

บทความวิชาการ (Academic Article)

การใช้แบบวัดบุคลิกภาพในกระบวนการคัดเลือกบุคคลในบริบททางการศึกษา USING PERSONALITY INVENTORIES IN THE EDUCATIONAL SELECTION

Received: June 19, 2020

Revised: August 4, 2020

Accepted: August 11, 2020

นุรซีตา เพอแผละ^{1*} ณัฐภรณ์ หลาวทอง² และสิวะโชติ ศรีสุทธิยากร³
Nurseeta Phoesalae^{1*} Nuttaporn Lawthong² and Siwachoat Srisuttiyakorn³

^{1,2,3}คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

^{1,2,3}Faculty of Education, Chulalongkorn University, Bangkok 10330, Thailand

*Corresponding Author, E-mail: Nurseeta_ph@hotmail.com

บทคัดย่อ

การวัดทางจิตวิทยาโดยเฉพาะด้านบุคลิกภาพเริ่มเข้ามามีบทบาทในกระบวนการคัดเลือกอย่างต่อเนื่องทั้งในภาคการศึกษาและอุตสาหกรรมองค์กร แต่เมื่อนำมาปฏิบัติจริงยังประสบปัญหาการตอบที่ตั้งใจบิดเบือนโดยเฉพาะการตอบตามความปรารถนาของสังคมเพื่อให้ตนเองผ่านการคัดเลือก จึงเริ่มมีการศึกษาและนำเสนอแนวทางในการป้องกันและแก้ไขปัญหาดังกล่าวอย่างหลากหลาย โดยในที่นี้การสร้างแบบวัดในรูปแบบบังคับเลือกเป็นหนึ่งในแนวทางที่มีผลการศึกษามากมาย ยืนยันถึงประสิทธิผลในการป้องกันการตอบแบบวัดอย่างบิดเบือน อย่างไรก็ตาม การใช้แบบวัดรูปแบบบังคับเลือกโดยใช้รูปแบบการคิดค่าคะแนนแบบดั้งเดิมยังก่อให้เกิดข้อจำกัดเกี่ยวกับปัญหาการไม่สามารถนำผลคะแนนไปใช้วัดความแตกต่างระหว่างบุคคลได้อย่างถูกต้อง ในระยะต่อมาจึงเริ่มมีการนำทฤษฎีการตอบสนองข้อสอบเข้ามาประยุกต์ใช้ในการคิดค่าคะแนนของแบบวัดรูปแบบบังคับเลือกเพื่อให้สามารถนำแบบวัดรูปแบบบังคับเลือกไปใช้ในกระบวนการคัดเลือกได้จริงในทางปฏิบัติ

คำสำคัญ: การวัดทางจิตวิทยา บุคลิกภาพ กระบวนการคัดเลือก

Abstract

Psychology testing, especially personality, has become an important role in selection practices in both education and industrial organization sectors. However, to implement psychology testing into the selection process has been challenged with social desirability response problems which respondents try to distort an answer in order to pass the selection. Consequently, there are currently many proposed solutions to this problem, in which one of those efficiency solutions is to construct measures with forced choice format. Nevertheless, forced choice format with traditional scoring has limitation in measuring individual difference that is a key point of the selection practices, which is called Ipsativity problem. Nowadays, Item Response Theory is applied to overcome this problem of forced choice format implementation. As a result of developed item response models, forced choice format with normative scoring approach can be practically implemented in selection context.

Keywords: Psychology Testing, Personality, Selection Process

บทนำ

การวัดทางจิตวิทยา (Psychological Testing) ถือเป็นหนึ่งในเครื่องมือสำคัญที่ถูกนำมาใช้ในกระบวนการคัดเลือกบุคคล โดยมีหลักฐานว่าถูกริเริ่มนำมาใช้ในกระบวนการคัดเลือกบุคคลเพื่อเข้าทำงานในกลุ่มชนชาติจีนกว่า 4000 ปีที่ผ่านมา (DuBois, 1970, 1972; Kaplan & Saccuzzo, 2009, p.11) ในระยะต่อมาจึงเริ่มแพร่หลายไปยังประเทศฝั่งตะวันตก องค์ความรู้ด้านการวัดทางจิตวิทยาจึงถูกศึกษาค้นคว้าและพัฒนาต่อยอดขึ้นอย่างรวดเร็ว โดยเฉพาะหลังจากการตีพิมพ์หนังสือ “The Origin of Species” ในปีคริสต์ศักราช 1859 ของ ชาร์ล ดาร์วิน ที่ได้นำเสนอทฤษฎีความแตกต่างระหว่างบุคคล (Individual Differences) แต่ในระยะแรกการนำองค์ความรู้ด้านการวัดทางจิตวิทยาไปใช้ในกระบวนการคัดเลือกบุคคลนั้น ยังคงถูกจำกัดอยู่ในคุณลักษณะที่เกี่ยวข้องกับความสามารถ (Ability Test) โดยเฉพาะการวัดเชาว์ปัญญาและการวัดผลสัมฤทธิ์ด้านต่างๆ ในขณะที่การวัดในกลุ่มจิตพิสัย (Non-Cognitive) ซึ่งหมายถึงรูปแบบของความรู้สึกนึกคิด คุณลักษณะทางจิต ความรู้สึก และบุคลิกภาพของบุคคล ยังไม่ถูกหยิบยกขึ้นมาให้ความสำคัญมากนัก

