สอดคล้องกับเมทริกซ์ของ gamma ที่ได้จากการวิเคราะห์องค์ประกอบเชิงยืนยัน

7. การกำหนดเวลาตอบของแบบทดสอบ

7.1 ช่วงเวลาตอบที่ 30 นาที เป็นช่วงเวลาตอบที่เหมาะสมสำหรับแบบทดสอบ โดย เป็นช่วงเวลาตอบที่คะแนนสอบมีการกระจายสูง และการกำหนดเวลาไม่มีผลกับคะแนนสอบ เนื่องจากเป็นช่วงเวลาที่จะคะแนนสอบมีความสัมพันธ์กับคะแนนสอบในช่วงที่ไม่จับเวลาอย่างมี นัยสำคัญทางสถิติ ทั้งนี้ แม้คะแนนรวมในช่วงเวลาตอบที่ 35 นาที จะมีความสัมพันธ์กับคะแนน รวมในช่วงที่ไม่จับเวลา ($r=.87$) และมีค่าเบี่ยงเบนมาตรฐานสูงสุด ($SD=7.07$) แต่ในช่วงเวลา ตอบที่ 30 นาที ก็มีค่าสถิติเหล่านี้ใกล้เคียงกันมาก ($r=.81$, $SD=6.88$) และคะแนนรวมของทั้งสอง ช่วงเวลามีค่าสัมประสิทธิ์สหสัมพันธ์ในระดับสูงมาก ($r=.97$) แสดงว่าช่วงเวลาตอบทั้ง 2 ช่วงนี้ มี ผลคะแนนที่สามารถเทียบเคียงกันได้ ดังนั้น การกำหนดเวลาตอบที่ 30 นาที จึงให้ประโยชน์ใน เรื่องของการลดเวลา ในการทดสอบมากกว่าการกำหนดเวลาตอบที่ 35 นาที

7.2 การพิจารณากำหนดเวลาตอบโดยตัดสินจากค่าเบี่ยงเบนมาตรฐาน เป็น แนวทางที่ดีสำหรับการกำหนดเวลาตอบของแบบทดสอบแบบให้เวลาจำกัดอย่างแท้จริง (pure speed test) (Nunnally and Bernstein, 1994, p. 350) แต่อาจไม่ใช่แนวทางที่ดีที่สุดสำหรับ

แบบทดสอบที่มีลักษณะของแบบทดสอบแบบให้เวลาเพียงพอ (power test) ร่วมด้วย เนื่องจากในแบบทดสอบแบบให้เวลาจำกัดอย่างแท้จริง ข้อสอบมีลักษณะง่ายมากและทุกคนสามารถทำได้ทุกข้อหากให้เวลาตอบอย่างเพียงพอ โอกาสของการได้คะแนนหรือการกระจายของคะแนนของข้อสอบประเภทนี้จึงขึ้นอยู่กับความเร็วของบุคคลแต่เพียงอย่างเดียว แต่สำหรับแบบทดสอบที่สร้างขึ้น เป็นแบบทดสอบที่มีลักษณะของแบบทดสอบแบบให้เวลาเพียงพอ (power test) ร่วมด้วย โดยข้อสอบมีลักษณะตั้งแต่ง่ายมากจนถึงค่อนข้างยาก โอกาสของการได้คะแนนในแต่ละช่วงเวลาตอบจึงไม่ได้มาจากความเร็วในการทดสอบเพียงอย่างเดียว แต่เป็นผลมาจากการเพิ่มระดับความยากของข้อสอบด้วย เมื่อเวลาในการตอบมากขึ้น ระดับความยากของข้อสอบก็จะส่งผลให้การกระจายคะแนนสอบมีมากขึ้น เพราะผู้สอบมีความสามารถในการทำข้อสอบที่ยากขึ้นได้แตกต่างกัน ดังนั้น ในช่วงเวลาตอบที่สั้นกว่าค่าส่วนเบี่ยงเบนมาตรฐานย่อมต่ำกว่าช่วงเวลาตอบที่ยาวนานกว่าเสมอ ประกอบกับในช่วงเวลาตอบที่สั้นกว่า จำนวนข้อในการทดสอบก็ย่อมน้อย โอกาสที่คะแนนจะมีความหลากหลายก็ย่อมน้อยไปด้วย

