

การวิจัยอื่นที่เกี่ยวข้องกับการวิจัยนี้

การวิจัยที่เกี่ยวกับแบบทดสอบสติปัญญา ภาคคำศัพท์ เท่าที่พบตามค้นคว้ามาสำหรับในประเทศไทย ปรากฏว่ายังไม่มีผู้ใดทำการวิจัยเรื่องนี้ เพื่อนำออกมาใช้ทดสอบสติปัญญาของเด็กไทยเคย โดยเฉพาะแบบทดสอบสติปัญญาภาคคำศัพท์ของ เวดสเตอร์ ทั้งนี้ ผลการวิจัยที่เกี่ยวข้องกับการวิจัยนี้ ส่วนใหญ่จึงเป็นผลการวิจัยของนักวิจัยชาวต่างประเทศ ดังจะกล่าวต่อไป

แบบทดสอบสติปัญญาที่ใช้ในกันในปัจจุบันนี้ ได้มีการพัฒนาขึ้นมานับตั้งแต่ปีเนต ได้คิดเครื่องมือทดสอบสติปัญญาขึ้นในปี ค.ศ. 1905

ต่อมาในปี ค.ศ. 1916 เทอร์แมน¹ (Terman) ได้ปรับปรุงแบบทดสอบสติปัญญาของบีเนต และซีมอน (Simon) เพื่อให้แบบทดสอบนี้มีความสมบูรณ์และใช้ได้กว้างขวางยิ่งขึ้น โดยเพิ่มจำนวนข้อทดสอบจากแบบทดสอบเดิมของบีเนต ซึ่งมีเพียง 45 ข้อ เป็น 90 ข้อ แต่ยังคงรักษาระดับสมมติฐานเดิมของแบบทดสอบไว้ และได้เพิ่มความระมัดระวังในการเลือกข้อตัวอย่างให้รัดกุม เพื่อให้กลุ่มตัวอย่างที่ได้เป็นตัวแทนของประชากร ซึ่งนับว่าเป็นการริเริ่มที่จะควบคุมการเลือกตัวอย่างขึ้นเป็นครั้งแรก รวมทั้งการปรับปรุงข้อทดสอบให้มีความเที่ยงตรงและเชื่อถือได้มากขึ้น มีการนำเอาเกณฑ์ภาคเชาวน์ (IQ) เข้ามาใช้เป็นครั้งแรกในแบบทดสอบปี ค.ศ. 1916

ในปี ค.ศ. 1937 เทอร์แมน และเมอร์ริลล์² (Merrill) ได้ทำการปรับปรุงแบบทดสอบสติปัญญาของบีเนตอีกเป็นครั้งที่สอง โดยมีวัตถุประสงค์ที่จะให้บุคคลทั่วไปเข้าใจถึงคุณลักษณะที่แท้จริง

¹ L.M. Terman and M.A. Merrill., Stanford Binet Intelligence Scale, Manual , 3rd Rev., Boston (Houghton Mifflin Co. 1960) pp 5 - 7

² Ibid., pp 8 - 11

ของแบบทดสอบนี้ แบบทดสอบในปี 1937 แสดงให้เห็นว่า คำศัพท์ซึ่งเป็นข้อทดสอบย่อย (Subtest) ส่วนหนึ่งของแบบทดสอบ มีความเกี่ยวข้องกับความสามารถทางสติปัญญาของบุคคล เช่น เด็กอายุ 6 ปี มีค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคำศัพท์กับคะแนนรวมเป็น 0.65

ในปี ค.ศ. 1960 เฮอร์แมน และเมอร์ริล³ ได้ทำการปรับปรุงแบบทดสอบของบีเนตออกเป็นครั้งที่ 3 ซึ่งเป็นแบบทดสอบที่ใช้อยู่ในปัจจุบัน การปรับปรุงครั้งนี้ได้เน้นถึงอำนาจจำแนกของข้อทดสอบด้วย เพื่อแสดงให้เห็นลักษณะของสติปัญญาซึ่งมีอยู่หลายประการควบกัน ปรากฏว่าแบบทดสอบฉบับปี 1960 นี้ คะแนนที่เด็กทำได้สูงขึ้นกว่าแบบทดสอบฉบับปี 1937