ต่อมาในช่วงสงครามโลกครั้งที่ 2 การวัดทางจิตวิทยาในกลุ่มจิตพิสัยเริ่มถูกหยิบยกขึ้นมาในกระบวนการคัดกรองทหารในกองทัพสหรัฐอเมริกา ในระยะนั้นแบบวัดบุคลิกภาพซึ่งเป็นแบบวัดจิตพิสัยที่เป็นทางการฉบับแรกจึงถูกพัฒนาขึ้นมา โดยใช้ชื่อว่า “Woodworth Personal Data Sheet” (Kaplan & Saccuzzo, 2009, p.17) หลังจากนั้นนักวิจัยเริ่มค้นพบว่าการวัดตัวแปรในกลุ่มจิตพิสัยโดยเฉพาะการด้านบุคลิกภาพ สามารถทำนายประสิทธิภาพด้านต่างๆของบุคคลได้เพิ่มมากขึ้น เมื่อเปรียบเทียบกับการใช้ตัวแปรที่เกี่ยวกับความสามารถทางสมองเป็นตัวทำนายเพียงอย่างเดียว (Stark et al., 2014, p. 154) โดยมีการศึกษาจำนวนมากที่ค้นพบว่าการนำคุณลักษณะด้านบุคลิกภาพไปใช้เป็นส่วนหนึ่งในกระบวนการคัดเลือกบุคคล จะส่งผลโดยตรงต่อระดับสมรรถนะของบุคคลที่ผ่านการคัดเลือก และสุดท้ายแล้วจะส่งผลต่อเนื่องไปยังระดับสมรรถนะขององค์กรโดยรวม (Touze, 2005, p. 38) โดยความสัมพันธ์ดังกล่าวถูกยืนยันผ่านการศึกษาและงานวิจัยต่างๆ จำนวนมากอย่างต่อเนื่อง โดยเฉพาะหลังจากการกำเนิดขึ้นของทฤษฎีบุคลิกภาพ 5 องค์ประกอบ (Salgado et al., 2014, p.797) ทำให้ปัจจุบันการวัดทางจิตพิสัยโดยเฉพาะด้านบุคลิกภาพเพื่อเป็นส่วนหนึ่งของเครื่องมือที่ใช้ในกระบวนการคัดเลือกจึงได้รับความสนใจเพิ่มขึ้นและขยายวงกว้างอย่างต่อเนื่องทั้งในด้านจิตวิทยา ด้านการศึกษา ตลอดจนด้านอุตสาหกรรมองค์กร

บทความฉบับนี้นำเสนอถึงความสำคัญของการวัดบุคลิกภาพเพื่อใช้ในกระบวนการคัดเลือก และประเด็นปัญหาในทางปฏิบัติ โดยเฉพาะปัญหาการตั้งใจบิดเบือนคำตอบของผู้ตอบแบบวัด นอกจากนี้ ยังรวบรวมข้อสรุปจากงานวิจัยส่วนหนึ่งเพื่อนำเสนอถึงทิศทางและพรหมแดนขององค์ความรู้เกี่ยวกับแนวทางการป้องกันและแก้ปัญหาการตั้งใจบิดเบือนคำตอบในปัจจุบัน ซึ่งมีการนำเสนอรายละเอียดของแนวทางที่ใช้แบบวัดรูปแบบบังคับเลือกในการป้องกันปัญหา โดยกล่าวถึงรูปแบบต่างๆ ของรูปแบบบังคับเลือก การคิดค่าคะแนนและปัญหาที่เกิดจากการคิดค่าคะแนนแบบดั้งเดิมต่อการนำไปใช้ในกระบวนการคัดเลือก โมเดลการตอบสนองข้อสอบสำหรับแบบวัดรูปแบบบังคับเลือก และบทสรุปขององค์ความรู้เกี่ยวกับการวัดบุคลิกภาพในกระบวนการคัดเลือกในปัจจุบันและทิศทางในอนาคต

การวัดบุคลิกภาพกับการตั้งใจบิดเบือนคำตอบ

ถึงแม้ว่าการวัดบุคลิกภาพจะสามารถทำได้ด้วยวิธีการที่หลากหลาย แต่โดยทั่วไปโดยเฉพาะในบริบทของกระบวนการคัดเลือคนั้น นิยมวัดโดยใช้ “แบบวัดบุคลิกภาพแบบประเมินตนเอง” ซึ่งสามารถบริหารจัดการการวัดได้อย่างสะดวก รวมทั้งสามารถวัดบุคลิกภาพได้หลายลักษณะและวัดคนจำนวนมากได้ในเวลาเดียวกัน ปัจจุบันจึงมีแบบวัดบุคลิกภาพแบบประเมินตนเองที่เป็นมาตรฐานถูกนำเสนอและนำไปใช้ในการดำเนินงานด้านต่างๆออกมามากมายอย่างต่อเนื่อง เช่น The Occupational Personality Questionnaire (2009), 16PF-Adolescent Personality Questionnaire (2008), Revised Neo Personality Inventory (2005), HEXACO Model of Personality Inventory (2000), Hogan Personality

Inventory (1980), Minnesota Multiphasic Personality Inventory (1939) เป็นต้น รวมถึงการวัดบุคลิกภาพแบบประเมินตนเอง ภายใต้ระบบ Tailored Adaptive Personality Assessment System (2005) ของกองทัพทหารสหรัฐอเมริกา ที่ถูกออกแบบขึ้นมาเพื่อใช้เป็นส่วนหนึ่งในกระบวนการคัดเลือกบุคคลเพื่อบรรจุแต่งตั้งเข้าปฏิบัติหน้าที่ในกองทัพทหารอเมริกาในปัจจุบัน ซึ่งทำให้เห็นว่าแบบวัดบุคลิกภาพถูกพัฒนาและนำไปใช้ประโยชน์ในด้านต่างๆ อย่างแพร่หลาย

อย่างไรก็ดี การใช้แบบวัดแบบประเมินตนเองสำหรับการวัดบุคลิกภาพซึ่งเป็นคุณลักษณะทางด้านจิตพิสัยนั้น มีความแตกต่างจากการวัดคุณลักษณะที่เกี่ยวข้องกับความสามารถ (Ability Test) โดยการวัดคุณลักษณะที่เกี่ยวข้องกับความสามารถนั้น หากวัดโดยใช้แบบสอบที่มีคุณภาพจะทำให้ผลของการวัดเป็นไปตามข้อเท็จจริงหรือความสามารถจริงของผู้ตอบอย่างตรงไปตรงมาและชัดเจนได้ง่ายกว่า ในขณะที่การวัดในกลุ่มจิตพิสัยยังอาจเผชิญความท้าทายจากการตั้งใจบิดเบือนคำตอบของผู้ตอบ โดยเฉพาะการตอบตามความปรารถนาของสังคม (Social Desirability) ซึ่งเป็นปัญหาหลักของการนำแบบวัดแบบประเมินตนเองไปใช้ในกระบวนการคัดเลือก และอาจทำให้การนำผลที่ได้จากการทำแบบวัดทางจิตวิทยาไปใช้มีความคลาดเคลื่อนอย่างรุนแรง (Kline, 2000, p. 254)