ทั้งนี้ จากการศึกษาเพิ่มเติมพบว่าวิธีการหนึ่งที่อาจนำมาประยุกต์ใช้ในการกำหนดเวลาตอบได้ คือ การเปรียบเทียบว่าช่วงเวลาตอบใดที่ใช้เวลาทดสอบน้อยที่สุดและคะแนนของผู้สอบเป็นผลมาจากความเร็วในการทดสอบน้อยที่สุด โดยในการพิจารณาว่าในแต่ละช่วงเวลาตอบคะแนนของผู้สอบเป็นผลมาจากความเร็วมากน้อยเพียงใด สามารถทำได้จากการหาค่าสัดส่วน (p) ระหว่างความแปรปรวนของจำนวนข้อที่แต่ละบุคคลทำเสร็จ (variance of number of items completed by different persons) หารด้วยความแปรปรวนของคะแนนสอบ (variance of total test scores) (Anastasi and Urbina, 1997, p. 104) ดังสูตร

$$p = \frac{SD_c^2}{SD_t^2}$$

ผลจากการคำนวณ จะบ่งชี้ถึงผลของความเร็วที่มีต่อความแปรปรวนของคะแนนสอบ ถ้าค่าสัดส่วนมีค่าเข้าใกล้ 0 มาก แสดงว่าความเร็วมีผลต่อการทดสอบต่ำ (มีลักษณะเป็น power test) แต่ถ้ามีค่าเข้าใกล้ 1 มาก แสดงว่าความเร็วมีผลต่อการทดสอบสูง (มีลักษณะเป็น speed test) ดังนั้น ในการพิจารณากำหนดช่วงเวลา จึงควรเลือกช่วงเวลาที่มีค่าสัดส่วนจากการคำนวณมีค่าต่ำที่สุดและใช้เวลาทดสอบน้อยที่สุด

8. ความเที่ยงตรงตามสภาพของแบบทดสอบ

ผลการศึกษาความเที่ยงตรงตามสภาพของแบบทดสอบโดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ระหว่างคะแนนรวมจากแบบทดสอบความสามารถทั่วไปทางสมอง และผลการปฏิบัติงานของทุกกลุ่มสายงานในภาพรวมและจำแนกตามสายงาน ไม่พบว่า คะแนนรวมของแบบทดสอบความสามารถทั่วไปทางสมองมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติกับผลการปฏิบัติงาน ทั้งในด้านผลงาน ด้านทักษะความสัมพันธ์ระหว่างบุคคล และด้านการคิด

ผลการศึกษานี้ ไม่สอดคล้องกับการศึกษาของฮันเตอร์และฮันเตอร์ (Hunter and Hunter, 1984, pp. 72-98, quoted in Ree, Carretta and Steindl, 2001, p. 225; Schmidt and Hunter, 2004, pp.162-173) และชมิทท์และฮันเตอร์ (Schmidt and Hunter, 1998, pp. 262-274) ซึ่งพบว่าความสามารถทั่วไปทางสมองมีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติในงานที่หลากหลายทุกระดับ ซึ่งความไม่สอดคล้องดังกล่าว อาจเป็นผลมาจากปัญหาของวิธีการวิจัย ดังนี้