คำศัพท์ในแบบทดสอบของบีเนต มีจุดมุ่งหมายเพื่อต้องการทราบว่า บุคคลรู้จักความหมายของคำศัพท์หรือไม่ โดยไม่คำนึงถึงว่าจะต้องเป็นคำศัพท์ที่ถูกสอนตามหลักสูตรและปรัจฉา คำศัพท์ในแบบทดสอบทั้งหมดมี 45 คำ ระดับการให้คะแนนก็ไม่เท่ากัน เช่น เด็กอายุ 6 ปีตอบถูก 6 คำ ถือว่าใช้ได้ เด็กอายุ 8 ปีต้องตอบถูก 8 คำ เด็กอายุ 10 ปีต้องตอบถูก 11 คำ ดังนี้ เป็นต้น

สำหรับคำศัพท์ในแบบทดสอบปี 1960 ปรากฏว่าส่วนใหญ่มีค่าความเชื่อถือได้ (Reliability) สูงกว่าแบบทดสอบปี 1937 เช่นแบบทดสอบฉบับปี 1937 เด็กอายุ 6 ปี มีค่าสัมประสิทธิ์ของความเชื่อถือได้ (Coefficient of Reliability) เท่ากับ .65 ($r = .65$) แต่แบบทดสอบฉบับปี 1960 มีค่าสัมประสิทธิ์ของความเชื่อถือได้เท่ากับ .67 ($r = .67$) โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์แบบไบเซรียล (Biserial Correlation) แสดงให้เห็นว่า การมีการปรับปรุงข้อทดสอบอยู่เสมอ จะทำให้แบบทดสอบนี้มีความเชื่อถือได้มากขึ้น

³ Ibid pp 13 - 40

นอกจากแบบทดสอบของบีเนตต์แล้ว ในปี ค.ศ. 1939 เวคสเลอร์ ได้สร้างแบบทดสอบสติปัญญา (Wechsler Bellevue Intelligence Scale (W - B Scale) ขึ้นโดยแบ่งออกเป็น 2 หมวดใหญ่ ๆ คือ หมวดคำพูด (Verbal) และหมวดการกระทำ (Performance) ซึ่งประกอบด้วยข้อทดสอบย่อย 11 หัวข้อ และค่าที่พบก็เป็นข้อทดสอบย่อยหัวข้อหนึ่งในจำนวนข้อทดสอบย่อยทั้งหมด จัดไว้ในหมวดคำพูด

ต่อมาในปี ค.ศ. 1949 เวคสเลอร์ ได้สร้างแบบทดสอบสติปัญญาสำหรับเด็ก (Wechsler Intelligence Scale for Children) มีชื่อย่อซึ่งเป็นที่รู้จักกันดีว่า วิสค์ (WISC) ซึ่งเป็นแบบทดสอบที่ผู้เขียนใช้เป็นหลักในการศึกษาและวิจัยครั้งนี้

วิสค์ เป็นแบบทดสอบที่พัฒนามาจากแบบทดสอบสติปัญญา (W - B Scale) แม้ว่าโดยลักษณะทั่วไปจะคล้ายคลึงกัน แต่วิสค์ก็เป็นแบบทดสอบที่แยกออกต่างหากจากแบบทดสอบสติปัญญา (W - B Scale) และมีมาตรฐานของตนเองเป็นแบบอิสระ แบบทดสอบวิสค์นี้ประกอบด้วยข้อทดสอบย่อย 12 หัวข้อ⁵ ซึ่งแบ่งออกเป็น 2 หมวด คือ หมวดคำพูด (Verbal) และหมวดการกระทำ (Performance) เช่นเดียวกับแบบทดสอบสติปัญญา (W - B Scale) ดังนี้

