คุณภาพของแบบประเมินความเสี่ยงและความจำเป็นในการคัดกรองเด็กและเยาวชนที่กระทำความผิด
The quality of the "risks and needs assessments scale" in case classification of juvenile offenders

ดร.ธัญลักษ์ รุจิภักดิ์[*]

▣ บทคัดย่อ

รายงานการวิจัยนี้มีวัตถุประสงค์เพื่อทดสอบคุณภาพของแบบประเมินความเสี่ยงในการกระทำผิดซ้ำและความจำเป็นในการบำบัดฟื้นฟูของเด็กและเยาวชนที่กระทำความผิด โดยการทดสอบค่าอำนาจจำแนก ความตรงตามสภาพ ความตรงเชิงทำนาย และความเที่ยงของแบบประเมิน การศึกษาครั้งนี้ใช้ระเบียบวิธีวิจัยเชิงปริมาณ กลุ่มตัวอย่างในการศึกษาครั้งนี้ ได้แก่ เด็กและเยาวชนจำนวนทั้งหมด 1,181 ราย การทดสอบคุณภาพของเครื่องมือประเมินความเสี่ยงและความจำเป็นโดยใช้สถิติในการวิเคราะห์ข้อมูล การตรวจสอบค่าอำนาจจำแนกพบว่า ค่าอำนาจจำแนกรายข้ออยู่ในระดับต้องปรับปรุงจนถึงดี ซึ่งสามารถแยกกลุ่มเด็กและเยาวชนได้ ผลการทดสอบความตรงตามสภาพ พบว่าเด็กและเยาวชนในกลุ่มความเสี่ยง และกลุ่มความรุนแรง มีความแตกต่างกันในปัจจัยด้านต่าง ๆ ทั้งการประเมินด้านความเสี่ยงและการประเมินด้านความจำเป็นอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ยกเว้น ปัจจัยด้านสภาวะกายและจิต ปัจจัยด้านปัญหาพฤติกรรม และปัจจัยด้านการคุ้มครองสวัสดิภาพเด็ก แสดงให้เห็นว่าแบบประเมินนั้นมีความตรงตามสภาพในระดับปานกลาง ส่วนการประเมินความตรงเชิงทำนาย พบว่า ค่าความสัมพันธ์ระหว่างปัจจัยด้านต่าง ๆ ของคะแนนความเสี่ยงและ

[*] *อาจารย์ประจำหลักสูตรอาชญาวิทยา การบริหารงานยุติธรรมและสังคม ภาควิชาสังคมศาสตร์ มหาวิทยาลัยมหิดล*

ความจำเป็นมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และไม่มีความสัมพันธ์คู่ใดมีความสัมพันธ์กันสูงกว่า 0.80 ซึ่งแสดงให้เห็นว่าปัจจัยด้านต่าง ๆ นั้นวัดในสิ่งที่แตกต่างกันอย่างเห็นได้ชัด นอกจากนี้ จากการทดสอบค่าความสามารถในการทำนายการแบ่งกลุ่มของเด็กจากปัจจัยด้านต่าง ๆ พบว่าค่าความสามารถในการทำนายอยู่ในระดับปานกลางถึงดี คือมีค่าตั้งแต่ ร้อยละ 62 ถึงร้อยละ 77 ผลการทดสอบค่าความเที่ยงของแบบประเมินพบว่า ค่าความเที่ยงทั้งฉบับเท่ากับ 0.62 ซึ่งอยู่ในระดับที่สามารถยอมรับ จากการทดสอบแสดงให้เห็นว่าแบบประเมินความเสี่ยงและความจำเป็นมีประสิทธิภาพในระดับดี สามารถจำแนกกลุ่มเด็กและเยาวชนได้ในระดับที่ยอมรับได้ แม้ว่าบางข้อมีค่าอำนาจจำแนกต่ำ อย่างไรก็ตามข้อคำถามที่ไม่สามารถจำแนกกลุ่มเด็กและเยาวชนได้ หน่วยงานที่เกี่ยวข้องจะทำการปรับเนื้อหาและค่าการให้คะแนนอีกครั้งหนึ่งเพื่อความเหมาะสมต่อไป

คำสำคัญ: คุณภาพเครื่องมือการวิจัย / การทดสอบความตรง / การทดสอบความน่าเชื่อถือ

Abstract

The purpose of this research is to test the construct and develop the validity and reliability of the assessment scale for assessing risks (of re-offending) and needs (for correction) in juvenile offender classification. The discrimination power, concurrent and predictive validity and internal consistency were statistically analyzes. A quantitative approach was applied for data collection. A total sample of 1,181 juvenile offenders was collected. The discrimination index of the assessment was ranged from low to excellent levels, indicating acceptable level of classification. The concurrent validity presented there was a significant difference between risk and severity of crime groups in all factors, excluding physical and psychological harm factor, behavior problems factor and children protection and welfare factor. The predictive validity was also found the moderately significant relationship between risk and need assessment, and the problem of multicollinearity was not found. The prediction of the model accounts for 62% and 77% of the variation in assessing risk and needs classification. These validity tests presented that the assessment was moderate accuracy valid. The internal consistency reliability coefficient of the assessment was acceptable (alpha = .62). The efficiencies of the measurement validity and reliability were acceptable. The Interview questionnaire to “Estimate Risks and Needs” is an extensive instrument that is intended to predict risk and severity in Thai adolescents. Although few items are low discrimination index, which is less likely affect. Some items’ contents and item’ scores need to be adjusted, in order to improve the accuracy of the assessment.

Key Words: Quality of the Questionnaire / Validity / Reliability

▣ บทนำ

กรมพินิจและคุ้มครองเด็กและเยาวชน เป็นหน่วยงานที่มีภารกิจเกี่ยวข้องกับเด็กและเยาวชนกระทำผิดกฎหมายที่เข้าสู่กระบวนการยุติธรรม โดยเริ่มตั้งแต่การสืบเสาะและพินิจเรื่องอายุ ประวัติ ความประพฤติ สติปัญญา การศึกษา สุขภาพ ภาวะแห่งจิต นิสัย อาชีพของเด็กและเยาวชน ซึ่งต้องสืบเสาะการกระทำความผิด และของบิดามารดา ผู้ปกครอง หรือบุคคลที่เด็กและเยาวชนอาศัยอยู่ ตลอดจนสิ่งแวดล้อมทั้งปวงที่เกี่ยวข้องกับเด็กและเยาวชน รวมถึงสาเหตุของการกระทำความผิดเพื่อรายงานต่อศาล ในขณะเดียวกันก็มีหน้าที่ให้การสงเคราะห์ แก้ไข และบำบัดเด็กและเยาวชนในระหว่างควบคุมหรือภายหลังปล่อย รวมทั้งจัดการศึกษา ฝึกและอบรมวิชาชีพสามัญ ตลอดจนดูแลอบรม สั่งสอนเด็กและเยาวชนซึ่งอยู่ในความควบคุมด้วย โดยสรุปแล้วภารกิจของกรมพินิจและคุ้มครองเด็กและเยาวชน คือ การแก้ไข บำบัด ฟื้นฟูเด็กและเยาวชนที่กระทำความผิด เพื่อให้เด็กและเยาวชนสามารถกลับตนเป็นคนดี และไม่หวนกลับไปกระทำผิดซ้ำอีก

นอกจากนี้กรมพินิจและคุ้มครองเด็กและเยาวชนยังต้องทำงานร่วมกับกรมคุมประพฤติ โดยการประสานส่งต่อข้อมูลที่ได้จากการจำแนก และการแก้ไข บำบัด ฟื้นฟูให้แก่กรมคุมประพฤติ เพื่อใช้ข้อมูลในการกำหนดแผนการคุมประพฤติของเด็กและเยาวชนได้อย่างเหมาะสม ตลอดจนใช้เป็นข้อมูลในการแยกเด็กและเยาวชนออกเป็นกลุ่มต่าง ๆ ซึ่งจะช่วยให้สามารถกำหนดสถานที่ควบคุมตัว และสถานฝึกและอบรมเด็กและเยาวชนที่เหมาะสมกับลักษณะของเด็กและเยาวชนแต่ละราย ทั้งนี้เพื่อให้สามารถจัดระดับการควบคุมและสอดส่องให้สอดคล้องกับระดับปัญหาของเด็กและเยาวชน รวมทั้งเพื่อกำหนดแผนการแก้ไขฟื้นฟูในชุมชนได้สอดคล้องกับสภาพปัญหา และความต้องการของผู้ถูกคุมประพฤติแต่ละราย กรมพินิจและคุ้มครองเด็กและเยาวชน กรมคุมประพฤติ และศาลเยาวชนและครอบครัว จึงจำเป็นต้องมีการพัฒนาและปรับปรุงเครื่องมือการจำแนกเด็กและเยาวชนเพื่อใช้ในการปฏิบัติงาน และสามารถนำมาแก้ไขและพัฒนาระบบ และเครื่องมือการจำแนกเด็กและเยาวชนที่เข้าสู่กระบวนการยุติธรรมได้ต่อไป

ในการพัฒนาแบบคัดกรองในการจำแนกเด็กและเยาวชนนั้น แบบประเมินความเสี่ยงและความจำเป็นจัดได้ว่าเป็นเครื่องมือที่มีความสำคัญมากในการนำไปพัฒนาระบบการปฏิบัติงานของกรมพินิจและคุ้มครองเด็กและเยาวชน แบบประเมินความเสี่ยงและความจำเป็นสำหรับคัดกรองเด็กและเยาวชนพัฒนาขึ้นเพื่อใช้ในการประเมินความเสี่ยงต่อการกระทำผิดซ้ำและความจำเป็นในการแก้ไข บำบัด ฟื้นฟูของเด็กและเยาวชน และเพื่อใช้ประเมินเด็กและเยาวชนที่เข้าสู่กระบวนการดำเนินคดีอาญาตามปกติให้ได้รับโอกาสในการใช้มาตรการพิเศษแทนการดำเนินคดีอาญา และเพื่อให้พนักงานคุมประพฤติได้มีเครื่องมือที่ใช้ในการรวบรวมข้อมูลเกี่ยวกับเด็กและเยาวชนและครอบครัว ใช้ประกอบการจัดทำรายงานการสืบเสาะที่มีความครบถ้วนและครอบคลุมปัจจัยที่มีความเกี่ยวข้องกับความเสี่ยงและความจำเป็นของเด็กและเยาวชนที่เข้าสู่กระบวนการยุติธรรม และลดความซ้ำซ้อนในการสืบเสาะข้อมูล โดยจัดให้ข้อมูลที่มีความสำคัญอยู่ในแบบประเมินชุดเดียวกัน