1. ผลกระทบจากพิสัยที่ถูกจำกัด (restricted range effect) กลุ่มตัวอย่างในการวิจัยครั้งนี้ ส่วนใหญ่เป็นผู้ที่เคยสอบผ่านแบบทดสอบความสามารถทั่วไปที่ใช้ในการคัดเลือกบุคคลเข้ารับราชการมาแล้ว ดังนั้นพิสัยของคะแนนที่ได้จากแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นจึงถูกจำกัด และอาจมีผลทำให้ค่าความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบความสามารถทั่วไปทางสมองกับคะแนนผลการปฏิบัติงานมีค่าต่ำกว่าความเป็นจริง ในการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์จึงจำเป็นต้องใช้สูตรปรับแก้เพื่อลดกระทบของพิสัยของความสามารถทั่วไปทางสมองที่ถูกจำกัด ซึ่งข้อมูลที่ใช้สำหรับการคำนวณ คือ ค่าเบี่ยงเบนมาตรฐานของคะแนนสอบจากกลุ่มประชากรที่ไม่ถูกจำกัดด้วยเครื่องมือคัดเลือกประเภทเดียวกันมาก่อน (unselected population) ("Range restriction adjustments in the prediction of military ," online.)

2. การใช้ผลการปฏิบัติงานโดยหัวหน้างาน เป็นเกณฑ์ในการทำนายผลการปฏิบัติงานจากผลการศึกษาของฮันเตอร์ (Hunter, 1983a, pp. 257-266 quoted in Ree, Carretta and Steindl, 2001, pp. 227-228; Schmidt and Hunter, 2004, pp.162-173) และบอร์แมน ไวท์ ปุลาโกส และออปเพลอร์ (Borman, White, Pulakos and Oppler, 1991, 863-872) ไม่พบว่าความสามารถทั่วไปทางสมองมีอิทธิพลโดยตรงต่อการประเมินผลการปฏิบัติงานของข้าราชการพลเรือนโดยหัวหน้างาน (standardized path coefficient = .14) แต่พบว่ามีอิทธิพลโดยอ้อมผ่านการได้มาซึ่งความรู้ในงาน

(acquisition of job knowledge) (standardized path coefficient = .80) และยังพบว่า ความสามารถทั่วไปทางสมองมีอิทธิพลโดยตรงต่อผลการปฏิบัติงานจากตัวอย่างงาน (standardized path coefficient = .31) ดังนั้น การไม่พบความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบ ความสามารถทั่วไปทางสมองกับผลการปฏิบัติงานในการวิจัยครั้งนี้ อาจเป็นเพราะความสามารถทั่วไปทางสมองไม่ได้มีอิทธิพลต่อผลการปฏิบัติงานที่ได้จากการประเมินโดยผู้บังคับบัญชาระดับต้นโดยตรง แต่มีอิทธิพลโดยอ้อมผ่านการได้มาซึ่งความรู้ในงาน การหาความเที่ยงตรงตามสภาพของแบบทดสอบโดยใช้ผลการปฏิบัติงานจากผู้บังคับบัญชาระดับต้น (หัวหน้างาน) เป็นเกณฑ์ จึงควรศึกษาอิทธิพลของความรู้ในงานที่อาจมีผลต่อการอธิบายความสัมพันธ์ระหว่างแบบทดสอบ ความสามารถทั่วไปทางสมองและผลการปฏิบัติงานที่ประเมินโดยผู้บังคับบัญชาระดับต้นด้วย หรืออาจใช้ผลการปฏิบัติงานจากแหล่งอื่นเช่นจากตัวอย่างงานเป็นเกณฑ์ในการหาความเที่ยงตรงตามสภาพของแบบทดสอบแทนผลการปฏิบัติงานที่ประเมินโดยผู้บังคับบัญชาระดับต้น