การตั้งใจบิดเบือนคำตอบของผู้ตอบแบบวัดแบบประเมินตนเองในกระบวนการคัดเลือก มักเกิดจากการตั้งใจตอบตามความปรารถนาของสังคม ซึ่งหมายความว่าผู้ตอบคำนึงถึงบรรทัดฐานของคนในสังคม แล้วแสวงหาคำตอบให้เป็นที่พึงปรารถนาและเป็นที่ยอมรับของสังคม (Jantawarn, 2013, p. 2) งานวิจัยจำนวนมากยืนยันว่าคะแนนการทำแบบวัดบุคลิกภาพในฐานะที่เป็นส่วนหนึ่งของกระบวนการคัดเลือกให้ผลคะแนนที่เพ้อฝันเกินจริงเมื่อเปรียบเทียบกับการทำแบบวัดในสถานการณ์ทั่วไป (Cao, 2016, p. 3) นอกจากนี้ การตั้งใจบิดเบือนคำตอบไปในทิศทางบวกของผู้ตอบแบบวัดยังทำให้ส่งผลอื่นๆต่อเนื่องตามมา โดยมีการศึกษาจำนวนมากได้ให้ข้อสรุปที่ตรงกันว่าคะแนนที่ได้จากการทำแบบวัดของกลุ่มผู้สมัครงาน ทำให้ความตรงเชิงโครงสร้างและความตรงเชิงเกณฑ์สัมพันธ์ของแบบวัดมีความคลาดเคลื่อน โดยมีองค์ประกอบที่เกี่ยวกับการเป็นพนักงานในอุดมคติ (Ideal Employee) เกิดขึ้นเป็นองค์ประกอบใหม่นอกเหนือจากองค์ประกอบที่ได้จากคะแนนของกลุ่มบุคคลทั่วไป (Cao, 2016, p. 4) ซึ่งจะส่งผลต่อเนื่องไปถึงการจัดลำดับคะแนนของผู้สมัครและการตัดสินใจในการคัดเลือก โดย Mueller-Hanson et al. (2003, pp. 351-352) ได้ทำการวิจัยเชิงทดลองโดยใช้สถานการณ์จริงในกระบวนการคัดเลือกบุคคลเข้าทำงานของหน่วยหนึ่งและทำการติดตามผล พบว่า เมื่อมีการใช้ผลจากการทำแบบวัดบุคลิกภาพเป็นเกณฑ์ในการคัดเลือก กลุ่มตัวอย่างที่ถูกเปิดโอกาสให้ทำการตอบแบบวัดบุคลิกภาพอย่างตั้งใจบิดเบือนคำตอบไปในทิศทางบวก (Faker) มีโอกาสในการถูกคัดเลือกเข้าทำงานมากกว่ากลุ่มควบคุม แต่เมื่อปฏิบัติงานไปในระยะเวลาหนึ่งพบว่ากลุ่มดังกล่าวมีระดับประสิทธิภาพในการปฏิบัติงานโดยเฉลี่ยต่ำกว่ากลุ่มควบคุม ซึ่งปัญหาดังกล่าวทำให้การใช้แบบวัดบุคลิกภาพในกระบวนการคัดเลือกไม่มีประสิทธิภาพได้อย่างที่ควรเป็น

รูปแบบบังคับเลือก: ทางออกของปัญหาการตั้งใจบิดเบือนคำตอบในแบบวัดบุคลิกภาพ

จากปัญหาการตั้งใจบิดเบือนคำตอบในการทำแบบวัดทางจิตในกระบวนการคัดเลือก ทำให้เกิดความพยายามที่จะศึกษาค้นคว้าถึงวิธีการจัดการกับปัญหาดังกล่าวอย่างต่อเนื่องตั้งแต่ในอดีตจนถึงปัจจุบัน โดยมีวิธีการและกระบวนการต่างๆ ถูกเสนอออกมามากมาย เช่น เทคนิคการสร้างข้อคำถามด้วยแบบต่างๆ การใช้ข้อความแจ้งเตือนระหว่างการทำแบบวัด ตลอดจนการแทรกข้อความหรือใช้เทคนิคทางสถิติต่างๆ เข้ามาตรวจจับและปรับคะแนนที่ได้จากการตั้งใจบิดเบือนคำตอบ อย่างไรก็ตามทิศทางการวิจัยในปัจจุบันให้น้ำหนักไปกับการสร้างแบบวัดในรูปแบบบังคับเลือก (Forced Choice) เพื่อป้องกันและขัดขวางการตั้งใจบิดเบือนคำตอบตั้งแต่ต้น โดยคาดว่าจะจะเป็นวิธีการที่ง่ายกว่าการจัดการกับคำตอบที่เกิดจากการตั้งใจบิดเบือนจากแบบวัดในรูปแบบอื่น (Stark et al., 2012, p. 465)

การสร้างแบบวัดแบบบังคับเลือกเป็นเทคนิคการสร้างแบบวัดที่นำเสนอตัวเลือกตั้งแต่ 2 ตัวเลือกขึ้นไปให้แก่ผู้ตอบ เพื่อให้ผู้ตอบเปรียบเทียบแล้วบังคับให้ผู้ตอบเลือกคำตอบตามวิธีการที่กำหนดไว้ให้เท่านั้น (Brown, 2015, p. 4) จากลักษณะของรูปแบบบังคับเลือกนั้นจึงถือว่าเป็นการสร้างแบบวัดเพื่อให้ได้ผลที่จะนำไปใช้ตัดสินใจเชิงเปรียบเทียบ (Comparative Judgments) ซึ่งแตกต่างจากการสร้างแบบวัดในรูปแบบมาตราประมาณค่า (Rating Scale) ที่มีการสร้างข้อความคำถามในรูปแบบข้อความเดียว (Single Statement) ผลที่ได้จากการทำแบบวัดจึงอยู่บนพื้นฐานของการตัดสินใจเชิงสัมบูรณ์ (Absolute Judgment) ที่ผู้ตอบสามารถตอบข้อความแต่ละข้อได้โดยไม่ต้องใช้การเปรียบเทียบกับข้อความอื่น