การนำแบบประเมินนี้มาใช้เป็นเครื่องมือในการจำแนกเด็กและเยาวชนตามสภาพความเสี่ยงในการกระทำผิดกฎหมายซ้ำ และความจำเป็นที่จะต้องได้รับการแก้ไข บำบัด ฟื้นฟู โดยนำข้อมูลที่ได้มาสรุปเป็นรายงานเพื่อเสนอวิธีการสำหรับเด็กและเยาวชน และแนวทางการแก้ไข บำบัด ฟื้นฟูเด็กและเยาวชนต่อศาลเยาวชนและครอบครัว แบบประเมินดังกล่าวนี้จึงมีความสำคัญอย่างมากในการทำหน้าที่เป็นเครื่องมือที่ช่วยในการเก็บและรวบรวมข้อมูลที่เกี่ยวข้องกับความเสี่ยงและความจำเป็นดังกล่าวอย่างเป็นหมวดหมู่ เพื่อใช้ในการตัดสินใจกำหนดวิธีการดังกล่าวสำหรับเด็กและเยาวชน โดยการดำเนินการในขั้นนี้เริ่มต้นจากการกำหนดข้อคำถามในการกำหนดประเด็นหลักในการตั้งข้อคำถามเพื่อให้ครอบคลุมทั้งปัจจัยความเสี่ยงในการกระทำผิดกฎหมายซ้ำ และปัจจัยความจำเป็นที่จะต้องได้รับการแก้ไข บำบัด ฟื้นฟูของเด็กและเยาวชน จึงทำให้แบบประเมินที่มีการพัฒนามานั้นต้องมีความถูกต้องและความแม่นยำ (Accuracy) ในการนำมาจำแนกกลุ่มเด็กและเยาวชนต่อไป อย่างไรก็ตาม ในการนำแบบประเมินไปใช้นั้นอาจไม่เป็นไปตามความคาดหวังได้อย่างถูกต้องสมบูรณ์ อาจจะเกิดความผิดพลาดในการจำแนกกลุ่มพฤติกรรมเด็กและเยาวชนอาจเกิดคลาดเคลื่อนได้ เพราะฉะนั้นควรมีการตรวจสอบคุณภาพของเครื่องมือในการประเมินในด้านความแม่นยำในการทำนายเพื่อจำแนกความเสี่ยงในการประทำผิดซ้ำและความจำเป็นในการบำบัดฟื้นฟูเพื่อพัฒนาคุณภาพของการนำไปใช้จริงต่อไป และสิ่งที่มีความสำคัญอีกประการก็คือความน่าเชื่อถือของเครื่องมือที่จะนำมาใช้ในการจำแนกกลุ่มเด็กและเยาวชนว่ามีความเหมาะสมและความน่าเชื่อถือเพียงใด กล่าวคือในการเก็บรวบรวมข้อมูลเพื่อให้ได้ข้อมูลที่ถูกต้อง น่าเชื่อถือ ยอมรับได้ ถือว่าเป็นสิ่งสำคัญ ซึ่งจะต้องมีความรู้และความชำนาญในการใช้เครื่องมือให้ได้มาซึ่งข้อมูลที่ตรงตามความต้องการ สามารถนำไปวิเคราะห์และสรุปผลได้ตรงตามวัตถุประสงค์ (Bryman, 2008) ถ้าเครื่องมือ หรือมาตรวัดสามารถวัดสิ่งที่ต้องการวัดได้ ก็กล่าวได้ว่ามาตรวัดหรือเครื่องชี้วัดนั้น ๆ มีความแม่นตรง (Valid) ซึ่งคำว่า *"ความแม่นตรง"* (Thomas, 2007) นักวิชาการไทยบางท่านเรียกว่า *"ความตรงประเด็น"* หมายถึง การวัดสิ่งที่ต้องการวัดได้อย่างถูกต้องแม่นยำ ในการวัดหรือการทดสอบเรื่องหนึ่ง ๆ ถ้าเครื่องมือในการวัดขาดความแม่นตรงแล้ว ผลสรุป ย่อมต้องผิดพลาดอย่างแน่นอน (บุญเรียง, 2543)

การตรวจสอบคุณภาพของเครื่องมือสามารถตรวจสอบได้สองแนวทางคือ แนวทางแรกได้แก่ การตรวจสอบเครื่องมือโดยอาศัยหลักเกณฑ์ทฤษฎี และความคิดเห็นของผู้เชี่ยวชาญตัดสินว่าเหมาะสมหรือไม่ จะเป็นการพิจารณาความเหมาะสมของเนื้อหา และมีการปรับสำนวนภาษาโดยผู้ที่เกี่ยวข้อง และผู้เชี่ยวชาญเรียบร้อยแล้วนั้น แบบประเมินได้ถูกนำไปทำการทดสอบจากผู้ที่มีส่วนเกี่ยวข้องในการใช้เครื่องมือ ได้แก่ เจ้าหน้าที่คุมประพฤติ และเจ้าหน้าที่สถานพินิจและคุ้มครองเด็กและเยาวชน พนักงานคุมประพฤติ นักวิชาการอบรมและฝึกวิชาชีพ พยาบาลวิชาชีพ ซึ่งจะเป็นการตรวจสอบในเรื่องของสำนวนภาษา และความยาก-ง่ายในการถาม-ตอบของเด็กที่อยู่ในสถานพินิจและคุ้มครองเด็กและเยาวชน ซึ่งผลจากการประเมินการทดลองใช้เครื่องมือของผู้ที่เกี่ยวข้องที่มีต่อข้อคำถามแต่ละข้อนั้น ทำให้ได้มีการปรับแบบประเมินเพื่อให้มีความง่ายและสะดวกในการ

นำไปใช้ โดยการตัดข้อคำถามที่ซับซ้อน และปรับสำนวนภาษาในบางข้อคำถาม และมีการปรับเปลี่ยนในเรื่องความหมาย และองค์ประกอบของข้อคำถาม เพื่อให้เกิดความเหมาะสมยิ่งขึ้นต่อการนำไปใช้ต่อไป จากนั้นก็จะเป็นการนำแบบประเมินไปทดลองใช้กับกลุ่มตัวอย่าง เพื่อการพัฒนาคุณภาพของแบบประเมินในขั้นตอนต่อไป ซึ่งวิธีการดังกล่าวเป็นวิธีการในส่วนที่สองซึ่งต้องแนวทางในการวิเคราะห์สถิติเพื่อตรวจสอบคุณภาพของเครื่องมือ ได้แก่ การหาค่าความเที่ยง (Validity) ความเชื่อมั่นหรือความตรง (Reliability) และค่าอำนาจจำแนก (Power of Discrimination) ของแบบประเมินเพื่อนำผลทางสถิติไปใช้ในการปรับปรุงแก้ไขแบบประเมินต่อไป

กล่าวได้ว่าในการประเมินคุณภาพของเครื่องมือในการคัดกรองเด็กและเยาวชน โดยใช้แบบประเมินความเสี่ยงและความจำเป็นเป็นขั้นตอนหนึ่งในการตรวจสอบว่าแบบประเมินนี้สามารถนำไปใช้ในการจำแนกกลุ่มเด็กและเยาวชนได้จริงหรือไม่ โดยใช้หลักสถิติเป็นเกณฑ์ในการตัดสินใจ และผลการประเมินจากเครื่องมือประเมินความเสี่ยงและความจำเป็นสามารถนำไปใช้ในการปรับปรุงและพัฒนาแบบประเมินให้มีประสิทธิภาพมากยิ่งขึ้น นอกจากนี้ผลการวิจัยสามารถช่วยให้ผู้เชี่ยวชาญหรือผู้ที่เกี่ยวข้องมีความมั่นใจในการตัดสินใจในการใช้เครื่องมือในการจำแนกเด็กและเยาวชนมากขึ้น โดยเฉพราะในการจำแนกเด็กและเยาวชนที่กระทำผิดว่าต้องได้รับการดูแล แก้ไข บำบัด ฟื้นฟู หรือควรมีการดำเนินการอย่างไรต่อไป เพื่อประโยชน์สูงสุดในการป้องกันและฟื้นฟูการกระทำความผิดของเด็กและเยาวชนต่อไปในอนาคต

▣ วัตถุประสงค์

วัตถุประสงค์ของการศึกษาครั้งนี้เพื่อนำเสนอการทดสอบคุณภาพของแบบประเมินที่ใช้ในการคัดกรองเด็กและเยาวชน คือ *"แบบประเมินความเสี่ยงและความจำเป็นในการคัดกรองเด็กและเยาวชนที่กระทำความผิด"* ที่พัฒนาโดยกรมพินิจและคุ้มครองเด็กและเยาวชน กระทรวงยุติธรรม

▣ ขอบเขตของการศึกษา

ในการตรวจสอบคุณภาพของเครื่องมือในการคัดกรองเด็กและเยาวชน โดยแบบประเมินความเสี่ยงและความจำเป็นนั้น ผู้วิจัยได้กำหนดขอบเขตในการตรวจสอบคุณภาพแบบประเมินดังต่อไปนี้

แผนภาพที่ 1 ขอบเขตการทดสอบคุณภาพแบบประเมินความเสี่ยงและความจำเป็น

การทดสอบคุณภาพของแบบประเมินความเสี่ยงและความจำเป็นในการจำแนกเด็กและเยาชนในการศึกษาครั้งนี้ได้ใช้วิธีการทางสถิติในการตัดสินใจในการปรับปรุงแบบประเมิน หลังจากได้มีการตรวจสอบความตรงเชิงเนื้อหาของแบบประเมินความเสี่ยงและความจำเป็นแล้ว โดยวิธีการทดสอบค่าอำนาจจำแนก (Item discriminant) การทดสอบความตรงใช้วิธีการทดสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion-Related validity analysis) ซึ่งการทดสอบความตรงแบบนี้เป็นการแสดงให้เห็นถึงความสามารถในการวัดของแบบทดสอบในเรื่องคัดกรองเด็กและเยาวชนโดยการทำนายพฤติกรรมของกลุ่มตัวอย่าง โดยทำการทดสอบ 2 แบบ คือ ความตรงตามสภาพ (Concurrent validity) และความตรงตามทำนาย (Predicted validity) และการทดสอบความเที่ยงหรือความเชื่อมั่นของเครื่องมือ โดยการทดสอบความสอดคล้องภายใน (Internal consistency reliability)