3. แบบประเมินผลการปฏิบัติงานที่ใช้เป็นเกณฑ์ในการหาความเที่ยงตรงตามสภาพของแบบทดสอบในการวิจัยนี้ เป็นแบบประเมินผลการปฏิบัติงานโดยรวมสำหรับทุกสายงาน ข้อประเมินบางข้อจึงอาจไม่สอดคล้องกับลักษณะงานที่ปฏิบัติจริง ทำให้ผู้บังคับบัญชาไม่มีข้อมูลเพียงพอที่จะประเมินผลการปฏิบัติงานในข้อประเมินนั้นได้ ดังจะเห็นได้จากผลการวิเคราะห์ ข้อประเมินที่ผู้บังคับบัญชาประเมินว่า “ข้อมูลไม่พอ” (แสดงในภาคผนวก ญ) ซึ่งพบว่า องค์ประกอบด้านทักษะความสัมพันธ์ระหว่างบุคคล และองค์ประกอบด้านการคิด มีร้อยละของผู้ที่ประเมินว่า “ข้อมูลไม่พอ” ค่อนข้างสูง ผลการวิเคราะห์ข้อประเมินดังกล่าวแสดงให้เห็นว่า แต่ละสายงานอาจไม่มีโอกาสที่จะแสดงพฤติกรรมทุกข้อ เนื่องจากข้อประเมินบางข้อไม่สอดคล้องกับหน้าที่ในงาน เช่น บางสายงานอาจไม่มีหน้าที่ให้บริการ หรืออาจไม่มีบทบาทในการเสนองานในที่ประชุม เป็นต้น ความไม่สอดคล้องกันระหว่างลักษณะงานที่ปฏิบัติจริงกับข้อประเมินในแบบประเมินผลการปฏิบัติงาน จึงอาจเป็นสาเหตุหนึ่งที่ทำให้ไม่พบความสัมพันธ์ระหว่างคะแนนจากแบบทดสอบความสามารถทั่วไปทางสมองและผลการปฏิบัติงาน ดังนั้น แบบประเมินผลการปฏิบัติงานที่ใช้เป็นเกณฑ์ในการหา ความเที่ยงตรงตามสภาพของแบบทดสอบจึงควรเป็นแบบประเมินที่ข้อประเมินสอดคล้องกับหน้าที่หรือลักษณะงานในแต่ละสายงานโดยเฉพาะ อย่างไรก็ตาม การประเมินว่า “ข้อมูลไม่พอ” อาจเกิดขึ้นได้ด้วยในกรณีที่ข้อประเมินบางข้อเป็นการประเมินพฤติกรรมที่สังเกตได้ไม่ชัดเจน (เช่น ข้อที่ 52 จัดการกับความเครียดของตนเองได้) และในกรณีที่ผู้บังคับบัญชาไม่ได้ใกล้ชิดเพียงพอที่จะเห็นการทำงานของผู้ใต้บังคับบัญชา (เช่น ข้อคำถามที่ 41 “ในการให้บริการ เข้าไปถามผู้รับบริการก่อนที่ผู้รับบริการจะถาม”) การสร้างหรือ

ปรับปรุงแบบประเมินผลการปฏิบัติงาน จึงควรสร้างหรือปรับปรุงข้อประเมินให้เป็นพฤติกรรมที่สังเกตได้ชัดเจน และควรจะต้องมีการชี้แจงผู้บังคับบัญชาให้เข้าใจหลักของการประเมินผลการปฏิบัติงานด้วย เช่น การสังเกตและจดบันทึกพฤติกรรมการทำงานของผู้ใต้บังคับบัญชาอย่างสม่ำเสมอ การระมัดระวังเกี่ยวกับอคติในการประเมินผลการปฏิบัติงาน เป็นต้น เพื่อให้ความคลาดเคลื่อนของการประเมิน ผลการปฏิบัติงานโดยผู้บังคับบัญชาเกิดขึ้นน้อยที่สุด

เหตุผลดังกล่าวข้างต้น จึงอาจยังไม่สามารถสรุปได้ว่าแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้น มีความเที่ยงตรงตามสภาพในการทำนายผลการปฏิบัติงานในระดับใด ดังนั้นจึงควรมีการศึกษาเพิ่มเติมเพื่อยืนยันถึงคุณภาพของแบบทดสอบก่อนที่จะมีการนำแบบทดสอบที่สร้างขึ้นไปใช้งานจริง