มีผลการศึกษาจำนวนมากได้ยืนยันถึงคุณภาพและข้อดีด้านต่างๆของการใช้รูปแบบบังคับเลือก ประการแรก คือ การที่รูปแบบบังคับเลือกตั้งอยู่บนพื้นฐานของการตัดสินใจเชิงเปรียบเทียบ (Comparative Judgments) จะช่วยลดปัญหาต่างๆของผลการตอบอันเกิดขึ้นได้บ่อยจากการใช้แบบวัดในรูปแบบข้อความเดียว เช่น ปัญหาการไม่ใส่ใจในการตอบจนนำไปสู่การตอบไปในทางเดียวกันทั้งหมดโดยไม่คำนึงถึงเนื้อหาของข้อความ (Acquiescence Bias) การตอบแบบสูงเกินไปหรือต่ำเกินไปโดยไม่มีเหตุผลอันควร (Severity Effect) การตอบแบบเน้นการตอบแบบกลางๆมากเกินไป (Central Tendency Responding) ซึ่งสุดท้ายแล้วการใช้รูปแบบบังคับเลือกจะช่วยลดปัญหาการเบี่ยงเบนคำตอบอย่างเป็นระบบดังที่กล่าวถึงข้างต้นได้ให้ดีขึ้น (Cheung & Chan, 2002; Brown, 2015, p.6) รวมทั้งลดปัญหาการตอบแบบเกิดปรากฏเหมารวม (Halo Effect) นอกจากนี้ จากการที่รูปแบบบังคับเลือกจัดให้มีการเปรียบเทียบแต่ละข้อความกันโดยตรงจะทำให้ผู้ตอบมีการตัดสินใจที่ง่ายกว่าการที่ต้องให้คะแนนในแต่ละข้อความ และประการสุดท้ายคือการช่วยป้องกันการเกิดปัญหาการตั้งใจบิดเบือนคำตอบ เช่น การตอบตามความปรารถนาของสังคม (Socially Desirable Responding) หรือการบิดเบือนคำตอบเพื่อจุดประสงค์บางประการ (Faking) เป็นต้น (Brown, 2015, p.7) ด้วยข้อดีของการสร้างแบบวัดในรูปแบบบังคับเลือกที่ช่วยแก้ปัญหาและจุดอ่อนของการสร้างแบบวัดในรูปแบบข้อความเดียว จึงทำให้ปัจจุบันองค์ความรู้เกี่ยวกับรูปแบบบังคับเลือกถูกศึกษา นำเสนอและขยายขอบเขตขององค์ความรู้ออกไปอย่างต่อเนื่อง

ภายใต้คำว่า “รูปแบบบังคับเลือก” นั้น มีรูปแบบย่อยที่แตกต่างกันออกไปหลากหลายรูปแบบซึ่งสามารถแบ่งได้ตามเงื่อนไขต่างๆ เช่น หากพิจารณาจากคุณลักษณะทางจิตมิติของแต่ละข้อความภายในข้อเดียวกัน รูปแบบย่อยของ รูปแบบบังคับเลือกสามารถแบ่งออกเป็น แบบเอกมิติ (Unidimensional Forced Choice) ที่แต่ละข้อความที่เสนอให้เลือกตอบภายในข้อเดียวกันมีการวัดคุณลักษณะทางจิตในองค์ประกอบเดียวกัน และแบบพหุมิติ (Multidimensional Forced Choice) ที่ข้อความภายในแต่ละข้อ มีการวัดคุณลักษณะทางจิตในองค์ประกอบที่แตกต่างกัน และหากพิจารณาด้านวิธีการตอบที่กำหนดให้ผู้ตอบ สามารถแบ่งได้เป็น 3 รูปแบบย่อย ได้แก่ รูปแบบ “PICK” ซึ่งกำหนดให้ผู้ตอบเลือกข้อความ ที่ตรงกับตัวเองมากที่สุดขึ้นมา 1 ข้อความ ในขณะที่รูปแบบ “MOLE” มีการกำหนดให้ผู้ตอบทำการเลือกข้อความขึ้นมาจำนวน 2 ข้อความที่ตรงกับตัวเองมากที่สุดและน้อยที่สุด และรูปแบบ “RANK” ซึ่งมีการกำหนดให้ผู้ตอบทำการเรียงลำดับข้อความที่ตรงกับตัวเองจากมากไปหาน้อย เป็นต้น

ตาราง 1 ตัวอย่างการสร้างข้อคำถามรูปแบบบังคับเลือกในรูปแบบต่างๆ

รูปแบบย่อย	ตัวอย่างการพัฒนาข้อคำถาม	
PICK	จงเลือกข้อความที่อธิบายได้ตรงกับตัวคุณมากที่สุด	
	☐ A: ฉันมีความสุขเมื่ออยู่ท่ามกลางผู้คน	
	☐ B: ห้องนอนของฉันเป็นระเบียบเรียบร้อย	
MOLE	เลือกข้อที่ตรงกับตัวคุณมากที่สุด (Most) 1 ข้อความ และน้อยที่สุด (Least) 1 ข้อความ	
	ตรงมากที่สุด	ตรงน้อยที่สุด
	A: ฉันเป็นผู้นำกลุ่ม ☐	☐
	B: ฉันเคารพในความคิดผู้อื่น ☐	☐
	C: ฉันเข้าใจสิ่งที่เพื่อนร่วมงานได้สื่อสาร ☐	☐
	D: ฉันมีความคิดสร้างสรรค์อยู่เสมอ ☐	☐
RANK	เรียงลำดับข้อความที่ “ฉันชอบ” จากมากไปหาน้อย	
		ลำดับ
	A: การเจรจากับผู้คน ☐	
	B: สร้างสรรค์งานศิลป์ ☐	
	C: ซ่อมอุปกรณ์หรือสิ่งของเครื่องใช้ ☐	
	D: อ่านและทำความเข้าใจทฤษฎีที่ซับซ้อน ☐	

จากตาราง 1 จะเห็นได้ว่ารูปแบบบังคับเลือกมีรูปแบบย่อยที่แตกต่างกันออกไป อย่างไรก็ตาม ในงานวิจัยเชิงสังเคราะห์ของ Cao (2016, p. 69) ที่ได้รวบรวมและสังเคราะห์งานวิจัยที่เกี่ยวกับรูปแบบของแบบวัดที่สามารถป้องกันการตั้งใจบิดเบือนคำตอบได้ ผลการศึกษาพบว่าการใช้แบบวัดในรูปแบบบังคับเลือกป้องกันการตั้งใจบิดเบือนคำตอบได้จริง โดยเฉพาะรูปแบบที่เป็นแบบบังคับเลือกแบบพหุมิติ (Multidimensional Forced Choice) ในขณะที่การอภิปรายผลการศึกษาของ Hontangas et al. (2015, p. 609) ที่ศึกษาพบว่า การใช้แบบวัดรูปแบบบังคับเลือกในรูปแบบ PICK, MOLE และ RANK สามารถป้องกันการตั้งใจบิดเบือนคำตอบไปในทิศทางบวกได้จริง แต่ในทางปฏิบัติ การใช้รูปแบบ PICK ที่นำเสนอข้อความให้ผู้ตอบเพียง 2 ข้อความเพื่อให้ผู้ตอบเลือกข้อความใดข้อความหนึ่งอาจให้ผลที่ดีที่สุด เนื่องจากมีแหล่งของความคลาดเคลื่อนจากจุดต่างๆ น้อยที่สุด จึงอาจกล่าวได้ว่าผลการศึกษาในปัจจุบันให้น้ำหนักต่อข้อสรุปไปในทิศทางที่ว่า รูปแบบบังคับเลือกที่มีประสิทธิภาพในการป้องกันการตั้งใจบิดเบือนคำตอบได้ดีที่สุด คือ รูปแบบที่นำคุณสมบัติเกี่ยวกับความเป็นพหุมิติของข้อความภายในข้อเดียวกัน (Multidimensional Forced Choice) ร่วมกับการกำหนดรูปแบบการตอบของผู้ตอบให้เป็นแบบเลือกข้อความใดข้อความหนึ่งขึ้นมา (PICK) จากสองข้อความที่มีระดับความปรารถนาของสังคมเท่ากัน โดยปัจจุบันมีการกำหนดชื่อให้แก่รูปแบบย่อยที่มีคุณสมบัติทั้ง 2 คุณสมบัติดังกล่าวไว้ว่า “แบบวัดพหุมิติแบบคู่เทียบ” หรือ Multidimensional Pairwise Preference (MDPP) (Stark et al., 2005, pp. 1-21) หรือ Multidimensional Pairwise Comparison (MPC) (Wang et al. 2017, pp. 600-613)