▣ ระเบียบวิธีวิจัย

การทดสอบคุณภาพของเครื่องมือในการศึกษาครั้งนี้ได้ประยุกต์ใช้วิธีการศึกษาเชิงปริมาณ (Quantitative approach) ในการดำเนินการเก็บข้อมูล โดยใช้แบบสอบถามในการสัมภาษณ์ (Interview questionnaire) กลุ่มประชากรที่ใช้ในการศึกษาเพื่อตรวจสอบคุณภาพของแบบประเมินความเสี่ยงและความจำเป็นคือ กลุ่มเด็กและเยาวชนที่กระทำผิดกฎหมาย โดยอยู่ในระหว่างรอการพิจารณาคดีและอยู่ในระหว่างการฝึกอบรมในศูนย์ฝึกและอบรมเด็กและเยาวชนชายบ้านอุเบกขา ทั้งหมด 27 ศูนย์ รวมกลุ่มตัวอย่างจำนวนทั้งหมด 1,181 คน ในการกำหนดกลุ่มตัวอย่างในการศึกษานั้น กำหนดได้ 4 ลักษณะ โดยพิจารณาจากความเสี่ยงที่จะกระทำผิดซ้ำ (Risk) ความรุนแรงของการกระทำผิดซ้ำ (Severity) คือ 1) ความเสี่ยงที่จะกระทำผิดซ้ำสูง (High risk) คือ เด็กและเยาวชนที่มีประวัติถูกดำเนินคดีตั้งแต่ 4 ครั้งขึ้นไป 2) ความเสี่ยงที่จะกระทำผิดซ้ำต่ำ (Low risk) คือ เด็กและเยาวชนที่มีประวัติเพิ่งถูกดำเนินคดีเป็นครั้งแรก 3) การกระทำผิดซ้ำมีความรุนแรงมาก (High severity) คือ เด็กและเยาวชนที่เคยถูกดำเนินคดี (หรือเป็นข้อหาปัจจุบันในกรณีที่ถูกดำเนินคดีเป็นครั้งแรก) ในข้อหาใดข้อหาหนึ่งหรือหลายข้อหาในจำนวนข้อหาต่อไปนี้ ฆ่า รุมโทรม ข่มขืน และทำร้ายร่างกาย และ 4) การกระทำผิดซ้ำมีความรุนแรงน้อย (Low severity) คือ เด็กและเยาวชนที่ไม่เคยถูกดำเนินคดี (หรือเป็นข้อหาปัจจุบันในกรณีที่ถูกดำเนินคดีเป็นครั้งแรก) ในข้อหาใดข้อหาหนึ่งหรือหลายข้อหาในจำนวนข้อหาต่อไปนี้ ฆ่า รุมโทรม ข่มขืน และทำร้ายร่างกาย ซึ่งในการทดสอบคุณภาพเครื่องมือจะจำแนกเด็กและเยาวชนออกเป็น 4 กลุ่มจำแนกตามความเสี่ยงในการกระทำผิดซ้ำ (เสี่ยงสูง-ต่ำ) และความรุนแรงในการกระทำความผิด (รุนแรงสูง-ต่ำ) ได้แก่ (1) กลุ่มเสี่ยงสูง-รุนแรงสูง (2) กลุ่มเสี่ยงสูง-รุนแรงต่ำ (3) กลุ่มเสี่ยงต่ำ-รุนแรงสูง (4) กลุ่มเสี่ยงต่ำ-รุนแรงต่ำ

การวิเคราะห์ข้อมูลโดยใช้สถิติในการวิเคราะห์คือ 1) การวิเคราะห์โดยการหาค่าอำนาจจำแนกรายข้อ เพื่อประเมินความสัมพันธ์ระหว่างข้อคำถามรายข้อในแบบประเมินกับคะแนนรวมของแบบประเมิน (Corrected item-total correlation) ถ้าค่าความสัมพันธ์ระหว่างข้อคำถามกับคะแนนรวมมีความสัมพันธ์ตั้งแต่ 0.2 ขึ้นไปหมายถึงข้อคำถามนั้นมีค่าอำนาจ

ในการจำแนกรายข้อ 2) การวิเคราะห์ความตรงสภาพ ใช้สถิติ t-test และ One-way ANOVA ในการเปรียบเทียบกลุ่มเด็กและเยาวชน 3) การวิเคราะห์ความตรงตามทำนายได้ทำการทดสอบโดยการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson product moment coefficients) ในการวิเคราะห์ความสัมพันธ์ของตัวแปรในแต่ละด้านของคะแนนด้านความเสี่ยงและความจำเป็น 4) การวิเคราะห์จำแนกประเภท (Discriminant analysis) เพื่อที่จะวิเคราะห์ได้ว่าปัจจัยใดบ้างที่สามารถระบุกลุ่มความเสี่ยงและความรุนแรงในการกระทำความผิดของเด็กและเยาวชนได้ 5) การวิเคราะห์ค่าความเที่ยงเชิงความสอดคล้องภายใน (Internal consistency reliability) ด้วยวิธีการหาค่าสัมประสิทธิ์อัลฟ่าของครอนบาค (Cronbach's alpha coefficients) ของแบบประเมินทั้งฉบับและองค์ประกอบแต่ละด้าน

▣ ผลการทดสอบคุณภาพเครื่องมือและการอภิปรายผล

1. ข้อมูลทั่วไปของกลุ่มเด็กและเยาวชน

กลุ่มเด็กและเยาวชนที่ศึกษาส่วนใหญ่เป็นเพศชาย (ร้อยละ 95.3) มากกว่าร้อยละ 80 อายุมากกว่า 15 ปี และเมื่อพิจารณาในเรื่องของลักษณะในการกระทำความผิดพบว่าคดียาเสพติดจะมีจำนวนเด็กและเยาวชนกระทำความผิดมากที่สุด (ร้อยละ 37.26) รองลงมาคือ คดีเกี่ยวกับชีวิตและร่างกาย (ร้อยละ 26.33) คดีเกี่ยวกับทรัพย์สิน (ร้อยละ 23.37) คดีอื่น ๆ เช่น ขับขี่, บุกรุก, อาวุธ (ร้อยละ 13.04) ตามลำดับ และเมื่อจำแนกเด็กและเยาวชนออกเป็น 4 กลุ่มจำแนกตามความเสี่ยง และความรุนแรงพบว่า ประมาณครึ่งหนึ่ง (ร้อยละ 52.2) เด็กและเยาวชนจะอยู่ในกลุ่มเสี่ยงต่ำ-รุนแรงต่ำ รองลงมาคือ กลุ่มเด็กและเยาวชนเสี่ยงสูง-รุนแรงต่ำ (ร้อยละ 19.8) ซึ่งจะใกล้เคียงกับกลุ่มเด็กและเยาวชนเสี่ยงต่ำ-รุนแรงสูง (ร้อยละ 19.2) และมีเพียงร้อยละ 8.4 ที่มีเป็นกลุ่มเด็กและเยาวชนเสี่ยงสูง-รุนแรงสูง

2. การทดสอบอำนาจจำแนก

การวิเคราะห์ค่าสหสัมพันธ์ระหว่างข้อคำถามแต่ละข้อกับคะแนนรวมของข้ออื่น ๆ พบว่าแบบประเมินในเรื่องของคะแนนความเสี่ยงและความจำเป็นมีค่าอำนาจจำแนกตั้งแต่ 0.006 ถึง 0.578 มีข้อคำถามบางข้อที่มีค่าอำนาจในการจำแนกต่ำกว่า 2 และเมื่อพิจารณาความแตกต่างระหว่างกลุ่มกับข้อคำถามรายข้อพบว่า มีคำถามในบางข้อ ไม่มีความแตกต่างระหว่างกลุ่ม แสดงให้เห็นว่าคำถามในบางข้อไม่สามารถจำแนกกลุ่มของเด็กและเยาวชนได้ ซึ่งในกรณีนี้ผู้ที่มีส่วนเกี่ยวข้องควรจะพิจารณาปรับปรุงในเรื่องข้อคำถามใหม่และบางคำถามควรพิจารณาในเรื่องของการกำหนดการให้คะแนนรายข้อใหม่ ตัวอย่างผลการศึกษาบางส่วนแสดงในตารางที่ 1 ข้อคำถามที่มีการประเมินว่าควรปรับปรุงคือ ข้อคำถามนั้นไม่สามารถจำแนกกลุ่มของเด็กและเยาวชนได้อย่างแน่นอน กล่าวคือผู้ที่มีส่วนเกี่ยวข้องต้องพิจารณาข้อคำถามนั้น ๆ ใหม่ทั้งในส่วนของเนื้อหา และในส่วนของการให้คะแนนรายข้อ เพื่อให้มีความสอดคล้องกับเนื้อหาในการประเมินมากขึ้น ส่วนข้อคำถามที่มีการประเมินว่าพิจารณาใหม่นั้น ก็คือข้อคำถามที่มีค่าอำนาจในการจำแนกรายข้อค่อนข้างต่ำมาก หมายความว่าข้อคำถามนั้นถึงแม้ว่าจะมีค่าอำนาจจำแนกที่ต่ำแต่ก็สามารถ

แบ่งกลุ่มเด็กและเยาวชนได้ในระดับปานกลาง ดังนั้นข้อคำถามที่ไม่ผ่านการประเมิน คือไม่สามารถจำแนกกลุ่มได้ จะต้องมีการปรับปรุงด้านเนื้อหา การให้คะแนน รวมทั้งสำนวนภาษาให้เหมาะสมในการพิจารณาข้อคำถามแต่ละข้อ เพื่อความสอดคล้องในเรื่องการวิเคราะห์การประเมินความเสี่ยงและความจำเป็นต่อไป

ตารางที่ 1 ตัวอย่างผลการทดสอบค่าอำนาจจำแนกรายข้อในการประเมินคะแนนด้านความเสี่ยง

ข้อคำถาม		r	F-test	การประเมิน
rf1	สถานภาพของบิดามารดา	0.250	65.557	ใช้ได้
rf2	ลักษณะการอยู่อาศัย	0.165	22.827	พิจารณาใหม่
rf3	วิธีการควบคุมพฤติกรรมของเด็กและเยาวชน	0.506	263.271	ใช้ได้
rf4	ลักษณะความสัมพันธ์ระหว่างบิดา มารดา ผู้ปกครอง	0.186	36.911	พิจารณาใหม่
rf5	ลักษณะความสัมพันธ์ระหว่างเด็ก/เยาวชนกับบิดามารดา/ผู้ปกครอง	0.522	411.616	ใช้ได้
rf6	แบบอย่างที่ไม่ดีของบุคคลในครอบครัวและบุคคลที่เด็ก/เยาวชนอาศัยอยู่	0.267	53.353	ใช้ได้
rf7	ความเจ็บป่วยทางสุขภาพกายและจิตของบุคคลในครอบครัว	0.121	10.490	พิจารณาใหม่
rc5	ผลร้ายที่เกิดกับชีวิตร่างกายของผู้เสียหาย	0.052	0.503	ควรปรับปรุง

จากผลการศึกษากล่าวได้ว่าข้อคำถามที่มีค่าอำนาจจำแนกค่อนข้างต่ำนั้นหมายถึงคำถามในข้อนั้น ๆ อาจจะมีความไม่เหมาะสมในการจำแนกกลุ่มของเด็กและเยาวชน เนื่องจากเป็นลักษณะข้อคำถามที่มีความกำกวม หรือเป็นข้อคำถามที่มีความไม่ชัดเจนในการวัดการประเมิน ซึ่งทำให้ไม่สามารถใช้ในการจำแนกกลุ่มของเด็กและเยาวชนได้ อย่างไรก็ตาม การพัฒนาให้ค่าอำนาจจำแนกมีค่าสูงขึ้นนั้นสามารถทำได้ (สุวิมล ติรกานันท์, 2551) โดยการปรับข้อคำถามและการให้พิจารณาค่าคะแนนของข้อคำถามนั้นใหม่ เนื่องจากสาเหตุอีกประการหนึ่งที่ทำให้แบบประเมินความเสี่ยงและความจำเป็นมีค่าอำนาจจำแนก มีค่าที่ค่อนข้างต่ำนั้น เนื่องจากการหาค่าอำนาจจำแนกด้วยวิธีนี้นั้นค่า r จะมีค่าอำนาจจำแนกสูงในกรณีที่ข้อมูลมีระดับการวัดเป็นช่วง หรือข้อมูลที่มีระดับการวัดแบบ Likert Scale ซึ่งแบบประเมินความเสี่ยงและความจำเป็นในครั้งนี้ ระดับการวัดส่วนใหญ่จะเป็นระดับการวัดแบบ Nominal และ Ordinal Scale ซึ่งอาจจะเป็นประเด็นที่ทำให้ค่าอำนาจค่อนข้างต่ำกว่าเกณฑ์มาตรฐาน ดังนั้น นอกจากเนื้อหาและการให้คะแนนแล้ว ระดับการวัดในข้อคำถามแต่ละด้านควรที่จะพิจารณาในการปรับเปลี่ยนต่อไป