หมวดคำพูด (Verbal)	หมวดการกระทำ (Performance)
1. ความรู้ทั่วไป (General Information)	7. การเติมรูปภาพ (Picture Completion)
2. ความเข้าใจทั่วไป (General Comprehension)	8. การเรียงคำกับรูปภาพ (Picture Arrangement)
3. เลขคณิต (Arithmetic)	9. การเรียงคูกบาทตามแบบ (Block Design)
4. ความคล้ายคลึง (Similarities)	10. การต่อรูป (Object Assembly)
5. คำศัพท์ (Vocabulary)	11. การเขียนรหัส (Coding)
6. การอ่านตัวเลข (Digit Span)	12. การเขียนทางวนเวียน (Maze)

⁴ David Wechsler, Wechsler Intelligence Scale for Children, Manual, New York, (The Psychological Corporation, 1949) pp 1 - 7

⁵ Ibid. p 6

สำหรับการอ่านตัวเลขและการเขียนทางวงเวียนเป็นข้อทดสอบย่อยที่จะนำมาใช้แทนข้อทดสอบอื่น ๆ ใน 10 หัวข้อนั้น ในกรณีที่ข้อทดสอบเหล่านี้ใช้ไม่ได้ผล เพราะไม่เหมาะสมกับบุคคลทดสอบ อันเนื่องมาจากสาเหตุทางวัฒนธรรม เป็นต้น แต่อาจขอทดสอบทั้ง 10 หัวข้อใช้กันไต่แล้วจะไม่ใช้การอ่านตัวเลขแทน แม้ว่าข้อทดสอบจะทำให้คะแนนค่าเพียงใดก็ตาม ทั้งนี้ เพราะความสัมพันธ์ของสหสัมพันธ์ระหว่างการอ่านตัวเลขกับข้อทดสอบย่อยข้ออื่น ๆ อย่างไรก็ดี แบบทดสอบวิเศษที่สถาบันต่าง ๆ ของไทยนำมาใช้นั้นปรากฏว่าในภาคคำศัพท์ (Vocabulary) นำมาใช้ไม่ได้ เพราะความแตกต่างทางด้านภาษาและวัฒนธรรมดังกล่าว ซึ่งเป็นเหตุทำให้ต้องใช้การอ่านตัวเลขแทนภาคคำศัพท์

จากการศึกษาของเวคสเลอร์ ปรากฏว่าสัมพันธ์ของสหสัมพันธ์ระหว่างคำศัพท์กับข้อทดสอบย่อยอื่น ๆ มีค่าสูง ดังแสดงในตารางข้างล่างนี้

ตารางที่ 1 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคำศัพท์กับข้อทดสอบย่อยอื่น ๆ ในแบบ

ทดสอบของเวคสเลอร์

Vocabulary Age	Information	Comprehension	Arithmetic	Similarities	Picture Completion	Picture Arrangement
10 ปี	.75	.75	.62	.64	.47	.56
13 ปี	.74	.65	.46	.66	.36	.44

นอกจากนี้ ค่าสัมประสิทธิ์ของสหสัมพันธ์ระหว่างคำศัพท์กับส่วนที่ทดสอบโดยใช้ คำพูดสูงกว่าค่าสัมประสิทธิ์ของสหสัมพันธ์ระหว่างคำศัพท์กับส่วนการทดสอบโดยการกระทำ นอกจากนี้ ค่าสัมประสิทธิ์ของสหสัมพันธ์ระหว่างคำศัพท์กับคะแนนรวม (Full Scale Score) ก็มีค่าสูง ดังแสดงไว้ในตารางที่ 2

ตารางที่ 2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคำศัพท์กับคะแนนของหมวดคำพูด (Verbal Scale Score), คะแนนของหมวดการกระทำ (Performance scale Score) และคะแนนรวม (Full Scale Score) ในระดับอายุต่าง ๆ กัน 004365

Vocabulary Age	Verbal Scale Score	Performance Scale Score	Full Scale Score
10 ปี	.82	.68	.83
13 ปี	.75	.51	.70

จากการหาความเชื่อถือได้แบบแบ่งครึ่ง (Split - Half) ปรากฏว่าคำศัพท์ใน วิสทัศน์นี้มีค่าความเชื่อถือได้ในระดับสูง ตัวอย่างเช่น ในเด็กอายุ 7 ปี 10 ปี และ 13 ปี มีค่าสัมประสิทธิ์ความเชื่อถือได้ เท่ากับ .77 , .91 และ .90 ตามลำดับ