ข้อเสนอแนะ

จากการสร้างและพัฒนาแบบทดสอบความสามารถทั่วไปทางสมองที่วัดทั้งความสามารถและความรวดเร็วในการตอบ สำหรับข้าราชการพลเรือนระดับแรกบรรจุ ผู้วิจัยมีข้อเสนอแนะในการนำแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างไปใช้ในการคัดเลือก และมีข้อเสนอแนะในการวิจัยครั้งต่อไป ดังนี้

1. การนำแบบทดสอบความสามารถทั่วไปทางสมองไปใช้ในการคัดเลือก

การนำแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นไปใช้ในการคัดเลือกข้าราชการพลเรือนระดับแรกบรรจุ ควรมีการดำเนินการดังนี้

1.1 ตรวจสอบความเที่ยงตรงตามสภาพของแบบทดสอบซ้ำ เพื่อยืนยันถึงคุณภาพของแบบทดสอบในการทำนายผลการปฏิบัติงาน โดยการนำปัจจัยด้านความรู้ในงานมาใช้ในการทำนายผลการปฏิบัติงานที่ประเมินโดยผู้บังคับบัญชาระดับต้น ร่วมกับแบบทดสอบความสามารถทั่วไปทางสมอง หรืออาจใช้ผลการปฏิบัติงานจากตัวอย่างงานเป็นเกณฑ์ผลการปฏิบัติงาน และวิเคราะห์สมการถดถอยที่มีคะแนนจากแบบทดสอบเป็นตัวทำนายผลการปฏิบัติงาน เพื่อใช้สำหรับการกำหนดเกณฑ์การคัดเลือก

1.2 จัดทำคู่มือแบบทดสอบเพื่อใช้เป็นแหล่งข้อมูลและมาตรฐานสำหรับการดำเนินการทดสอบ ซึ่งเนื้อหาในคู่มือควรรวให้ข้อมูลที่ชัดเจนเกี่ยวกับขั้นตอนการสร้าง

แบบทดสอบ คุณภาพของแบบทดสอบทั้งในด้านความเชื่อมั่นและความเที่ยงตรง ขั้นตอนในการดำเนินการทดสอบ กำหนดเวลาตอบ ปัญหาและข้อควรระวังเกี่ยวกับการใช้แบบทดสอบ การตรวจให้คะแนน และการแปลผลคะแนน

1.3 กำหนดวิธีการนำคะแนนที่ได้จากแบบทดสอบมาใช้งาน เพื่อกำหนดรูปแบบในการรายงานผลการทดสอบ เช่น ถ้าต้องการนำคะแนนไปใช้ตัดสินผู้สมัครโดยการจัดเรียงอันดับ (ranking) หรือต้องการนำคะแนนไปให้น้ำหนัก หรือรวมกับคะแนนจากเครื่องมืออื่น ควรใช้คะแนนรวมในการรายงานผล แต่ถ้าหากต้องการใช้คะแนนตัด (cut-scores) ควรรายงานผลคะแนนเป็นผ่านหรือไม่ผ่าน โดยอาจกำหนดคะแนนตัดโดยเทียบจากเกณฑ์ปกติของแบบทดสอบ

1.4 จัดเก็บข้อมูลผลการสอบของข้าราชการในสายงานต่างๆ เพื่อใช้ในการพัฒนาเกณฑ์ปกติให้เป็นปัจจุบัน และสมบูรณ์มากที่สุด ทั้งนี้ เมื่อมีกลุ่มตัวอย่างจำนวนมากพอ ควรมีการพัฒนาเกณฑ์ปกติโดยรวมเพศและกลุ่มอายุเข้าด้วยกัน เช่น เกณฑ์ปกติสำหรับเพศชาย อายุต่ำกว่า 31 ปี เพศหญิง อายุต่ำกว่า 31 ปี และควรมีการจำแนกเกณฑ์ปกติสำหรับสายงานด้วย เพื่อให้สามารถนำไปใช้แปลผลคะแนนได้เหมาะสมกับกับกลุ่มคนมากยิ่งขึ้น