3. การทดสอบความตรงของแบบประเมินความเสี่ยงและความจำเป็น

ความตรงตามสภาพ (Concurrent validity)

การทดสอบความตรงตามสภาพนั้นมีด้วยการหลากหลายวิธี สำหรับการศึกษาครั้งนี้ได้ทดสอบความตรงตามสภาพโดยวิธีการวิเคราะห์ความแตกต่างของกลุ่มเด็กและเยาวชนจำแนกตามกลุ่มความเสี่ยง (เสี่ยงสูง - เสี่ยงต่ำ) และกลุ่มความรุนแรง (รุนแรงสูง -

รุนแรงต่ำ) ในเรื่องความเสี่ยงในการกระทำผิดซ้ำ และความจำเป็นในการบำบัดฟื้นฟู ผลการศึกษาจากตารางที่ 2 พบว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงสูงนั้นมีคะแนนเฉลี่ยรวมด้านความเสี่ยงสูงกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มที่มีความเสี่ยงต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ($t = 12.326$, $p < 0.001$) อย่างไรก็ตาม เมื่อพิจารณารายด้านของคะแนนเฉลี่ยความเสี่ยง พบว่า เด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงสูงนั้นจะมีคะแนนเฉลี่ยในแต่ละด้านสูงกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ยกเว้นด้านสภาวะทางกายและจิต ซึ่งพบว่ากลุ่มเด็กและเยาวชนทั้งสองกลุ่มมีคะแนนด้านสภาวะทางกายและจิตไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

เมื่อพิจารณาคะแนนด้านความจำเป็นโดยรวม (ตารางที่ 2) พบว่า กลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงสูงนั้นมีคะแนนเฉลี่ยรวมด้านความจำเป็นสูงกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มที่มีความเสี่ยงต่ำ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ($t = 10.631$, $p < 0.001$) อย่างไรก็ตาม เมื่อพิจารณารายด้านของคะแนนเฉลี่ยด้านความจำเป็นพบว่าเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงสูงนั้น จะมีคะแนนเฉลี่ยในแต่ละด้านสูงกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ยกเว้นด้านสภาวะทางกายและจิต และด้านการคุ้มครองสวัสดิภาพเด็ก กล่าวคือกลุ่มเด็กและเยาวชนทั้งสองกลุ่มมีคะแนนด้านสภาวะทางกายและจิต และด้านการคุ้มครองสวัสดิภาพเด็กไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ซึ่งหมายความว่า ไม่ว่าเด็กและเยาวชนจะอยู่กลุ่มใด คะแนนด้านสภาวะทางกายและจิต และด้านการคุ้มครองสวัสดิภาพเด็ก ก็ไม่แตกต่างกัน แสดงให้เห็นว่าคะแนนความเสี่ยงด้านสภาวะทางกายและจิต และด้านการคุ้มครองสวัสดิภาพเด็ก ไม่สามารถวัดหรือจำแนกกลุ่มเด็กและเยาวชนที่มีระดับความเสี่ยงต่างกันได้

ตารางที่ 2 คะแนนความเสี่ยงและความจำเป็นระหว่างกลุ่มความเสี่ยงสูง-ความเสี่ยงต่ำ

แบบประเมินด้านต่างๆ	กลุ่มเสี่ยงสูง (n =305)		กลุ่มเสี่ยงต่ำ (n = 774)		t-value
	Mean	SD	Mean	SD	
คะแนนการประเมินด้านความเสี่ยง	**39.00**	**10.50**	**29.61**	**10.44**	**12.236**
1. ครอบครัว	14.58	6.08	12.15	5.61	6.036
2. ชุมชนและสภาพแวดล้อม	0.67	0.47	0.55	0.49	3.511
3. ภูมิหลังทางการศึกษาและสังคม	7.62	3.14	5.55	3.27	9.629
4. เพื่อนและคนใกล้ชิด	2.50	1.27	1.87	1.28	7.253
5. ประวัติการเกี่ยวข้องกับยาเสพติด	2.02	1.16	1.30	1.03	9.946
6. ประวัติการกระทำผิด	7.62	2.54	5.57	1.96	12.629
7. สภาวะทางกายและทางจิต	0.03	0.17	0.01	0.10	1.330
8. ปัญหาพฤติกรรม	3.96	2.58	2.60	1.96	8.303

แบบประเมินด้านต่างๆ	กลุ่มเสี่ยงสูง (n =305)		กลุ่มเสี่ยงต่ำ (n = 774)		t-value
	Mean	SD	Mean	SD	
คะแนนการประเมินด้านความจำเป็น	**31.35**	**11.03**	**24.79**	**10.32**	**10.631**
1. ครอบครัว	14.89	6.22	12.41	5.68	6.046
2. ชุมชนและสภาพแวดล้อม	0.67	0.47	0.55	0.49	3.511
3. ภูมิหลังทางการศึกษาและสังคม	8.23	3.25	6.00	3.40	9.792
4. เพื่อนและคนใกล้ชิด	2.50	1.27	1.87	1.28	7.253
5. ประวัติการเกี่ยวข้องกับยาเสพติด	2.02	1.16	1.30	1.03	9.946
6. สภาวะทางกายและทางจิต	0.08	0.29	0.05	0.24	1.346
7. ปัญหาพฤติกรรม	3.96	2.58	2.60	1.96	8.303
8. การคุ้มครองสวัสดิภาพเด็ก	0.01	0.08	0.00	0.05	0.967

หมายเหตุ: ค่า t-value น้อยกว่า 2 แสดงว่าไม่มีนัยสำคัญทางสถิติที่ระดับ 0.05

เมื่อพิจารณาความแตกต่างของกลุ่มเด็กและเยาวชนจำแนกตามความรุนแรงของการกระทำความผิด จากตารางที่ 3 พบว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงสูงและกลุ่มความรุนแรงต่ำนั้น มีค่าเฉลี่ยของคะแนนรวมด้านความเสี่ยงไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ($t = 0.745$, $p = 0.456$) แสดงว่าคะแนนด้านความเสี่ยงนั้นไม่สามารถวัดหรือจำแนกกลุ่มเด็กและเยาวชนที่กระทำความผิดในคดีที่มีความรุนแรงสูงและความรุนแรงต่ำได้ เมื่อพิจารณาเป็นรายด้านนั้น พบว่า เด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงสูงและความรุนแรงต่ำไม่มีความแตกต่างในคะแนนเฉลี่ยความเสี่ยงด้านภูมิหลังทางการศึกษาและสังคม ด้านสภาวะกายและจิต และด้านปัญหาพฤติกรรม ส่วนด้านอื่น ๆ นั้นผลการศึกษาพบว่า มีความแตกต่างกันระหว่างกลุ่มเด็กและเยาวชนทั้งสองกลุ่มอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อพิจารณาคะแนนเฉลี่ยด้านความจำเป็นโดยรวมพบว่า กลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงสูงมีคะแนนรวมด้านความจำเป็นเฉลี่ยต่ำกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 ($t = 3.753$, $p < 0.001$) อย่างไรก็ตามเมื่อพิจารณาคะแนนเฉลี่ยรายด้านของคะแนนด้านความจำเป็นพบว่า เด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงสูงมีคะแนนเฉลี่ยต่ำกว่ากลุ่มเด็กและเยาวชนที่อยู่ในกลุ่มความรุนแรงต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ยกเว้นด้านภูมิหลังทางการศึกษาและสังคม ด้านสภาวะทางกายและจิต ด้านปัญหาพฤติกรรม และด้านการคุ้มครองสวัสดิภาพเด็ก

นอกจากนี้ เมื่อพิจารณาความแตกต่างของกลุ่มเด็กและเยาวชนทั้ง 4 กลุ่ม (เสี่ยงต่ำ-รุนแรงต่ำ, เสี่ยงต่ำ-รุนแรงสูง, เสี่ยงสูง-รุนแรงต่ำ, และเสี่ยงสูง-รุนแรงสูง) กับความแตกต่างของคะแนนเฉลี่ยในการประเมินด้านความเสี่ยงและความจำเป็น ผลการศึกษาพบว่า กลุ่มที่มีคะแนนเฉลี่ยรวมในการประเมินด้านความเสี่ยง

แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยกลุ่มเด็กและเยาวชนที่มีความเสี่ยงสูง-ความรุนแรงต่ำจะมีคะแนนเฉลี่ยด้านความเสี่ยงสูงกว่ากลุ่มอื่น ๆ ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มเสี่ยงต่ำ-รุนแรงสูง จะมีคะแนนในการประเมินด้านความเสี่ยงต่ำสุด เมื่อพิจารณาเป็นรายด้านพบว่า เด็กและเยาวชนทั้ง 4 กลุ่มมีคะแนนเฉลี่ยของการประเมินด้านความเสี่ยงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ในทุกด้าน ยกเว้นคะแนนการประเมินด้านสภาวะทางกายและทางจิต ส่วนคะแนนการประเมินด้านความจำเป็นพบว่า กลุ่มเด็กและเยาวชนทั้ง 4 กลุ่มมีคะแนนเฉลี่ยรวมในการประเมินด้านความจำเป็นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยกลุ่มเด็กและเยาวชนที่มีความเสี่ยงสูง-ความรุนแรงต่ำจะมีคะแนนเฉลี่ยด้านความจำเป็นสูงกว่ากลุ่มอื่น ๆ ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มเสี่ยงต่ำ-รุนแรงสูงจะมีคะแนนในการประเมินด้านความจำเป็นต่ำสุด

ตารางที่ 3 คะแนนความเสี่ยงและความจำเป็นระหว่างกลุ่มความรุนแรงสูง-รุนแรงต่ำ

แบบประเมินด้านต่างๆ	กลุ่มรุนแรงสูง (*n* = 298)		กลุ่มรุนแรงต่ำ (n = 781)		t-value
	Mean	SD	Mean	SD	
คะแนนการประเมินด้านความเสี่ยง	**31.84**	**12.24**	**32.43**	**11.29**	**0.745**
1. ครอบครัว	11.67	6.13	13.28	5.67	4.061
2. ชุมชนและสภาพแวดล้อม	0.52	0.50	0.61	0.48	2.465
3. ภูมิหลังทางการศึกษาและสังคม	5.89	3.27	6.23	3.40	1.507
4. เพื่อนและคนใกล้ชิด	1.74	1.24	2.17	1.31	4.889
5. ประวัติการเกี่ยวข้องกับยาเสพติด	1.22	0.95	1.62	1.15	5.696
6. ประวัติการกระทำผิด	7.84	2.51	5.50	1.89	14.517
7. สภาวะทางกายและทางจิต	0.03	0.18	0.01	0.10	1.370
8. ปัญหาพฤติกรรม	2.93	2.32	3.00	2.21	0.486
คะแนนการประเมินด้านความจำเป็น	**24.82**	**11.63**	**27.73**	**10.73**	**3.753**
1. ครอบครัว	11.95	6.20	13.55	5.78	3.988
2. ชุมชนและสภาพแวดล้อม	0.52	0.50	0.61	0.48	2.465
3. ภูมิหลังทางการศึกษาและสังคม	6.40	3.43	6.72	3.53	1.338
4. เพื่อนและคนใกล้ชิด	1.74	1.24	2.17	1.31	4.889
5. ประวัติการเกี่ยวข้องกับยาเสพติด	1.22	0.95	1.62	1.15	5.696
6. สภาวะทางกายและทางจิต	0.05	0.24	0.06	0.26	0.384
7. ปัญหาพฤติกรรม	2.93	2.32	3.00	2.21	0.486
8. การคุ้มครองสวัสดิภาพเด็ก	0.01	0.82	0.00	0.51	0.818