การเลือกตัวอย่างในวิสทัศน์ เวกส์เลอร์คำนึงถึงลักษณะทางภูมิศาสตร์ อาชีพของบิดา มารดา และอายุของเด็ก เวกส์เลอร์ใช้กลุ่มตัวอย่างระดับอายุละ 200 คน เป็นชาย 100 คน หญิง 100 คน เป็นเด็กอายุตั้งแต่ 5 ปี ถึง 15 ปี รวมจำนวนตัวอย่างทั้งหมด 2200 คน เด็ก เหล่านี้มาจากรัฐต่าง ๆ 11 รัฐ และจากโรงเรียน 85 แห่ง ในจำนวนเด็กทั้งหมดนี้มีเด็กปัญญา อ่อนรวมอยู่ด้วย 55 คน

*Ibid pp. 98 - 99

ในแบบทดสอบวิเศษ มีคำศัพท์ทั้งหมด 40 คำ คะแนนเต็ม 80 คะแนน และในหลักการให้คะแนนจะให้คะแนนเป็น 2, 1 และ 0 ตามลำดับความถูกต้องและกระชับแจ่มชัดของคำตอบ และเพื่อให้การให้คะแนนเป็นไปอย่างถูกต้องและยุติธรรมที่สุด เวคส์เลอร์ จึงได้วางหลักเกณฑ์ในการให้คะแนนคำศัพท์ดังนี้.-

คำตอบที่จะให้คะแนนเต็ม 2 ควรจะมีลักษณะดังนี้.-

1. เป็นไวยากรณ์ที่ถูกต้อง
2. เป็นการนำส่วนที่สำคัญมาใช้
3. เป็นการให้คำนิยามลักษณะเฉพาะ หรือลักษณะเบื้องต้นของคำศัพท์นั้นๆ หรือมากกว่าหนึ่ง
4. เป็นการแบ่งประเภทของคำศัพท์โดยทั่ว ๆ ไปตามลักษณะของคำศัพท์
5. ใช้สัญลักษณ์ของคำศัพท์อย่างถูกต้อง
6. ในอำนันามหลาย ๆ อย่างที่คลุมเครือ แต่เป็นการอธิบายลักษณะโดยถูกต้องซึ่งจะแสดงให้เห็นถึงความเข้าใจคำศัพท์นั้น
7. (คำศัพท์เป็นคำกริยา) ให้นิยามตัวอย่างของการกระทำที่เป็นเหตุของความสัมพันธ์กัน

คำตอบที่จะได้คะแนน 1 ควรจะมีลักษณะดังนี้.-

ลักษณะคำตอบทั่ว ๆ ไปที่จะได้ 1 คะแนน คือคำตอบที่ไม่ได้แต่มีเนื้อความที่ไม่ถูกต้องสมบูรณ์ลักษณะต่อไปนี้.-

1. เป็นไวยากรณ์ที่คลุมเครือ หรือไม่ตรงเรื่อง
2. เป็นการนำส่วนที่ไม่สำคัญมาใช้ โดยไม่มีการขยายความ
3. อธิบายลักษณะโดยถูกต้อง แต่ไม่แจ่มแจ้งหรือแบ่งแยกลักษณะอย่างชัดเจน
4. ให้ตัวอย่างการใช้คำศัพท์นั้นโดยไม่มีการขยายความ

คำตอบที่ได้คะแนน 0 ควรจะมีลักษณะดังนี้.-

1. เป็นคำตอบที่เฉื่อย
2. อธิบายคำศัพท์โดยยาว โดยไม่แสดงความเข้าใจความหมายของคำศัพท์นั้น
3. เป็นคำตอบที่ไม่ถึงกับผิดความหมาย แต่เมื่อจัดกลุ่มให้อธิบายเพิ่มเติมก็ให้คำอธิบายอย่างคลุมเครือ ไม่เป็นสาระหรือมีเนื้อความไม่สมบูรณ์