ตัวอย่างของการสร้างข้อคำถามในแบบวัดพหุมิติแบบคู่เทียบ

----- A. กระเป๋าของฉันเป็นระเบียบเรียบร้อย

----- B. ฉันเชื่อมั่นในความสัมพันธ์

จาก 2 ข้อความที่ปรากฏในข้อหนึ่งของแบบวัดพหุมิติแบบคู่เทียบ จะเห็นว่าทั้งสองข้อความมาจากองค์ประกอบที่แตกต่างกัน แต่มีทิศทางและความน่าปรารถนาของสังคมไปในแนวทางเดียวกัน ซึ่งข้อความ A เป็นข้อความที่วัดองค์ประกอบ

ด้านความเป็นระเบียบเรียบร้อย (Organization) ในขณะที่ข้อความ B วัดองค์ประกอบด้านความไว้วางใจผู้อื่น (Interpersonal Trust) แล้วจึงกำหนดให้ผู้ตอบเลือกข้อความใดข้อความหนึ่ง ที่ผู้ตอบเห็นว่าตรงหรือใกล้เคียงกับผู้ตอบมากกว่าอีกข้อความหนึ่ง เมื่อการตอบเสร็จสิ้นแล้วจึงทำการประมวลผลและคิดค่าคะแนนแบบจำแนกตามแต่ละองค์ประกอบ ทุกองค์ประกอบที่ปรากฏในแบบวัด

Ipsativity: ข้อจำกัดในการเปรียบเทียบระหว่างบุคคลของแบบวัดรูปแบบบังคับเลือก

ถึงแม้ว่าแนวโน้มการศึกษาวิจัยในปัจจุบันพบว่าการสร้างแบบวัดในรูปแบบบังคับเลือกจะสามารถลดปัญหาต่างๆ อันเกิดจากการใช้แบบวัดในรูปแบบข้อความเดียว อย่างไรก็ตาม การใช้รูปแบบบังคับเลือกในทุกรูปแบบยังคงมีข้อจำกัดต่อการนำไปใช้ เมื่อข้อมูลที่ได้ถูกนำไปคิดค่าคะแนนแบบดั้งเดิม (Traditional Scoring Approach) กล่าวคือ การให้คะแนนเป็น “1” สำหรับองค์ประกอบที่ถูกเลือก และให้คะแนนเป็น “0” สำหรับองค์ประกอบที่ไม่ถูกเลือก แล้วจึงรวมผลคะแนนของแต่ละองค์ประกอบ แต่การรวมคะแนนในลักษณะนี้จะทำให้คะแนนรวมของผู้ทำแบบวัดเป็นค่าคงที่ที่เท่ากันทุกคน จึงทำให้เกิดปัญหาที่เรียกว่า “Ipsativity” ตามเกณฑ์ของ Clemans (as cited in Salgado et al., 2015, p. 80) ที่ระบุไว้ว่า Ipsativity เป็นปัญหาที่เกิดจากการใช้รูปแบบการคิดค่าคะแนนที่ผลรวมของคะแนนทั้งหมดเป็นค่าคงที่ที่เท่ากันทุกคน และจะนำไปสู่ข้อจำกัดในการนำไปใช้วัดความแตกต่างระหว่างบุคคล

ตาราง 2 ตัวอย่างผลการตอบจากแบบวัดรูปแบบบังคับเลือกและคิดค่าคะแนนแบบดั้งเดิม

บุคคล	องค์ประกอบ		รวม
	ความสามารถในการเข้าสังคม	การมีความคิดสร้างสรรค์	
นาย A	6	4	10
นาย B	8	2	10
นาย C	9	1	10

จากตาราง 2 แสดงตัวอย่างข้อมูลที่ได้จากตอบแบบวัดรูปแบบหนึ่งของแบบบังคับเลือก จำนวน 10 ข้อ ซึ่งต้องการวัดคุณลักษณะทางจิตของผู้ตอบแบบวัด 2 องค์ประกอบ ได้แก่ ความสามารถในการเข้าสังคมและการมีความคิดสร้างสรรค์ของผู้ตอบแบบวัดจำนวน 3 คน เมื่อทำการคิดค่าคะแนนแบบดั้งเดิม (Traditional Scoring Approach) จะพบว่าผลรวมคะแนนมีค่าเท่ากับ 10 คะแนนเท่ากันทุกคน จึงไม่สามารถเปรียบเทียบความแตกต่างระหว่างบุคคลโดยใช้คะแนนรวมได้ ในขณะเดียวกันหากทำการเปรียบเทียบและสรุปความแตกต่างระหว่างบุคคลโดยใช้ข้อมูลเพียงบางองค์ประกอบอาจให้ผลสรุปที่ผิดพลาดอย่างรุนแรงได้ เช่น การสรุปว่า “นาย A มีความสามารถในการเข้าสังคมมากกว่านาย C” อาจเป็นการสรุปที่ยังไม่สมเหตุสมผลนัก เนื่องจากการที่นาย A มีคะแนนด้านความสามารถในการเข้าสังคมเท่ากับ 6 คะแนน และมีคะแนนด้านการมีความคิดสร้างสรรค์เท่ากับ 4 คะแนน หากพิจารณาเพียงสององค์ประกอบดังกล่าว ผลคะแนนที่เกิดขึ้นอาจเกิดจากการที่นาย A มีความสามารถทั้งด้านการเข้าสังคมและการมีความคิดสร้างสรรค์ แต่ความสามารถด้านการเข้าสังคมสูงกว่าเล็กน้อย ในขณะที่นาย C ซึ่งมีคะแนนด้านความสามารถในการเข้าสังคมเท่ากับ 9 คะแนน และมีคะแนนด้านการมีความคิดสร้างสรรค์เท่ากับ 1 คะแนนนั้น หากพิจารณาเพียงแค่สององค์ประกอบดังกล่าว ผลคะแนนที่เกิดขึ้นอาจเกิดจากการที่นาย C ไม่มีความสามารถทั้งด้านการเข้าสังคมและการมีความคิดสร้างสรรค์เลย แต่เมื่อเปรียบเทียบเพียงสององค์ประกอบนี้แล้วความสามารถด้านการเข้าสังคมของนาย C อาจจะสูงกว่าความสามารถด้านการมีความคิดสร้างสรรค์เป็นอย่างมาก