หมายเหตุ: ค่า t-value น้อยกว่า 2 แสดงว่าไม่มีนัยสำคัญทางสถิติที่ระดับ 0.05

จากผลการศึกษาดังกล่าว เด็กและเยาวชนในกลุ่มความเสี่ยง และกลุ่มความรุนแรงมีความแตกต่างกันในปัจจัยด้านต่าง ๆ ทั้งการประเมินด้านความเสี่ยงในการกระทำผิดซ้ำและการประเมินด้านความจำเป็นในการบำบัดฟื้นฟู อย่างไรก็ตามมีปัจจัยในบางด้านซึ่งไม่มีความแตกต่างกันในกลุ่มเด็กและเยาวชน คือ ปัจจัยด้านสภาวะกายและจิต ปัจจัยด้านปัญหาพฤติกรรม และปัจจัยด้านการคุ้มครองสวัสดิภาพเด็ก ซึ่งหมายความว่าไม่ว่าเด็กและเยาวชนจะอยู่กลุ่มใด คะแนนด้านดังกล่าวก็ไม่แตกต่างกัน จึงไม่สามารถวัดหรือจำแนกกลุ่มเด็กและเยาวชนที่มีระดับความเสี่ยงในการกระทำผิดซ้ำรวมทั้งความจำเป็นในการบำบัดฟื้นฟูต่างกันได้ ดังนั้นผลการศึกษาสามารถยืนยันความตรงตามสภาพของแบบประเมินได้ในระดับปานกลางถึงระดับดี ดังนั้นกลุ่มข้อคำถามที่ไม่สามารถจำแนกเด็กและเยาวชนได้ ผู้ที่มีส่วนเกี่ยวข้องควรที่จะนำข้อคำถามเหล่านั้นมาพิจารณาใหม่ในประเด็นการให้คะแนนหรือการให้ค่าของคำตอบ รวมทั้งพิจารณาในเรื่องเนื้อหาของข้อคำถามใหม่เพื่อที่จะเพิ่มประสิทธิภาพของคุณภาพแบบทดสอบต่อไป

ความตรงเชิงทำนาย (Predictive validity)

ส่วนผลการศึกษาความตรงเชิงทำนายของแบบประเมินความเสี่ยงและความจำเป็นเป็นลักษณะการประเมินความสามารถของเครื่องมือที่จะบ่งบอกผลที่วัดได้ในขณะนั้นได้ถูกต้องตามสภาพจริงที่เกิดขึ้นในอนาคต ซึ่งการตรวจสอบความตรงเชิงทำนายนั้นเป็นการวัดความสามารถในการพยากรณ์ที่น่าจะเกิดขึ้นในอนาคต (บุญธรรม กิจปรีดาบริสุทธ์, 2553) ในการทดสอบความตรงเชิงพยากรณ์นั้นสามารถวัดได้หลายวิธี ในการศึกษาครั้งนี้ผู้วิจัยได้ประเมินความตรงเชิงทำนายโดยการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรด้านต่าง ๆ พบว่า ผลการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (ตารางที่ 4 และตารางที่ 5) พบว่า ความสัมพันธ์ระหว่างตัวแปรในแต่ละด้านของคะแนนด้านความเสี่ยงและความจำเป็นนั้น มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติระดับปานกลาง และไม่มีตัวแปรด้านใดมีความสัมพันธ์ในระดับสูงมาก ($r > 0.80$) ซึ่งแสดงให้เห็นว่า ความสัมพันธ์ของตัวแปรในแต่ละด้านนั้นมีความเป็นอิสระต่อกัน กล่าวคือ ตัวแปรในแต่ละด้านวัดในสิ่งที่แตกต่างกัน อย่างไรก็ตาม จากผลการศึกษาพบว่า ในการวัดความเสี่ยงนั้น ตัวแปรด้านสภาวะกายและจิต ไม่มีความสัมพันธ์กับด้านใดเลย ยกเว้นด้านภูมิหลังทางการศึกษา ด้านประวัติการกระทำผิด และด้านปัญหาพฤติกรรม ส่วนการวัดคะแนนด้านความจำเป็นนั้น พบว่าตัวแปรด้านสภาวะกายและจิตไม่มีความสัมพันธ์กับตัวแปรด้านใดเลย ยกเว้นด้านปัญหาพฤติกรรม และด้านการคุ้มครองสวัสดิภาพเด็ก

ตารางที่ 4 สัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยด้านต่าง ๆ ของคะแนนด้านความเสี่ยง

ด้านต่างๆ	1	2	3	4	5	6	7	8
1. ครอบครัว	1	.245**	.176**	.262**	.148**	.071*	.049	.061*
2. ชุมชนและสภาพแวดล้อม		1	.290**	.373**	.224**	.062*	.026	.051
3. ภูมิหลังทางการศึกษาและสังคม			1	.473**	.388**	.153**	.067*	.109**
4. เพื่อนและคนใกล้ชิด				1	.429**	.119**	.053	.127**
5. ประวัติการเกี่ยวข้องกับยาเสพติด					1	.110**	.056	.072*
6. ประวัติการกระทำผิด						1	.071*	.054
7. สภาวะทางกายและทางจิต							1	094**
8. ปัญหาพฤติกรรม								1

*** p < 0.01, *p< 0.05*

ตารางที่ 5 สัมประสิทธิ์สหสัมพันธ์ระหว่างปัจจัยด้านต่าง ๆ ของคะแนนด้านความจำเป็น

ด้านต่างๆ	1	2	3	4	5	6	7	8
1. ครอบครัว	1	.253**	.212**	.316**	.141**	.027	.087**	.049
2. ชุมชนและสภาพแวดล้อม		1	.272**	.373**	.224**	.058	.051	.021
3. ภูมิหลังทางการศึกษาและสังคม			1	.455**	.371**	.025	.109**	.047
4. เพื่อนและคนใกล้ชิด				1	.429**	.023	.127**	.066*
5. ประวัติการเกี่ยวข้องกับยาเสพติด					1	.046	.072*	.006
6. สภาวะทางกายและทางจิต						1	.089**	.220**
7. ปัญหาพฤติกรรม							1	102**
8. การคุ้มครองสวัสดิภาพเด็ก								1

*** p < 0.01, *p< 0.05*

จากผลการทดสอบแสดงให้เห็นว่าเมื่อข้อคำถามบางส่วนแต่ละด้านวัดในสิ่งแตกต่างกันในการจำแนกเด็กและเยาวชน จึงกล่าวได้ว่าข้อคำถามแต่ละข้อนั้นมีความตรงของข้อคำถาม (Item Validity) จากผลการศึกษาแสดงให้เห็นว่าแบบประเมินความเสี่ยงและความจำเป็นนั้นสามารถที่จะวัดหรือคาดคะเนความสัมพันธ์ของปัจจัยด้านต่าง ๆ ได้ในระดับปานกลาง ซึ่งเป็นสิ่งที่ยืนยันคุณภาพของเครื่องมือได้ในระดับหนึ่ง นอกจากนี้ ผู้วิจัยได้ทำตรวจสอบความตรงเชิงทำนายโดยการทดสอบการวิเคราะห์จำแนกประเภท (Discriminant Analysis) ซึ่งผู้วิจัยได้แบ่งผลการวิจัยออกเป็น 4 กลุ่ม

1) การประเมินด้านความเสี่ยงในการกระทำผิดซ้ำของเด็กและเยาวชนมีอยู่ในกลุ่มเสี่ยงสูงและเสี่ยงต่ำ พบว่า ปัจจัยด้านประวัติการกระทำความผิด ประวัติเกี่ยวข้องกับยาเสพติด และภูมิหลังทางการศึกษาและ

สังคม มีอิทธิพลในการจำแนกกลุ่มเด็กและเยาวชนจำแนกตามความเสี่ยงได้ดีที่สุดตามลำดับ ส่วนปัจจัยทางด้านสภาวะกายและจิต และปัจจัยทางด้านชุมชนและสภาพแวดล้อมนั้น มีอิทธิพลน้อยที่สุดในการจำแนกกลุ่มเด็กและเยาวชนเสี่ยงสูงและเสี่ยงต่ำ นอกจากนี้ผลการศึกษาพบว่าความสามารถในการจำแนกกลุ่มในการประเมินด้านความเสี่ยงอยู่ในระดับดีคือ 74% กลุ่มความเสี่ยงต่ำสามารถที่จะคาดคะเนความถูกต้องได้ 76% ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มความเสี่ยงสูง สามารถที่จะคาดคะเนความถูกต้องได้ 69% จากผลการวิเคราะห์แสดงให้เห็นว่าความสามารถในการทำนายกลุ่มความเสี่ยงของปัจจัยด้านความเสี่ยงในด้านต่าง ๆ นั้นสามารถนำไปใช้ในการคาดคะเนกลุ่มเด็กและเยาวชนที่มีความเสี่ยงสูงและความเสี่ยงต่ำได้ดีพอสมควร

2) การประเมินด้านความจำเป็นในการบำบัด ฟื้นฟูของเด็กและเยาวชนมีอยู่ในกลุ่มเสี่ยงสูงและเสี่ยงต่ำนั้น พบว่า ปัจจัยด้านประวัติเกี่ยวข้องกับยาเสพติด ภูมิหลังทางการศึกษาและสังคม และปัญหาพฤติกรรมมีอิทธิพลในการจำแนกกลุ่มเด็กและเยาวชนจำแนกตามความเสี่ยงได้ดีที่สุด ตามลำดับ ส่วนปัจจัยทางด้านสภาวะกายและจิต และปัจจัยทางด้านการคุ้มครองสวัสดิภาพเด็กนั้นมีอิทธิพลน้อยที่สุดในการจำแนกกลุ่มเด็กและเยาวชนเสี่ยงสูงและเสี่ยงต่ำ นอกจากนี้ผลการศึกษาพบว่า ความสามารถในการจำแนกกลุ่มในการประเมินด้านความจำเป็นอยู่ในระดับปานกลาง คือ 68.2% นอกจากนี้พบว่า กลุ่มเสี่ยงต่ำสามารถที่จะคาดคะเนความถูกต้องได้ 70.4% ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มเสี่ยงสูง สามารถที่จะคาดคะเนความถูกต้องได้ 62.2% จากผลการวิเคราะห์แสดงให้เห็นว่าความสามารถในการทำนายกลุ่มความเสี่ยงของตัวแปรด้านต่าง ๆ ด้านความจำเป็นนั้นสามารถนำไปใช้ในการคาดคะเนกลุ่มเด็กและเยาวชนที่มีความเสี่ยงสูงและความเสี่ยงต่ำได้ในระดับปานกลาง