การให้คะแนนจากข้อทดสอบแบบเติมความนี้ มีผู้ศึกษาพบว่ามี ความเชื่อถือได้ ในปี ค.ศ. 1957 แกรนท์ (Grant) และแคปแลน (Caplan) ได้ศึกษาพบว่า การสร้างคำถามเพื่อให้อุตอบ เพียงสั้น ๆ และให้คะแนน "2" สำหรับคำตอบที่ถูกต้อง คะแนน "1" สำหรับคำตอบที่ไม่ใคร่สมบูรณ์นัก และคะแนน "0" สำหรับคำตอบที่ใช้ไม่ได้ รวมทั้งผู้ที่ไม่ได้ตอบด้วย โดยผู้ศึกษาให้คะแนน 2 คน ปรากฏผลว่า ค่าสัมประสิทธิ์ของสหสัมพันธ์ของคะแนนที่ได้รับจากผู้ตรวจ 2 คนนั้นมีความเชื่อถือได้สูงพอสมควร คือได้ค่าสัมประสิทธิ์ของสหสัมพันธ์เป็น .78 , .96 , และ .87 จากคะแนน 2 วิชานี้ ในตอนกลางเทอมและตอนปลายเทอม จากผลการศึกษาชี้แจงว่า ข้อสอบแบบเติมความนี้ให้ คำตอบสั้น ๆ สามารถสร้างใหม่มีความเชื่อถือได้ในการให้คะแนน ดังเช่นแบบทดสอบภาคคำศัพท์ของ วิศล เป็นต้น

ถึงได้กล่าวมาแล้วว่า คำศัพท์นั้นขึ้นอยู่กับภาษาและวัฒนธรรม ตลอดจนสิ่งแวดล้อม ในปี ค.ศ. 1957 เลวินสัน⁸ (Levinson) ได้ทำการศึกษาพบว่า เด็กที่มาจากสภาพสังคมและเศรษฐกิจ- กิจจะทำการคะแนนหมวดคำพูดได้สูงกว่าเด็กที่มาจากสิ่งแวดล้อมที่ไม่ดี

⁷ Donald L. Grant and Nathan Caplan, "Studies in The Reliability of the Short - Answer Examination" J. of Ed. Res. Vol. 51, (1957) pp 109 - 116

⁸ Boris M. Levinson., "Traditional Jewish Cultural Values and Performance on the Wechsler Tests" J. of Ed. Psychol. 50 (Jan' 59) 61 - 65

ในปี ค.ศ. 1959 เควินสัน ได้ทำการทดสอบกับเด็กยิวในอิสราเอลพบว่า วัฒนธรรมมีอิทธิพลที่จะทำให้เกิดความแตกต่างระหว่างคะแนนหมวดคำพูด และคะแนนหมวดการกระทำในแบบทดสอบของเวกสเลอร์ได้

ในปี ค.ศ. 1961 ชวาล แพรัตกุล⁹ ได้ทำการสำรวจประสิทธิภาพของเครื่องมือทดสอบสำหรับใช้กับนิสิตปีที่ 1 ของมหาวิทยาลัยอินเดียนา เพื่อนำมาใช้ในโครงการแนะแนวพบว่า แบบทดสอบคำศัพท์เป็นแบบทดสอบที่ใช้ทำนายผลการเรียนได้ดีเป็นอันดับ 2 ในบรรดาแบบทดสอบทั้งหมด 5 แบบ

ในปี ค.ศ. 1964 เควินสัน¹⁰ ได้มีการศึกษาและพบว่า วัฒนธรรมมางประการมีอิทธิพลต่อการทำข้อทดสอบของเวกสเลอร์ โดยศึกษาจากกลุ่มคนผิวขาวที่มีที่อยู่อาศัย, คนผิวขาวที่ไร้ที่อยู่อาศัย และชนนิโกร ผลปรากฏว่า คะแนนจากข้อทดสอบของคะแนนรวมของกลุ่มคนผิวขาวที่มีที่อยู่อาศัยสูงกว่าคะแนนรวมของชนนิโกร และคนผิวขาวที่ไร้ที่อยู่อาศัยแสดงว่า ภาวะของสังคม เศรษฐกิจ และวัฒนธรรม มีส่วนเกี่ยวกับการทำแบบทดสอบของเวกสเลอร์มาก