อย่างไรก็ดี การคิดคะแนนในลักษณะนี้ ผลคะแนนที่ได้สามารถนำไปใช้ประโยชน์ได้ดีในสถานการณ์ที่ต้องการวัดความแตกต่างของแต่ละองค์ประกอบภายในบุคคลเดียวกัน (Intra-Individual Comparison) เช่น งานด้านการให้คำปรึกษา

และการแนะนำ แต่จะก่อให้เกิดข้อจำกัดต่อการนำมาใช้งานจริงโดยเฉพาะการนำไปใช้ในกระบวนการคัดเลือกซึ่งจำเป็นต้องให้ได้ผลการวัดที่สามารถแสดงให้เห็นความแตกต่างระหว่างบุคคลได้ (Inter-Individual Comparison) ด้วยเหตุนี้ถึงแม้ว่าแนวโน้มของข้อสรุปงานวิจัยในปัจจุบันพบว่าการสร้างแบบวัดในรูปแบบบังคับเลือกจะสามารถช่วยป้องกันปัญหาการตั้งใจบิดเบือนคำตอบซึ่งเป็นปัญหาสำคัญของการนำแบบวัดทางจิตพิสัยไปใช้ในกระบวนการคัดเลือก แต่ในขณะเดียวกันหากใช้รูปแบบการคิดค่าคะแนนแบบดั้งเดิม ผลคะแนนที่ได้ยังคงไม่สามารถนำไปใช้ในการหาความแตกต่างระหว่างบุคคลอันเป็นวัตถุประสงค์สำคัญของกระบวนการคัดเลือกได้

ทฤษฎีการตอบสนองข้อสอบ : ทางออกของปัญหา Ipsativity

ด้วยจุดเด่นของรูปแบบบังคับเลือกที่เหมาะสมต่อการนำไปใช้ในกระบวนการคัดเลือก ในระยะต่อมาจึงเริ่มมีการพัฒนากระบวนการคิดค่าคะแนนของรูปแบบบังคับเลือกขึ้นมาแทนการคิดค่าคะแนนแบบดั้งเดิมที่ก่อให้เกิดปัญหา Ipsativity อย่างต่อเนื่อง จนกระทั่งปัจจุบันจึงเริ่มมีการนำทฤษฎีความน่าจะเป็นและทฤษฎีการตอบสนองข้อสอบ (Item Response Theory) เข้ามาประยุกต์ใช้เพื่อให้สามารถประมวลผลข้อมูลที่ได้จากแบบวัดออกมาเป็นค่าตัวเลขที่ใช้เปรียบเทียบความแตกต่างระหว่างบุคคลได้และแก้ปัญหา Ipsativity ได้อย่างสมบูรณ์ ซึ่งสามารถเรียกการคิดค่าคะแนนในลักษณะนี้ว่า Model-based normative scoring โดยผลลัพธ์สุดท้ายที่ได้จากการประมวลผลภายใต้ทฤษฎีการตอบสนองข้อสอบจะแสดงถึงระดับคุณลักษณะทางจิตในแต่ละองค์ประกอบของผู้ตอบแต่ละบุคคล ซึ่งมีค่าเป็นตัวเลขที่มีการแจกแจงแบบปกติมาตรฐานที่มีค่าอยู่ในช่วง -3.00 ถึง 3.00 โดยค่าดังกล่าวเป็นค่าที่ไม่เกิดปัญหา Ipsativity จึงสามารถนำไปใช้เปรียบเทียบความแตกต่างระหว่างบุคคลได้อย่างมีประสิทธิภาพ

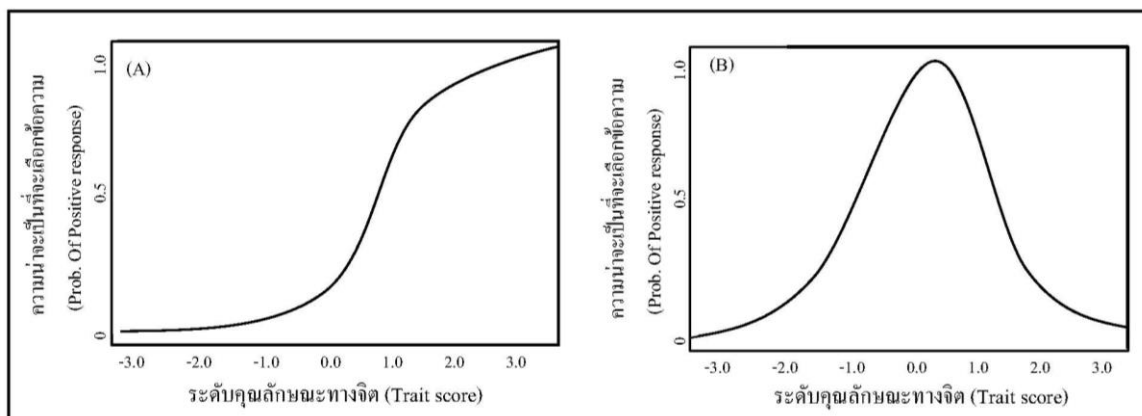
ปัจจุบันมีโมเดลการตอบสนองข้อสอบที่ถูกพัฒนาขึ้นเพื่อคิดค่าคะแนนของผู้ตอบแบบวัดรูปแบบบังคับเลือก โดยเฉพาะจำนวน 5 โมเดล ดังตาราง 3 โดยแต่ละโมเดลมีข้อกำหนดและแนวความคิดพื้นฐานบางประการที่แตกต่างกัน ได้แก่ 1) จำนวนของข้อความที่นำเสนอให้แก่ผู้ตอบในแต่ละข้อ (Block Size) 2) การกำหนดองค์ประกอบภายในข้อเดียวกัน (Dimensionality of comparisons) 3) รูปแบบการตอบ และ 4) ประเภทของโมเดลการวัด