3) การประเมินด้านความเสี่ยงในการกระทำผิดซ้ำของเด็กและเยาวชนที่อยู่ในกลุ่มรุนแรงสูงและรุนแรงต่ำ พบว่าปัจจัยด้านประวัติด้านการกระทำผิด ปัจจัยด้านประวัติที่เกี่ยวข้องกับยาเสพติด และปัจจัยด้านเพื่อนและบุคคลใกล้ชิด จะมีอิทธิพลต่อการจำแนกกลุ่มความรุนแรงของเด็กและเยาวชนมากที่สุด ส่วนปัจจัยที่มีอิทธิพลน้อยที่สุดได้แก่ ปัจจัยด้านปัญหาพฤติกรรม ปัจจัยด้านภูมิหลังการศึกษา และปัจจัยด้านสภาวะกายและจิต ตามลำดับ และผลการศึกษาในการคาดคะเนการทำนายความถูกต้องในการประเมินความเสี่ยงของกลุ่มเด็กและเยาวชนจำแนกตามความรุนแรงอยู่ในระดับดี คือ 77% และเมื่อพิจารณารายกลุ่มพบว่าข้อมูลของกลุ่มรุนแรงต่ำสามารถที่จะคาดคะเนความถูกต้องได้ถึง 80% ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มรุนแรงสูง สามารถที่จะคาดคะเนความถูกต้องได้ 69.5% จากผลดังกล่าวแสดงให้เห็นว่าความสามารถในการทำนายกลุ่มความเสี่ยงของตัวแปรด้านต่าง ๆ นั้นสามารถนำไปใช้ในการคาดคะเนกลุ่มเด็กและเยาวชนที่มีความรุนแรงสูงและความรุนแรงต่ำได้ดีพอสมควร

4) การประเมินความจำเป็นในการบำบัดฟื้นฟูของกลุ่มเด็กและเยาวชนจำแนกตามกลุ่มเด็กรุนแรงสูง และรุนแรงต่ำ พบว่าปัจจัยที่มีอิทธิพลต่อการจำแนกกลุ่มเด็กและเยาวชนมากที่สุด 3 อันดับแรกได้แก่ ปัจจัยด้านประวัติเกี่ยวข้องกับยาเสพติด ปัจจัยด้านเพื่อนและคนใกล้ชิด และปัจจัยด้านครอบครัว ส่วนปัจจัย

ที่มีอิทธิพลน้อยที่สุดได้แก่ ปัจจัยด้านสภาวะกายและจิต และปัจจัยด้านปัญหาพฤติกรรม เมื่อพิจารณาความสามารถในการจำแนกกลุ่มในการประเมินด้านความจำเป็น พบว่าอยู่ในระดับปานกลาง คือ 61.2% และเมื่อพิจารณารายกลุ่ม พบว่า ข้อมูลของกลุ่มรุนแรงต่ำสามารถที่จะคาดคะเนความถูกต้องได้เพียง 58.8% ส่วนเด็กและเยาวชนที่อยู่ในกลุ่มรุนแรงสูงสามารถที่จะคาดคะเนความถูกต้องได้ 67.4% จากผลการวิเคราะห์แสดงให้เห็นว่าความสามารถในการทำนายกลุ่มความรุนแรงของปัจจัยด้านต่าง ๆ ในการประเมินความจำเป็นนั้นสามารถนำไปใช้ในการคาดคะเนกลุ่มเด็กและเยาวชนที่มีความรุนแรงสูงและความรุนแรงต่ำได้ในระดับปานกลางเท่านั้น

จากผลการประเมินดังกล่าวสามารถสรุปได้ว่า ถึงแม้ว่าการวัดความตรงเชิงทำนายหรือการทำนายความถูกต้องของการจำแนกกลุ่มและเยาวชนเพื่อเป็นแนวทางในการพยากรณ์กลุ่มในการจำแนกต่อไป มีค่าระดับความถูกต้องในการทำนายที่ไม่สูงมากนัก แต่ก็เป็นการยืนยันระดับหนึ่งว่า แบบประเมินความเสี่ยงและความจำเป็นนั้นสามารถที่จะใช้ในการจำแนกกลุ่มเด็กและเยาวชนจำแนกตามความเสี่ยงและความรุนแรงได้ในระดับปานกลาง

4. ความเที่ยงของแบบประเมิน (Reliability Study)

จากตารางที่ 6 พบว่าค่าความสอดคล้องภายในขององค์ประกอบแต่ละด้านของการมีค่าตั้งแต่ 0.43-0.69 โดยที่ปัจจัยด้านภูมิหลังทางการศึกษาและสังคมมีค่ามากที่สุด และปัจจัยด้านประวัติที่เกี่ยวข้องกับยาเสพติด และประวัติการกระทำความผิด มีค่าความเที่ยงน้อยที่สุด และซึ่งแสดงให้เห็นว่าความน่าเชื่อถือของแบบประเมินในแต่ละด้านนั้นมีความน่าเชื่อถือในระดับต่ำถึงระดับปานกลาง และพบว่า ค่าความเที่ยงของแบบประเมินความเสี่ยงและความจำเป็นรวมทั้งฉบับ คือ 0.62 ซึ่งถือว่ามีความน่าเชื่อถือในระดับปานกลาง

ตารางที่ 6 ค่าความเที่ยงเชิงความสอดคล้องภายใน (Internal consistency)

ปัจจัยด้านต่างๆ	Alpha Coefficient
1. ครอบครัว	.55
2. ชุมชนและสภาพแวดล้อม	n.a
3. ภูมิหลังทางการศึกษาและสังคม	.69
4. เพื่อนและคนใกล้ชิด	.50
5. ประวัติการเกี่ยวข้องกับยาเสพติด	.43
6. ประวัติการกระทำความผิด	.43
7. สภาวะทางกายและทางจิต	n.a
8. ปัญหาพฤติกรรม	.66
9. การคุ้มครองสวัสดิภาพเด็ก	n.a
10. การค้ามนุษย์	n.a

n.a ไม่สามารถคำนวณได้

จากข้อค้นพบข้างต้นในเรื่องของปัจจัยที่มีความเที่ยงในระดับต่ำ (ต่ำกว่า 0.50) นั้นมีความเป็นไปได้เนื่องจากข้อคำถามในแต่ละด้านนั้นมีช่วงคำตอบที่ไม่เท่ากัน และในบางปัจจัยนั้นข้อคำถามมีเพียง 2 หรือ 3 ข้อเท่านั้น ด้วยเหตุผลนี้ทำให้ค่าความเที่ยงมีค่าน้อย และในบางด้านนั้นไม่สามารถหาค่าความเที่ยงได้เนื่องจากการให้คะแนนในข้อคำถามแต่ละด้านมีเพียง 1 ข้อเท่านั้น และช่วงของคะแนนไม่สามารถกระจายได้คือ มีคำตอบเพียง 2 คำตอบ ทำให้ไม่สามารถคำนวณหาค่า Alpha Coefficient ได้ อย่างไรก็ตามเพื่อเป็นการเพิ่มค่าความน่าเชื่อถือของแบบประเมินข้อคำถามในบ้างด้านควรที่จะต้องมีการปรับปรุง พิจารณาใหม่อีกครั้ง โดยที่ผู้เกี่ยวข้องต้องไปปรับในเรื่องของช่วงการให้คะแนนในแต่ละด้านให้เป็นไปในทิศทางเดียวกันและสอดคล้องกัน

นอกจากนี้ สาเหตุที่ความเที่ยงของแบบประเมินมีค่าน้อยถึงปานกลาง เนื่องจากการวัดตัวแปรหรือข้อคำถามที่มีลักษณะทางกายภาพ (ความจริง) หรือตัวแปรที่สังเกตได้ เช่น ข้อมูลด้านครอบครัว ข้อมูลด้านชุมชนและสิ่งแวดล้อม หรือประวัติในการกระทำความผิด ซึ่งลักษณะข้อคำถามในลักษณะนี้มักไม่มีการประมาณค่าความเที่ยง เพราะข้อคำถามในลักษณะนี้จะมีความคงที่ในตัวของข้อคำถามอยู่แล้ว การประมาณค่าความเที่ยงแบบความสอดคล้องภายใน (Internal Consistency) จึงไม่ค่อยมีความเหมาะสมในการประเมินความเที่ยงในแบบประเมินความเสี่ยงและความจำเป็นเท่าที่ควร (สุวิมล ติรกานันท์, 2551) ซึ่งเป็นประเด็นที่สามารถเกิดขึ้นได้ และไม่มีผลกระทบกับคุณภาพความน่าเชื่อถือของข้อมูลในแบบประเมินความเสี่ยงและความจำเป็นเนื่องจากข้อคำถามส่วนใหญ่เป็นข้อคำถามเชิงข้อเท็จจริงของตัวเด็กและเยาวชนเอง อย่างไรก็ตามการศึกษาครั้งนี้ใช้วิธีการประเมินความเที่ยงของแบบประเมินโดยการใช้วิธีการทดสอบซ้ำ (Test-retest) กล่าวคือ ก่อนที่จะมีการทดสอบในครั้งนี้ ทางคณะผู้ที่เกี่ยวข้องได้ทำการทดสอบข้อคำถาม รวมทั้งคุณภาพแบบสอบถามเบื้องต้น (Pilot test) โดยการเก็บข้อมูลกับเด็กและเยาวชน จำนวน 20 คน ซึ่งได้ตรวจสอบความเที่ยงของแบบประเมินความเสี่ยงและความจำเป็น โดยใช้วิธีการหาค่าความเที่ยงระหว่างผู้ประเมินหรือผู้สังเกต (Inter-rater Reliability) (สุชีรา ภัทรายุตวรรตน์, 2545) ผลการหาค่าความเชื่อมั่นของเครื่องมือจำแนกได้ดำเนินการโดยใช้วิธีการ inter rater reliability พบว่า ค่าความเที่ยงอยู่ในระดับดีมาก (ตารางที่ 7) แสดงให้เห็นว่าแบบประเมินความเสี่ยงและความจำเป็นมีค่าความเชื่อมั่นระดับสูงในเกือบทุกด้านของการประเมิน

ตารางที่ 7 ค่าความเชื่อมั่นของแบบประเมินโดยการทดสอบ inter rater reliability

เครื่องมือจำแนก	Kendall's Coefficient of concordance	p-value	การแปลค่า*
1. ด้านครอบครัว	.766	.000	Good
2. ด้านเพื่อนบุคคลใกล้ชิด	.503	.025	Fair
3. ด้านชุมชนสภาพแวดล้อม	.851	.000	Excellent
4. ด้านภูมิหลังทางสังคม	.903	.000	Excellent
5. ด้านประวัติเกี่ยวข้องกับยาเสพติด	.945	.000	Excellent
6. ด้านประวัติการกระทำผิด	.907	.000	Excellent
7. ด้านการประเมินสภาวะทางกาย	.572	.008	Fair
8. ด้านการประเมินสภาวะทางจิต	.718	.000	Good
ความเสี่ยงและความจำเป็นรวม	.896	.000	Excellent

▣ สรุปผลการศึกษา

1. ค่าอำนาจจำแนกรายข้อ ในการศึกษาครั้งนี้ได้ใช้วิธีการตรวจสอบค่าอำนาจจำแนกรายข้อโดยการทดสอบค่าสหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม และการหาค่าความแตกต่างระหว่างกลุ่ม ซึ่งผลการศึกษาพบว่า ค่าอำนาจจำแนกรายข้อในบางคำถามนั้นประมาณ ร้อยละ 80 ของข้อคำถามมีค่าอำนาจในการจำแนกดี (ค่า r มากกว่า 0.20) และสามารถแยกกลุ่มเด็กและเยาวชนได้ อย่างไรก็ตามข้อคำถามที่ไม่สามารถจำแนกได้นั้น หน่วยงานที่เกี่ยวข้องควรที่จะนำมาพิจารณาในการปรับข้อคำถามและการให้ค่าคะแนนอีกครั้งหนึ่งเพื่อความเหมาะสมต่อไป