⁹Chawal Paeratakul, "An investigation of the efficiency of the Indiana University Freshmen Orientation Test Battery and its implication for Counseling and Guidance" (Unpublished Ph. D. Thesis, Indiana University, 1961)

¹⁰Boris M. Levinson, "The "Beat" Phenomenon in Wechsler Tests" J. of Clin. Psychol. 50 (Apr' 64) 118 - 120

ในปี ค.ศ. 1964 จัสแทค (Jastak)¹¹ ได้สำรวจพบข้อบกพร่องบางประการของ คำศัพท์ในแบบทดสอบ วิสค์ จึงได้ทำการปรับปรุง และเปลี่ยนแปลงวิธีการเสียใหม่ จุดอ่อนของคำศัพท์ใน วิสค์ มีดังนี้

1. คำศัพท์บางคำยังไม่เหมาะสม และเชื่อถือไม่ได้ นอกจากนั้นคำศัพท์บางคำ ยังมีความหมายใกล้เคียงกันด้วย
 2. วิธีการให้คะแนนยังไม่รัดกุมพอ
 3. การละเว้นคำศัพท์บางคำที่รู้สึกว่าเป็น "ง่าย" และให้คะแนนเต็มไปโดยไม่ได้ทำการสอบจริง ๆ นั้น อาจจะทำให้เกิดความคลาดเคลื่อนได้ง่าย เพราะบางคนอาจจะตอบได้เพียง 1 คะแนนเท่านั้น ถ้าตอบจริง ๆ
 4. กฎที่ห้าให้เลือกทดสอบ ถ้าผู้ถูกทดสอบทำไม่ได้คิด ๆ กัน 5 ข้อ ยังไม่ยุติธรรมกับผู้ถูกทดสอบ เพราะศัพท์บางคำที่ง่ายอาจจะยากสำหรับบางคน และศัพท์ยากบางคำอาจจะง่ายสำหรับบางคนก็ได้ ดังนั้นการจำกัดครั้งเดียว จึงเป็นการวิครอนโอกาสของผู้ถูกทดสอบในการแสดงความสามารถที่แท้จริง
- ดังนั้นการศึกษาของ จัสแทค จึงมุ่งที่จะ
1. เพื่อลดจำนวนข้อทดสอบลง และกำจัดคำที่ไม่เหมาะสมออกไป
 2. เพื่อทำให้การให้คะแนนมีความแน่นอนขึ้น และเพื่อสร้างเกณฑ์การให้คะแนนรัดกุมยิ่งขึ้น
 3. เพื่อกำจัดข้อตกลง (Assumption) ในเรื่องข้อง่าย ข้อยาก และการจำกัดจำนวนข้อที่ตอบผิดแล้ว เลิกทำ
 4. เพื่อหาคำใหม่ที่น่าเชื่อถือได้ เท่ากับของเดิม แต่มีคุณสมบัติดีในการทดสอบการให้คะแนน และผลทางสถิติ

¹¹S.E. Jastak and S.R. Jastak, "Short forms of the WAIS and WISC Vocabulary Subtests" J. of Clin. Psychol. vol 20 (Apr' 64) 167 - 181

เพื่อเป็นการแสดงตัวอย่างคำตอบที่สมบูรณ์สำหรับคะแนน 2, 1 และ 0

วิธีดำเนินการของจัสแทค ใช้กลุ่มตัวอย่างจากเด็กนักเรียน 1248 คน ตั้งแต่อายุ 5 ถึง 15 ปี ประมาณ 65 % ของกลุ่มตัวอย่าง เป็นเด็กที่มารับการตรวจทางคำสนธิสัญญาที่สถานีแมะแนวในวิลมิงตัน เดลาแวร์ (Wilmington, Delaware)