1.5 สร้างแบบทดสอบชุดทางเลือก (alternative forms) เพื่อรองรับจำนวนการทดสอบที่อาจเพิ่มมากขึ้น เนื่องจากการใช้แบบทดสอบชุดเดียวซ้ำๆ อาจทำให้ข้อสอบจำเจ ซึ่งส่งผลต่อความเชื่อมั่นและความเที่ยงตรงของแบบทดสอบในการใช้ทดสอบ

1.6 พัฒนารูปแบบในการดำเนินการสอบแบบทดสอบไปสู่ระบบการสอบด้วยคอมพิวเตอร์ เพื่อให้สามารถดำเนินการสอบและประมวลผลการสอบได้สะดวกมากยิ่งขึ้น

1.7 ติดตามประเมินผลการนำแบบทดสอบไปใช้ เพื่อรวบรวมปัญหาสำหรับการพัฒนาแบบทดสอบ

2. การวิจัยครั้งต่อไป

2.1 การวิจัยเพื่อพัฒนาแบบทดสอบให้มีประสิทธิภาพ

2.1.1 ทำการวิจัยทำนองเดียวกันนี้ โดยปรับปรุงแบบทดสอบให้มีคุณภาพรายข้อมากขึ้น แต่ใช้วิธีการกำหนดเวลาตอบโดยเลือกช่วงเวลาตอบที่ใช้เวลาทดสอบน้อยที่สุดและคะแนนของผู้สอบเป็นผลมาจากความเร็วในการทดสอบน้อยที่สุด ตามวิธีการที่เสนอในกรอภิปราชข้อ 7.2 เพื่อให้สามารถกำหนดเวลาตอบที่รวดเร็วมากยิ่งขึ้น ทั้งนี้ อาจรวบขั้นตอนของการจัดทดสอบเพื่อกำหนดเวลาตอบ เพื่อตรวจสอบความเที่ยงตรงตามสภาพ และเพื่อสร้างเกณฑ์

ปกติเข้าด้วยกัน กล่าวคือ จัดทดสอบครั้งเดียวกัน แต่จับเวลาสอบเป็นช่วงคล้ายกับการวิจัยครั้งนี้ และเปรียบเทียบผลของช่วงเวลาตอบแต่ละช่วงในการทำนายผลการปฏิบัติงาน ซึ่งหากช่วงเวลาตอบแต่ละช่วงสามารถทำนายผลการปฏิบัติงานได้แตกต่างกัน ก็อาจใช้เป็นข้อพิจารณาเพิ่มเติมสำหรับการกำหนดเวลาตอบ

2.1.2 ศึกษาองค์ประกอบของข้อสอบรูปแบบต่างๆ ที่เกี่ยวข้องกับ การวัดความสามารถทั่วไปทางสมอง เพื่อใช้สำหรับการสร้างและพัฒนาแบบทดสอบให้วัดองค์ประกอบได้ตรงตามที่ต้องการ เนื่องจากผลการวิจัยครั้งนี้พบว่า รูปแบบข้อสอบด้านการสรุปความ (inference) ที่โดยทั่วไปเชื่อว่าวัดความสามารถด้านภาษา มีน้ำหนักในการวัดองค์ประกอบของความสามารถด้านตัวเลขมากกว่าด้านภาษา

2.2 การวิจัยเพื่อตรวจสอบคุณภาพของแบบทดสอบเพิ่มเติม

2.2.1 ศึกษาความเชื่อมั่นของแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นนี้ โดยใช้วิธีการอื่นๆ เช่น วิธีสอบซ้ำ (test retest method) วิธีแบ่งครึ่งข้อสอบ (split-half time method) ทั้งแบบแบ่งข้อ และแบ่งตามเวลา (Anastasi and Urbina, 1997, p. 104)