ตาราง 3 แสดงโมเดลการตอบสนองข้อสอบที่ใช้กับแบบวัดรูปแบบบังคับเลือก

ชื่อโมเดลการตอบสนองที่ใช้ในการวิเคราะห์	จำนวนข้อความ	องค์ประกอบ	รูปแบบการตอบ	ประเภทของโมเดลการวัด
Zinnes-Griggs model for unidimensional pairwise preferences (Zinnes & Griggs, 1974)	2	เอกมิติ	PICK	Ideal Point
McCloy-Heggestad-Reeve unfolding model for multidimensional ranking blocks (McCloy, Heggestad, & Reeve, 2005)	≥ 2	พหุมิติ	RANK	Ideal Point
Multi-unidimensional Pairwise Preference (MUPP) IRT Model (Stark, Chernyshenko & Drasgow, 2005)	2	พหุมิติ	PICK	Ideal Point
Thurstonian IRT model (Brown & Olivares, 2011)	≥ 2	เอกมิติ พหุมิติ	RANK	Dominance
Rasch ipsative model (RIM) (2017)	2	พหุมิติ	PICK	Dominance

ถึงแม้ว่าผลลัพธ์สุดท้ายที่ได้จากการคิดค่าคะแนนโดยใช้โมเดลการตอบสนองข้อสอบทุกโมเดลจะให้ผลไปในแนวทางเดียวกัน แต่การมีสมมติฐานเกี่ยวกับกระบวนการตอบของผู้ตอบ (Response Process) ที่แตกต่างกันนั้นจะนำมาสู่การมีรูปแบบของโมเดลการวัด (Measurement Model) ที่แตกต่างกัน ซึ่งหากจำแนกโมเดลในตารางที่ 3 ออกตามประเภทของโมเดลการวัด จะสามารถจำแนกได้เป็น 2 กลุ่ม ได้แก่โมเดลที่อยู่ในกลุ่ม Dominance IRT Model และ โมเดลที่อยู่ในกลุ่ม Ideal Point IRT Model (หรือเรียกอีกชื่อหนึ่งได้ว่า Unfolding Model) โดยโมเดลที่อยู่ในกลุ่ม Dominance IRT Model มีแนวความคิดพื้นฐานที่ว่า ในการวัดคุณลักษณะทางจิตในองค์ประกอบใดๆ หากระดับของคุณลักษณะทางจิตของผู้ถูกวัดอยู่ในระดับที่สูง คะแนนที่ได้จากการใช้โมเดลการวัดจะสูงตามไปด้วยเสมอ และในทางกลับกันหากระดับคุณลักษณะทางจิตของผู้ถูกวัดอยู่ในระดับที่ต่ำ คะแนนที่ได้จะต่ำตามไปด้วยเสมอ ด้วยเหตุนี้โมเดลการตอบสนองข้อสอบในกลุ่มนี้จึงเป็นโมเดลที่มีฟังก์ชันความน่าจะเป็นในการประมาณค่าคุณลักษณะทางจิตของผู้ถูกวัดเป็นแบบเพิ่มขึ้นทางเดียว (Monotonic)

ในขณะที่โมเดลที่อยู่ในกลุ่ม Ideal Point IRT Model มีแนวความคิดพื้นฐานที่ว่า ค่าของระดับคุณลักษณะทางจิตของผู้ถูกวัดและค่าของระดับคุณลักษณะของข้อความที่ใช้วัด เป็นค่าที่สามารถทำให้อยู่บนเส้นจำนวนต่อเนื่อง (Continuum) เดียวกันได้ จึงสามารถนำมาเปรียบเทียบกันได้ โดยคะแนนที่ได้จากการใช้โมเดลจะมีค่าสูงที่สุด เมื่อใช้กับผู้ถูกวัดที่มีคุณลักษณะทางจิตที่เท่ากับหรือใกล้เคียงกับระดับคุณลักษณะของข้อความนั้นๆ และคะแนนที่ได้จะน้อยลงเรื่อยๆ เมื่อระดับคุณลักษณะทางจิตของผู้ถูกวัดยิ่งห่างหรือแตกต่างจากระดับคุณลักษณะของข้อความ ซึ่งความแตกต่างนี้อาจเกิดจากการที่ระดับคุณลักษณะทางจิตของผู้ถูกวัดต่ำกว่าระดับคุณลักษณะของข้อความ (Disagree from below) หรืออาจเกิดจากการที่ระดับคุณลักษณะทางจิตของผู้ถูกวัดสูงกว่าระดับคุณลักษณะของข้อความ (Disagree from above) ด้วยเหตุนี้โมเดลการตอบสนองข้อสอบในกลุ่ม Ideal Point IRT Model จึงเป็นโมเดลที่มีฟังก์ชันความน่าจะเป็นที่ใช้ในการประมาณค่าคุณลักษณะทางจิตของผู้ถูกวัดเป็นแบบไม่เพิ่มขึ้นทางเดียว (Nonmonotonic)



ภาพ 1 ตัวอย่าง Item characteristic Curve (ICC) ของโมเดลในกลุ่ม Dominance IRT Model (A) และ Ideal point IRT Model (B)

ถึงแม้ว่าแต่ละโมเดลที่ถูกนำเสนอเพื่อใช้ในการคิดค่าคะแนนของแบบวัดรูปแบบบังคับเลือกจะมีข้อกำหนดและแนวความคิดเบื้องหลังที่ต่างกันออกไป แต่จุดเด่นของทุกโมเดลภายใต้ Model based normative scoring คือ การพยายามแก้ไขปัญหาค Ipsativity ให้ได้อย่างสมบูรณ์ จนกระทั่งสามารถเอาชนะข้อจำกัดของนำแบบวัดรูปแบบบังคับเลือกไปใช้วัดความแตกต่างระหว่างบุคคลได้ อีกทั้งในการคิดค่าคะแนนจากการประมาณโดยใช้โมเดลทางคณิตศาสตร์เป็นพื้นฐานนั้นทำให้ได้สารสนเทศทางสถิติมากมาย อันนำไปสู่การประยุกต์ใช้กับเทคโนโลยีการวัดตลอดจนเทคนิคการวิเคราะห์ต่างๆ มากมาย เช่น Differential item functioning analysis (DIF), Computerized adaptive testing (CAT) ได้ต่อไป

อย่างไรก็ตาม ในการนำโมเดลการตอบสนองโมเดลใดมาใช้กับแบบวัดรูปแบบบังคับเลือก ควรพิจารณาถึงข้อกำหนดและแนวความคิดพื้นฐานของโมเดลให้สอดคล้องกับรูปแบบของแบบวัดที่พัฒนาขึ้นรวมถึงคุณลักษณะที่ต้องการวัด จึงจะทำให้การนำแบบวัดรูปแบบบังคับเลือกไปใช้เป็นไปได้อย่างมีประสิทธิภาพ