จัสแทค ได้ตั้งเกณฑ์ในการให้คะแนนไว้ว่า ¹²

คำตอบที่ได้คะแนน 2 จะต้องมีลักษณะดังนี้.-

1. เป็นไวพจน์ที่เหมาะสม โดยมีหรือไม่มีการแบ่งแยกที่ชัดเจนจากหุ้บไว
2. ภาวะจับประเภทเฉพาะของคำศัพท์ อาจมีหรือไม่มีการแบ่งขอบ
3. เกี่ยวกับส่วนที่เพิ่มคำพูด เช่น คำถามอธิบายในรูปของคำถามกริยาในรูปของกริยา คำศัพท์ในรูปของวลีศัพท์, กาชอกระทง นั้นไม่เป็นการหรือกริยา คำตอบที่ได้รับเป็นอย่างไรโดยบางหนึ่งถือเป็นคำตอบที่ถูกต้อง

คำตอบที่ได้คะแนน 1 จะต้องมีลักษณะดังนี้.-

1. เป็นไวพจน์ที่ค้อย หรือการแบ่งแยกประเภทที่ไม่แน่นอน
2. เป็นการอธิบายประโยค และหน้าที่ของคำศัพท์นั้น
3. อธิบายคำศัพท์นั้นโดยแยกออกเป็นส่วน ๆ การแดงและคุณลักษณะอื่น ๆ
4. การใช้ขอความซึ่งได้อธิบายแล้วแต่ส่วนหนึ่งของคำพูด
5. การแสดงตัวอย่างของวลี หรือประโยค โดยอาจจะใสหรือไมใส่คำศัพท์นั้นในประโยค แต่ใจความที่มีความหมายเกี่ยวข้องกับคำเดิม ซึ่งแสดงถึงความเข้าใจคำศัพท์ที่ถูกต้อง

นอกจากเกณฑ์ดังกล่าวข้างต้นนี้แล้ว จัสแทคยังอาศัยพหุนุกรมเป็นเครื่องมือในการตัดสินอีกด้วยการเลือกคำที่เหมาะสม จัสแทคใช้วิธีเลือกจากจำนวนคำตอบที่ถูกต้อง ปรากฏว่าจากแบบทดสอบสติปัญญาสำหรับผู้ใหญ่ของเวคสเลอร์ (WAIS) เลือกได้ 30 คำ และจากแบบทดสอบสติปัญญาสำหรับเด็กของเวคสเลอร์เลือกได้ 25 คำ ในจำนวน 40 คำของเวคสเลอร์พบว่า คำตัวกลาง เฉลยผิด และค่าเบี่ยงเบนมาตรฐานระหว่างข้อทดสอบใหม่กับข้อทดสอบแบบเก่ามีค่าใกล้เคียงกันมาก และเช่นเดียวกัน ค่าสัมประสิทธิ์ของสหสัมพันธ์ของแบบทดสอบทั้ง 2 ก็มีค่าสูง ($r = .895, .911$ และ $.929$ สำหรับเด็กชาย และ $.886, .890$ และ $.863$ สำหรับเด็กหญิง)

จัสแทคได้แสดงการเรียงลำดับคำศัพท์ของข้อทดสอบใหม่ตามความยากง่าย โดยเรียงจากจำนวนที่มีผู้ตอบถูกมากที่สุด ลงมาถึงข้อที่มีผู้ตอบถูกน้อยที่สุด เมื่อเปรียบเทียบกับแบบทดสอบเดิม ปรากฏว่า มีการสลับลำดับที่กันบ้างในบางข้อ

อายุของเด็กมีส่วนสัมพันธ์กับการให้คะแนน 2, 1 หรือ 0 ถ้าเด็กอายุ 5 - 7 ปี คำตอบที่ได้รับจะมีลักษณะเป็นการอธิบายธรรมดา แต่ถ้าเด็กโตที่อายุมากขึ้น ตั้งแต่ 8 - 15 ปี คำตอบจะเป็นการบรรยายที่แจ่มแจ้ง และแสดงถึงความเจริญทางสติภาวะมากขึ้น