2.2.2 ศึกษาความเที่ยงตรงเชิงเอคนัย (convergent validity) ของแบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้น โดยนำไปหาค่าสัมประสิทธิ์สหสัมพันธ์กับแบบทดสอบที่วัดความสามารถทั่วไปทางสมองอื่นๆ เช่นแบบทดสอบภาคความรู้ความสามารถทั่วไปของภาครัฐ เพื่อยืนยันถึงคุณภาพของแบบทดสอบว่าวัดความสามารถทั่วไปทางสมองได้สอดคล้องกับเครื่องมืออื่นๆ

2.2.3 ศึกษาความเที่ยงตรงตามสภาพของแบบทดสอบที่สร้างขึ้นในระดับสายงาน โดยใช้การประเมินผลการปฏิบัติงานที่ได้จากผู้บังคับบัญชาระดับต้นเป็นเกณฑ์และจำแนกแบบประเมินผลการปฏิบัติงานให้มีความสอดคล้องกับสายงาน ซึ่งแบบประเมินดังกล่าวอาจได้มาจากการวิเคราะห์งานหรือจากมาตรฐานกำหนดตำแหน่ง ทั้งนี้ ควรตรวจสอบข้อประเมินให้เป็นการประเมินพฤติกรรมที่สังเกตได้ชัดเจน และชี้แจงผู้บังคับบัญชาให้เข้าใจหลักการประเมินผลการปฏิบัติงานด้วย เพื่อให้ความคลาดเคลื่อนของการประเมินผลการปฏิบัติงานเกิดขึ้นน้อยที่สุด และอาจเพิ่มตัวแปรความรู้ในงานในการทำนายผลการปฏิบัติงานร่วมกับแบบทดสอบความสามารถทั่วไปทางสมอง เพื่อตรวจสอบว่าคะแนนจากแบบทดสอบความสามารถทั่วไปทางสมองสามารถทำนายผลการปฏิบัติงานที่ได้จากผู้บังคับบัญชาระดับต้นโดยผ่านความบังเจ้ายด้านความรู้ในงานหรือไม่ นอกจากนี้ ควรมีการเก็บข้อมูลผลการทดสอบในกลุ่มประชากรที่ไม่ถูกจำกัดจากเครื่องมือคัดเลือกประเภทแบบทดสอบความสามารถทั่วไปทางสมอง (unselected

population) ด้วย เพื่อให้ได้ข้อมูลเกี่ยวกับค่าเบี่ยงเบนมาตรฐานสำหรับการคำนวณเพื่อปรับแก้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบที่สร้างขึ้นกับผลการปฏิบัติงาน โดยกลุ่มประชากรดังกล่าว อาจเป็นกลุ่มผู้สมัครสอบเข้ารับราชการจริง หรือกลุ่มนักศึกษาที่กำลังจะจบการศึกษาในระดับปริญญาตรี หรือระดับปริญญาโท ซึ่งเป็นกลุ่มบุคคลที่มีสิทธิ์สอบเข้ารับราชการในอนาคต

2.2.4 ศึกษาความเที่ยงตรงตามสภาพของแบบทดสอบที่สร้างขึ้นนี้ ในการทำนายผลการปฏิบัติงานในสายงานหรือกลุ่มสายงานอื่นของข้าราชการพลเรือน และสร้างเกณฑ์ปกติเพื่อใช้ในการคัดเลือกบุคคลเข้าสู่ตำแหน่ง

2.2.5 ศึกษาความเที่ยงตรงตามสภาพของแบบทดสอบที่สร้างขึ้นนี้ ในการทำนายผลการปฏิบัติงานในภาคเอกชน เพื่อเปรียบเทียบกับผลการศึกษาในองค์การภาครัฐ

2.2.6 ศึกษาผลของการใช้แบบทดสอบความสามารถทั่วไปทางสมองที่สร้างขึ้นนี้ ในการทำนายผลการปฏิบัติงานร่วมกับเครื่องมือในการคัดเลือกอื่นๆ เพื่อสร้างสมการสำหรับการทำนายผลการปฏิบัติงาน