2. การประเมินความตรงตามเกณฑ์ (Criterion-Related Validity) ในการประเมินความตรงของแบบประเมินนั้นได้มีการทดสอบค่าความตรงตามเกณฑ์ ซึ่งจะแบ่งเป็น 2 ด้าน ได้แก่ การประเมินความตรงตามสภาพ หรือความตรงร่วมสมัย (Concurrent Validity) และความตรงเชิงทำนาย (Predictive Validity)

1) การประเมินความตรงตามสภาพนั้นได้ประเมินจากความแตกต่างระหว่างกลุ่มของเด็กและเยาวชน พบว่า เด็กและเยาวชนในกลุ่มความเสี่ยง และกลุ่มความรุนแรงมีความแตกต่างกันในปัจจัยด้านต่าง ๆ ทั้งการประเมินด้านความเสี่ยงในการกระทำผิดซ้ำ และการประเมินด้านความจำเป็นในการบำบัดฟื้นฟูอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 (ยกเว้น ปัจจัยด้านสภาวะกายและจิต ปัจจัยด้านปัญหาพฤติกรรม และปัจจัยด้านการคุ้มครองสวัสดิภาพเด็ก) แสดงให้เห็นว่าแบบประเมินนั้นมีความตรงตามสภาพในระดับปานกลาง อย่างไรก็ตาม ปัจจัยด้านที่ไม่มีความแตกต่างผู้ที่เกี่ยวข้องได้ทำการปรับข้อคำถามและการให้ค่าคะแนนของคำถามในด้านนั้น ๆ ใหม่

2) การประเมินความตรงเชิงทำนาย พบว่า ตัวแปรรายด้านมีความสัมพันธ์กันระดับปานกลาง นอกจากนี้ ผลการศึกษาพบว่า ไม่มีปัญหาในเรื่องของความสัมพันธ์ที่สูงมากของคู่ตัวแปรในแต่ละด้าน (Multi-colinearity) เนื่องจากไม่มีตัวแปรด้านใดมีความสัมพันธ์ในระดับสูงมาก ซึ่งแสดงให้เห็นว่าความสัมพันธ์ของตัวแปรในแต่ละด้านนั้นมีความเป็นอิสระต่อกัน อย่างไรก็ตามจากผลการศึกษาพบว่า ในการวัดความเสี่ยงในการกระทำผิดซ้ำนั้น ตัวแปรด้านสภาวะกายและจิต ไม่มีความสัมพันธ์กับด้านใดเลย ยกเว้นด้านภูมิหลังทางการศึกษา ประวัติการกระทำผิด และด้านปัญหาพฤติกรรม ส่วนการวัดคะแนนด้านความจำเป็นในการบำบัดฟื้นฟูพบว่า ตัวแปรด้านสภาวะกายและจิตไม่มีความสัมพันธ์กับตัวแปรด้านใดเลย ยกเว้นด้านปัญหาพฤติกรรมและการคุ้มครองสวัสดิภาพเด็ก จากผลการศึกษาแสดงให้เห็นว่าแบบประเมินความเสี่ยงและความจำเป็นนั้นสามารถที่จะวัดหรือคาดคะเนความสัมพันธ์ของปัจจัยด้านต่าง ๆ ได้ในระดับปานกลาง ซึ่งเป็นสิ่งที่ยืนยันคุณภาพของเครื่องมือได้ในระดับหนึ่ง

นอกจากนี้ผู้วิจัยได้ทำตรวจสอบความตรงเชิงทำนายโดยการทดสอบการวิเคราะห์จำแนกประเภท (Discriminant Analysis) จากการทดสอบค่าความสามารถในการทำนายการแบ่งกลุ่มของเด็กจากปัจจัยด้านต่าง ๆ พบว่า ค่าความสามารถในการทำนายการจำแนกเด็กและเยาวชนอยู่ในระดับปานกลางถึงดี คือมีค่าตั้งแต่ ร้อยละ 62 ถึง ร้อยละ 77 แสดงให้เห็นว่าความสามารถในการทำนายกลุ่มความรุนแรงของปัจจัยด้านต่าง ๆ ในการประเมินความจำเป็น นั้นสามารถนำไปใช้ในการคาดคะเนกลุ่มเด็กและเยาวชนที่มีความรุนแรงสูงและความรุนแรงต่ำได้ในระดับปานกลาง ถึงแม้ว่าการวัดความตรงเชิงทำนาย มีค่าระดับความถูกต้องในการทำนายที่ไม่สูงมากนัก แต่ก็เป็นการยืนยันระดับหนึ่งว่า แบบประเมินความเสี่ยงและความจำเป็นนั้นสามารถที่จะใช้ในการจำแนกกลุ่มเด็กและเยาวชนจำแนกตามความเสี่ยงและความรุนแรงได้

3. ความเที่ยงของแบบประเมิน (Reliability) เมื่อพิจารณาความน่าเชื่อถือของเครื่องมือโดยการทดสอบค่าความเที่ยงของแบบประเมินความเสี่ยงและความจำเป็นพบว่า ค่าความเที่ยงทั้งฉบับเท่ากับ 0.62 ซึ่งอยู่ในระดับที่สามารถยอมรับ และเพื่อพิจารณาการหาค่าความเที่ยงแบบ inter rater พบว่าค่าความเที่ยงอยู่ระดับดีมาก คือ 0.89 ถึงแม้ว่าค่าความสอดคล้องภายในแบบประเมินความเสี่ยงและความจำเป็นในการศึกษาครั้งนี้มีค่าที่ไม่สูงมากนัก ซึ่งเป็นประเด็นที่สามารถเกิดขึ้นได้และไม่มีผลกระทบกับคุณภาพความน่าเชื่อถือของข้อมูลในแบบประเมินมากนัก เนื่องจากข้อคำถามส่วนใหญ่เป็นข้อคำถามเชิงข้อเท็จจริงของตัวเด็กและเยาวชนเองไม่ได้เป็นข้อคำถามที่วัดทัศนคติแต่อย่างใด

จากข้อค้นพบในการตรวจสอบคุณภาพของเครื่องมือในการประเมินความเสี่ยงในการกระทำผิดซ้ำและความจำเป็นในการบำบัดฟื้นฟู พบว่า เครื่องมือดังกล่าวมีคุณภาพในระดับปานกลางถึงดี ในการจำแนกเด็กและเยาวชนที่กระทำความผิด แต่อย่างไรก็ตามข้อคำถามในบางประเด็นควรพิจารณาในเรื่องเนื้อหา จำนวนข้อคำถามในแต่ละด้าน และการให้ค่าคะแนนในแต่ละด้านเพื่อให้เกิดความเหมาะสมต่อไป

▣ ข้อเสนอแนะจากการศึกษา

1. เนื่องจากแบบประเมินความเสี่ยงและความจำเป็นมีข้อคำถามเป็นจำนวนมาก ซึ่งบางข้อคำถามไม่ได้นำมาใช้ในการให้ค่าคะแนน และบางข้อคำถามไม่สามารถนำมาจำแนกกลุ่มเด็กและเยาวชนได้ เพราะฉะนั้นข้อคำถามเหล่านั้นน่าจะพิจารณาตัดออกจากแบบประเมิน อย่างไรก็ตามการตัดข้อคำถามโดยอิงจากค่าสถิตินั้นเป็นการพิจารณาในเบื้องต้นเท่านั้น ผู้ที่มีส่วนเกี่ยวข้องควรจะพิจารณาข้อคำถามโดยพิจารณาในส่วนของเนื้อหาอีกครั้งหนึ่งว่าข้อคำถามดังกล่าวมีความสำคัญหรือมีส่วนเกี่ยวข้องกับประเด็นในการประเมินหรือไม่ ถ้าประเด็นคำถามดังกล่าวมีความสำคัญในเชิงเนื้อหา หรือทฤษฎีก็ควรจะพิจารณาปรับปรุงข้อคำถาม แทนการตัดข้อคำถามนั้นออกไป แต่ถ้าหากข้อคำถามดังกล่าวสามารถพิจารณาตัดออกไปได้ ก็ควรพิจารณาตัดออกจากแบบประเมิน ซึ่งการปรับลดข้อคำถามนั้นเพื่อให้แบบประเมินมีความกระชับมากยิ่งขึ้น และเป็นการประหยัดเวลาของพนักงานคุมประพฤติในการสัมภาษณ์เด็กและเยาวชน

2. เนื่องจากคะแนนที่ได้จากแบบประเมินความเสี่ยงและความจำเป็นนำมาใช้ในการจัดกลุ่มเด็กและเยาวชนเป็น 3 กลุ่ม ได้แก่ กลุ่มที่มีความเสี่ยงในการกระทำผิดซ้ำต่ำ-ปานกลาง-สูง และกลุ่มที่มีความจำเป็นในการบำบัดฟื้นฟูต่ำ-ปานกลาง-สูง แต่การออกแบบข้อคำถามปลายปิดของแบบประเมินบางส่วนระดับการวัดมีเพียง 2 ระดับเท่านั้น และในข้อคำถามด้านอื่นมีระดับการวัดที่ไม่เท่ากัน ซึ่งปัญหาเหล่านี้อาจมีผลกระทบต่อความเที่ยงและความตรงของแบบประเมินได้ เพื่อที่จะปรับปรุงในส่วนนี้ ผู้ที่มีส่วนเกี่ยวข้องในการพัฒนาแบบประเมินดังกล่าว ควรที่จะคำนึงถึงระดับการวัดของข้อคำถามให้เป็นไปในรูปแบบเดียวกัน หรือมีระดับการวัดที่เหมือนกัน หรือเพื่อให้เป็นมาตรฐานเดียวกันนั้นเอง นอกจากนี้ ถ้าระดับการวัดมีการปรับปรุงให้เป็นระดับเดียวกันและมีความสอดคล้องกันในทุกด้านของแบบประเมิน ผู้วิจัยสามารถที่จะวิเคราะห์ความตรงเชิงโครงสร้างของแบบประเมินในแต่ละด้านได้ โดยใช้การวิเคราะห์องค์ประกอบ (Factor Analysis) ในการพิจารณาความตรงเชิงโครงสร้าง ซึ่งจะเป็นการทดสอบได้ว่าข้อคำถามแต่ละข้อนั้นสามารถวัดโครงสร้างด้านนั้นได้หรือไม่อย่างไร