บทสรุป

จากองค์ความรู้และข้อค้นพบจากงานวิจัยจำนวนมากยืนยันว่าคุณลักษณะทางจิตพิสัยของบุคคลมีความสัมพันธ์โดยตรงกับผลสัมฤทธิ์ทั้งทางการศึกษาและการทำงาน โดยเฉพาะการวัดด้านบุคลิกภาพภายใต้ทฤษฎีบุคลิกภาพห้าองค์ประกอบ จึงเริ่มมีการนำการวัดทางจิตพิสัยเข้าไปใช้ในกระบวนการคัดเลือกควบคู่กับการวัดความสามารถ แต่ยังมีประสบการณ์การตั้งใจบิดเบือนคำตอบของผู้ทำแบบวัดโดยมีวัตถุประสงค์เพื่อให้ตนเองได้รับการคัดเลือก จึงเริ่มมีการคิดค้นและนำแบบวัดรูปแบบบังคับเลือกเข้ามาใช้เพื่อแก้ปัญหาดังกล่าว อย่างไรก็ตามการใช้แบบวัดรูปแบบบังคับเลือกกับการคิดค่าคะแนนแบบดั้งเดิม ทำให้ผลคะแนนที่ได้เกิดปัญหา Ipsativity ที่ผลคะแนนเหมาะสมกับการใช้วัดความแตกต่างของแต่ละองค์ประกอบภายในบุคคลเดียวกันเท่านั้น แต่ไม่สามารถนำไปใช้วัดความแตกต่างระหว่างบุคคลได้อย่างถูกต้อง ในระยะต่อมาจึงเริ่มมีการนำทฤษฎีการตอบสนองข้อสอบเข้ามาประยุกต์ใช้กับการคิดค่าคะแนนของแบบวัดรูปแบบบังคับเลือก นำไปสู่การคิดค่าคะแนนในรูปแบบใหม่ที่เป็นแบบ Model-based normative scoring เพื่อแก้ปัญหา โดยปัจจุบันมีโมเดลที่แนะนำให้เสนอเข้ามาเพื่อเป็นตัวเลือกในการใช้งานอยู่จำนวนหนึ่ง

อย่างไรก็ดี องค์ความรู้เกี่ยวกับการใช้แบบวัดทางจิตวิทยาในกระบวนการคัดเลือกยังคงต้องการการศึกษาเพิ่มเติมอีกในหลายประเด็นเพื่อให้สามารถนำไปใช้ได้อย่างมีคุณภาพ เช่น การออกแบบและพัฒนา การทดสอบคุณภาพ การศึกษาเปรียบเทียบคุณภาพของรูปแบบย่อยภายใต้รูปแบบบังคับเลือก รูปแบบการคิดค่าคะแนน และโมเดลการตอบสนองข้อสอบสำหรับแบบวัดรูปแบบบังคับเลือกรูปแบบย่อยต่างๆ ตลอดจนการศึกษาเพื่อกำหนดสถานการณ์ที่เหมาะสมต่อการนำแบบวัดรูปแบบบังคับเลือกรูปแบบย่อยต่างๆไปใช้ เพื่อให้สามารถนำการวัดทางจิตพิสัยไปใช้ในกระบวนการคัดเลือกได้อย่างมีคุณภาพ

References

- Brown, A., & Olivares, A. M. (2011). Item Response Modeling of forced-choice questionnaires, *Educational and Psychological Measurement*, 71(3), 460-502.
- Brown, A. (2015). Personality assessment forced choice. In Wright, J.D., ed. *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier.
- Cao, M. (2016). *Examining the fixability of forced-choice individual differences measures* (Doctoral dissertation). Illinois: University of Illinois.
- Clemans, W. V. (1966). *An analytical and empirical examination of some properties of ipsative measures*. Richmond, VA: Psychometric Society.
- Clemans, W. V. (1967). *An analytical and empirical examination of some properties of ipsative measures*. Richmond, VA: Psychometric Society.
- Jantawarn, S. (2013). *Development of a socially desirable responding scale of Thai undergraduate students* (Doctoral dissertation). Bangkok: Chulalongkorn University. [in Thai]
- Kaplan, R. M., & Saccuzzo, D. P. (2009) *Psychological testing: Principles, applications, and issues*. Belmont, CA: Wadsworth Cengage Learning,

- Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). London: United Kingdom Routledge.
- Lee, P., Lee, S., & Stark, S. (2018). Examining validity evidence for multidimensional forced choice measures with different scoring approaches. *Personality and Individual Differences*, 123, 229-235. 10.1016/j.paid.2017.11.031.
- Mueller-Hanson, R., Heggstad, E. D., & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88(2), 348-355.
- Hontangas, P. D., Torre, J. D., Ponsoda, V., Leenen, I., Morillo, D., & Abad, F. J. (2015). Comparing traditional and IRT scoring of forced-choice tests. *Applied Psychological Measurement*, 39(8), 598-612.
- Salgado, J. F., Anderson, N., & Tauriz, G. (2015). The validity of ipsative and quasi-ipsative forced-choice personality inventories for different occupational groups: A comprehensive meta-analysis. *Journal of Occupational and Organizational Psychology*, 88(4), 797-834.
- Stark, S., Chernyshenko, O. S., Drasgow, F., & White, A. (2012). Adaptive Testing with Multidimensional Pairwise Preference Items: Improving the efficiency of personality and other noncognitive assessment. *Organizational Research Methods*, 15(3), 463-487.
- Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, D., White, A., Heffner, T., & Farmer, L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology*, 26(3), 153-164.
- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2005). An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: An application to the problem of faking in personality assessment. *Applied Psychological Measurement*, 29, 184- 201.
- Stark, S., Chernyshenko, O. S., & Drasgow, F., & Williams, B. A. (2006). Item responding in personality assessment: Should ideal point methods be considered for scale development and scoring? *Journal of Applied Psychology*, 91, 25-3.
- Touze, P. (2005). *Personality and prediction of performance in the workplace*. *Personality and performance*, 68(1), 37-53.
- Wang, W. C., Qiu, X. L., Chen, C. W., Ro, S., & Jin, K. Y. (2017). Item Response Theory Models for Ipsative Tests with Multidimensional Pairwise. Comparison Items. *Applied Psychological Measurement*, 41(8), 600-613.
- Zhang, B. (2018). *Though forced, still valid: Psychometric equivalence of forced-choice and single-statement measures* (Doctoral dissertation). Illinois: University of Illinois.