ผลของการศึกษา จัสแทคพบว่า การที่ไม่ทดสอบคำถามที่ง่ายก็อยู่ถูกทดสอบนั้นทำให้ผล การทดสอบมีความเชื่อถือไม่ได้ประมาณ 12.5 % - 31 % ของจำนวนกลุ่มตัวอย่างที่สามารถทำ ข้อทดสอบข้อหนึ่ง ๆ ได้ หลังจากที่ทำไม่ได้คิด ๆ กันมาแล้ว 5 ข้อ และข้อทดสอบแบบใหม่ให้ เพิ่มการจำกัดโอกาสในการตอบเป็น 10 ข้อ ปรากฏว่าทำให้คะแนนดิบของนักวิจัยแต่ละคนเพิ่มขึ้น 1 - 6 คะแนน

ค่าสัมประสิทธิ์ความเชื่อถือได้ (Reliability Coefficient) ของข้อทดสอบ คำศัพท์ใหม่ที่สูงเท่า ๆ กับ ข้อทดสอบแบบเก่า บางระดับสูงกว่าเล็กน้อย แบบที่ปรับปรุงใหม่ มีค่าสัมประสิทธิ์ของความเชื่อถือได้ .77 .95

ความแตกต่างกันในเชิงคุณภาพของตัวจำกัดความขึ้นอยู่กับตัวการหลายอย่าง เช่น เรื่องเพศของผู้ถูกทดสอบ จากการศึกษาของจัสแทกพบว่า ผู้หญิงตอบคำศัพท์เกี่ยวกับเรื่องวรรณคดี ศิลปะ ศาสนา การแต่งกาย เครื่องประดับ และความทุกข์ได้ดี ส่วนผู้ชายตอบคำถามเกี่ยวกับพวก ความรุนแรง สงคราม อาวุธ เครื่องจักรกล ความตลกโปกส และการปฏิบัติได้ดีกว่า ความแตกต่างระหว่างเพศนี้เป็นเครื่องแสดงว่า ประเพณีและวัฒนธรรมเข้ามามีบทบาทอยู่ด้วย นอกเหนือจากพื้นฐานทางสรีระวิทยาแล้ว

ต่อมาได้มีผู้นำแบบทดสอบคำศัพท์ที่จัสแทกปรับปรุงใหม่ไปทดสอบกับคนไข้โรคจิตในโรงพยาบาลปรากฏว่า ผลที่ได้ใกล้เคียงกับการใช้แบบทดสอบเดิมของเวกสเลอร์

หลักฐานจากการวิจัยดังกล่าวมาเดิมนั้น ส่วนใหญ่เป็นการสร้างและพัฒนาแบบทดสอบสติปัญญาขึ้น โดยคำนึงถึงส่วนประกอบของมันสมองมนุษย์ วัฒนธรรม และสิ่งแวดล้อม ซึ่งนักจิตวิทยาและนักการศึกษาเชื่อว่า เป็นเรื่องบังเอิญ และมีอิทธิพลต่อระดับสติปัญญาของบุคคล นอกจากนั้น ยังแสดงให้เห็นว่า แบบทดสอบที่สร้างขึ้นมามีข้อบกพร่องบางประการที่ควรแก้ไข ปรับปรุงใหม่ เพื่อให้แบบทดสอบมีประสิทธิภาพยิ่งขึ้น

จากผลการวิจัยของนักวิจัยต่าง ๆ ดังกล่าวมาเดิมนั้นว่าเป็นประโยชน์อย่างยิ่ง ซึ่งผู้เขียนได้นำมาเป็นแนวทางในการศึกษาและวิจัย เพื่อสร้างแบบทดสอบภาคคำศัพท์เป็นภาษาไทยสำหรับใช้กับเด็กไทยโดยเฉพาะในการวินิจฉัย และหวังว่าจะมีนักวิจัยอื่น ๆ นำผลการวิจัยนี้ไปปรับปรุงแก้ไขให้ดียิ่งขึ้น เพื่อเป็นประโยชน์ต่อวงการศึกษาระดับของเยาวชนไทยในโอกาสต่อไป