

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบคุณภาพของวิธีการวัดคะแนนพัฒนาการตามแนวคิดแบบดั้งเดิม 3 วิธี ได้แก่ วิธีการวัดคะแนนพัฒนาการจากความแตกต่างระหว่างคะแนนดิบ (DS) วิธีการวัดคะแนนพัฒนาการสัมพัทธ์ (RG) และวิธีการวัดคะแนนพัฒนาการสัมพัทธ์สมดุล (BRG) โดยใช้โมเดลโค้งพัฒนาการที่มีตัวแปรแฝงเป็นเกณฑ์ เมื่อข้อมูลมีเงื่อนไขการเปลี่ยนแปลงเป็นแบบเส้นตรงและแบบไม่เป็นเส้นตรง และจำนวนครั้งของการวัดซ้ำแตกต่างกัน ซึ่งข้อมูลที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลที่ได้มาจากการจำลอง (simulate) ข้อมูลด้วยเทคนิคมอนติคาร์โลโดยใช้โปรแกรม MPLUS version 4.1 ตามเงื่อนไข 4 ประการ ได้แก่ 1) รูปแบบของการเปลี่ยนแปลงของข้อมูล 2) จำนวนครั้งของการวัดซ้ำ 3) ค่าเฉลี่ยตัวแปรแฝงระดับคะแนนในการวัดครั้งแรก (ML) และ 4) ค่าอัตราพัฒนาการ (MS) รวมทั้งสิ้น 108 สถานการณ์ และในแต่ละสถานการณ์มีกลุ่มตัวอย่างเป็น 500 ชุดข้อมูล การวิเคราะห์ข้อมูลใช้การประมาณค่าคะแนนพัฒนาการ และวิเคราะห์เปรียบเทียบคุณภาพของวิธีการวัดคะแนนพัฒนาการในด้านความตรงตามเกณฑ์สัมพัทธ์และค่าความคลาดเคลื่อนด้วยสถิติทดสอบ Hotelling และแผนภาพการกระจาย

ผลการวิจัยสรุปได้ดังนี้

- 1) คะแนนพัฒนาการที่ได้จากการประมาณค่าด้วยวิธีการวัดทั้ง 3 วิธี มีความตรงตามเกณฑ์สัมพัทธ์อยู่ในระดับสูง และให้ค่าความคลาดเคลื่อนของคะแนนพัฒนาการต่ำ
- 2) เมื่อข้อมูลมีการเปลี่ยนแปลงเป็นแบบเส้นตรง และมีค่าสัมประสิทธิ์พื้นฐานเป็น $[0, 1, 2, 3, 4]$ ในภาพรวมคะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่น รองลงมาได้แก่คะแนนพัฒนาการจากวิธีการวัด DS และ BRG เมื่อพิจารณาจากเงื่อนไขจำนวนครั้งของการวัดซ้ำ พบว่าคะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่นทุกจำนวนครั้งของการวัดซ้ำ
- 3) เมื่อข้อมูลมีลักษณะการเปลี่ยนแปลงเป็นแบบเส้นตรง และมีค่าสัมประสิทธิ์พื้นฐานเป็น $[0, 0.2, 0.4, 0.6, 0.8]$ ในภาพรวมคะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่น รองลงมาได้แก่คะแนนพัฒนาการจากวิธีการวัด BRG และ DS เมื่อข้อมูลมีจำนวนครั้งของการวัดซ้ำเป็น 3 ครั้งและ 4 ครั้ง คะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่น ส่วนข้อมูลที่มีจำนวนครั้งของการวัดซ้ำเป็น 5 ครั้ง คะแนนพัฒนาการจากวิธีการวัด BRG มีคุณภาพมากกว่าวิธีอื่น
- 4) สำหรับข้อมูลที่มีลักษณะของการเปลี่ยนแปลงไม่เป็นแบบเส้นตรง ในภาพรวมคะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่น รองลงมาได้แก่คะแนนพัฒนาการจากวิธีการวัด BRG และ DS เมื่อข้อมูลมีจำนวนครั้งของการวัดซ้ำเป็น 3 ครั้ง คะแนนพัฒนาการจากวิธีการวัด BRG มีคุณภาพมากกว่าวิธีอื่น ส่วนข้อมูลที่มีจำนวนครั้งของการวัดซ้ำเป็น 4 ครั้ง และ 5 ครั้ง คะแนนพัฒนาการจากวิธีการวัด RG มีคุณภาพมากกว่าวิธีอื่น

The purposes of this research were to compare the quality of the three measuring growth methods based on the classical test theory which were 1) observed different score method (DS), 2) relative gain score (RG), and 3) balanced relative gain score (BRG) using the latent growth curve model as the criterion for the data that had changed in linear and nonlinear growth model and had different time-point-measurement. The 108 situations were built up by monte carlo technique using MPLUS version 4.1 from 4 conditions which were 1) form of data-changing, 2) time-point-measurement, 3) mean of level (ML), and 4) mean of slope (MS). The data were analyzed in order to estimate growth scores, and the scores quality using criterion-related validity and measurement error were compared by Hotelling test and scatter plot.

The research results indicated that:

1) The growth scores from all of the three measuring growth methods had high level in criterion-related validity and low measurement errors.

2) For the data that had changed in linear growth model and basis coefficient is $[0,1,2,3,4]$, it showed that the method of RG had better quality than DS and BRG, respectively. Even if in the different time-point-measurement condition.

3) For the data that had changed in linear growth model with basis coefficient $[0,0.2,0.4,0.6,0.8]$, it showed that the method of RG had better quality than BRG and DS, respectively. It still had a better quality for 3 or 4 time-point-measurement conditions. But whenever the data had 5 time-point-measurement, the method of BRG had better quality than others.

4) For the data that had changed in nonlinear growth model, it showed that the method of RG had better quality than BRG and DS, respectively. It still had a better quality for 5 time-point-measurement conditions. But whenever the data had 3 or 4 time-point-measurement, the method of BRG had better quality than others.