3. จากการศึกษาพบว่า แบบประเมินบางด้าน เช่น ด้านที่ 2 ชุมชนและสภาพแวดล้อม ด้านที่ 9 การคุ้มครองสวัสดิภาพเด็ก และด้านที่ 10 การค้ามนุษย์ มีข้อคำถามเพียงข้อเดียวเท่านั้นที่ใช้วัดในด้านดังกล่าว ซึ่งในด้านเหล่านี้จะมีผลกระทบโดยตรงต่อการหาความตรงและความเที่ยงของเครื่องมือ ในการวัดองค์ประกอบทางโครงสร้างของแบบประเมินในแต่ละด้านควรที่จะมีคำถามอย่างน้อย 4 ข้อ (Kline, 2005) เพื่อที่จะสามารถนำมาตรวจสอบคุณภาพให้ชัดเจนได้มากขึ้น นอกจากนี้จำนวนข้อคำถามที่เพิ่มมากขึ้นจะทำให้ค่าความเที่ยงของแบบประเมินเพิ่มขึ้นได้ อย่างไรก็ตามการเพิ่มข้อคำถามก็ควรที่จะคำนึงในด้านเนื้อหาประกอบด้วยเช่นกัน นอกจากนี้การเพิ่มข้อคำถามให้มีคุณภาพดีหมายถึงข้อคำถามที่มีความเหมาะสมและมีค่าอำนาจจำแนกสูง (สุวิมล ติรกานันท์, 2551) ซึ่งในทางตรงกันข้ามถ้าข้อคำถามที่เพิ่มเข้ามา ไม่มีความสอดคล้องกับองค์ประกอบหรือด้านนั้น ๆ ก็อาจจะส่งผลให้คุณภาพของเครื่องมือลดลงด้วย

4. นอกจากนี้จากข้อค้นพบดังกล่าว ผู้วิจัยได้นำเสนอข้อเสนอแนะของการพัฒนาคุณภาพเครื่องมือที่ใช้ในการวิจัย หรือการประเมินของศิริชัย กาญจนวาสี (2544: 118-122) และ สุวิมล ติรกานันท์ (2551) ซึ่งสรุปได้ว่าปัจจัยที่ส่งผลต่อความเที่ยงตรงของแบบประเมินนั้นขึ้นอยู่กับปัจจัยต่าง ๆ หลากหลายด้วยกัน ถ้าต้องการปรับปรุงแบบประเมินหรือแบบสอบถามให้มีความน่าเชื่อถือมากยิ่งขึ้นควรที่จะพิจารณาในเรื่องต่อไปนี้คือ

4.1 ความชัดเจนของข้อคำถาม และการตอบคำถามของแบบประเมิน ควรมีความชัดเจนใช้ภาษาที่รัดกุม ไม่กำกวม ซึ่งถ้าภาษาที่ใช้ในคำถามและคำตอบมีความคลุมเครือมักจะสร้างความสับสนให้แก่ผู้ใช้แบบสอบถามหรือแบบประเมินได้ และทำให้ผู้ที่นำแบบประเมินไปใช้มีความสับสน และมีความยากในการใช้แบบประเมินในการสัมภาษณ์หรือแม้กระทั่งการตอบของผู้ตอบ ก็จะทำให้ส่งผลต่อคำตอบได้เช่นกัน ในประเด็นนี้สามารถทำให้ค่าอำนาจจำแนกรายข้อของแบบประเมินมีค่าต่ำได้ ซึ่งจะมีผลต่อการใช้แบบประเมินในการวัดได้ นอกจากนี้ข้อคำถามควรจะตรงกับเรื่องที่จะทำการประเมิน ส่วนในเรื่องของการกำหนดคำตอบเป็นคำถามปลายปิดนั้น ควรมีความชัดเจนในการให้คะแนนของคำตอบแต่ละข้อ และควรมีการให้คำจำกัดความของคำตอบแต่ละหัวข้อให้ชัดเจน ซึ่งถ้าผู้ถามและผู้ตอบมีความเข้าใจที่ชัดเจนและตรงกันนั้น สามารถที่จะเพิ่มความน่าเชื่อถือของแบบสอบถามได้ คือคำถามสามารถวัดในสิ่งที่ผู้วิจัยต้องการรู้หรือศึกษาได้

4.2 ความเหมาะสมของเนื้อหาที่จะทำการวัดหรือทำการประเมิน ในเรื่องดังกล่าวได้แก่เรื่องขอบเขต และประเด็นคำถามในแต่ละข้อว่ามีความชัดเจนในการวัดหรือไม่ และข้อคำถามดังกล่าวย่อมมีผลต่อการตัดสินของผู้เชี่ยวชาญเกี่ยวกับความครอบคลุมด้านเนื้อหาและความเป็นตัวแทนของข้อคำถามที่จะใช้วัดองค์ประกอบในแต่ละด้าน นอกจากนี้แล้วการวัดความเหมาะสมของสิ่งเหล่านี้เป็นเกณฑ์ทางด้านปัจจัยที่เกี่ยวข้องกับทฤษฎีที่ใช้ในการพัฒนาข้อคำถามในแต่ละด้าน ซึ่งใช้ค่าสถิติเป็นเกณฑ์ในการอ้างอิงในการตัดสินใจควบคู่กันไป

ในกรณีดังกล่าว ทางผู้ที่มีส่วนเกี่ยวข้องในการพัฒนาแบบประเมินในการประเมินความเสี่ยงและความจำเป็นได้ตระหนักถึงเรื่องดังกล่าว และได้มีการปรับปรุงด้านภาษา และการให้ค่าคะแนนในระดับเบื้องต้นแล้ว รวมทั้งการใช้สถิติในการวิเคราะห์เพื่อเป็นเกณฑ์ในการตัดสินใจหรืออ้างอิงในเรื่องคุณภาพของเครื่องมือในระดับหนึ่ง แต่อย่างไรก็ตาม จากผลการศึกษาพบว่า คำถามบางประเด็นสมควรที่จะพิจารณาและปรับปรุงแก้ไขก่อนที่จะนำไปใช้เป็นแบบสอบถามมาตรฐานต่อไป ซึ่งหากพิจารณาและทำการควบคุมปัจจัยด้านการใช้ภาษา การให้ค่าคะแนนที่เหมาะสม รวมทั้งการใช้สถิติเป็นเกณฑ์ในการอ้างอิงให้มีความเหมาะสม ก็จะทำให้ได้เครื่องมือที่สามารถวัดในสิ่งที่ต้องการได้อย่างมีประสิทธิภาพ

▣ การพัฒนาการศึกษาในอนาคต (Future Study)

จากข้อค้นพบต่าง ๆ ในงานวิจัยนี้ นำไปสู่การพัฒนาการศึกษาต่อไป ดังต่อไปนี้

1. การศึกษาครั้งนี้เป็นการศึกษาที่ใช้กับกลุ่มตัวอย่างในสถานพินิจและคุ้มครองเด็กและเยาวชนที่เป็นกลุ่มตัวอย่างเท่านั้น ด้วยเหตุนี้

นักวิชาการ หรือผู้ที่มีส่วนเกี่ยวข้องกับการบำบัด ฟื้นฟูเด็กและเยาวชน ควรที่จะพัฒนารูปแบบการประเมินให้สามารถใช้กับกลุ่มเด็กทั่วไปที่มีความเสี่ยงในการกระทำผิด เช่น เด็กนักเรียนในสถานศึกษา เพื่อที่จะสามารถประเมินระดับความเสี่ยงในการกระทำความผิดของกลุ่มเด็กเหล่านี้ได้ ซึ่งเมื่อได้ทราบปัจจัยความเสี่ยงที่เกี่ยวข้องกับแนวโน้มของการกระทำความผิดแล้ว นักวิชาการที่เกี่ยวข้องก็สามารถนำมาวิเคราะห์ เพื่อหารูปแบบหรือแนวทางในการป้องกันการกระทำความผิดที่อาจจะเกิดขึ้นในอนาคต

2. นักวิชาการหรือผู้มีส่วนเกี่ยวข้องอาจพัฒนารูปแบบของแบบประเมินความเสี่ยงและความจำเป็นของเด็กและเยาวชนเป็นแบบประเมินขนาดย่อหรือฉบับย่อ (Shot form version) ที่สามารถนำมาใช้แทนแบบประเมินเต็มรูปแบบได้ โดยการพัฒนาแบบประเมินขนาดย่อนั้น เพื่อเป็นการประหยัดเวลาในการคัดกรองเด็กและเยาวชนเบื้องต้นให้มีความกระชับมากยิ่งขึ้น แต่อย่างไรก็ตามในการพัฒนาแบบประเมินขนาดย่อนั้น ก็ควรที่จะยึดหลักการพัฒนาเหมือนฉบับเต็ม และต้องมีการตรวจสอบความสอดคล้องของแบบประเมินฉบับเต็มกับฉบับย่อด้วยว่าสามารถใช้แทนกันได้หรือไม่ โดยใช้วิธีการทางสถิติเข้ามาช่วยในการทดสอบดังกล่าว

3. นักวิชาการหรือผู้มีส่วนเกี่ยวข้องควรที่จะนำรูปแบบเครื่องมือในการคัดกรองเด็กที่ใช้ในต่างประเทศ หรือในกลุ่มประเทศอาเซียน หรือแบบที่เป็นสากลมาพัฒนาให้เหมาะสมกับรูปแบบบริบทของสังคมไทย เพื่อที่จะสามารถเปรียบเทียบรูปแบบในการพัฒนา หรือการคัดกรองเด็กและเยาวชนในประเทศกับต่างประเทศได้ เนื่องจากถ้าเครื่องมือที่ใช้มีความใกล้เคียงกัน ก็สามารถนำแนวทางในการแบ่งกลุ่มเด็กและเยาวชนมาเปรียบเทียบกันได้ เพื่อการพัฒนาที่เป็นสากลต่อไป

4. เนื่องจากสังคมมีการเปลี่ยนแปลงตลอดเวลาและการกระทำผิดของเด็กและเยาวชนก็มีรูปแบบที่แตกต่างกันออกไปตามยุคตามสมัย เพราะฉะนั้นเครื่องมือที่นำมาใช้ในการคัดกรองเด็กและเยาวชนควรที่จะมีการพัฒนาทุก ๆ ปี หรือทุก ๆ 3 ปี เพื่อที่ปรับปรุงให้มีความเหมาะสมมากยิ่งขึ้น

5. เมื่อพิจารณาในเรื่องของข้อคำถามหรือตัวแปรที่ใช้ในการศึกษาครั้งนี้ พบว่าตัวแปรส่วนใหญ่จะเป็นตัวแปรที่มีลักษณะทางกายภาพ หรือตัวแปรที่มีลักษณะคงที่ ซึ่งในการศึกษาครั้งต่อไป ผู้ที่มีส่วนเกี่ยวข้องน่าจะมีการศึกษา หรือนำตัวแปรที่เป็นตัวแปรแอบแฝง (Latent Variable) หรือมาตรวัดทางจิตวิทยามาประยุกต์ใช้ในการทดสอบปัจจัยที่มีอิทธิพลในการจำแนกกลุ่มเด็กและเยาวชนตามความเสี่ยงและความรุนแรง ควบคู่กับปัจจัยที่ใช้ในการประเมินในปัจจุบัน ซึ่งนอกจากจะทำให้รู้ถึงปัจจัยภายนอกที่มีอิทธิพลต่อการจำแนกกลุ่มเด็กและเยาวชนแล้ว ยังทำให้ทราบถึงปัจจัยภายในตัวของเด็กและเยาวชนอีกด้วย ซึ่งน่าจะเป็นประโยชน์ในการแก้ไข บำบัด ฟื้นฟูเด็กและเยาวชนที่กระทำความผิด รวมทั้งเป็นข้อมูลในการป้องกันการกระทำผิดซ้ำของเด็กและเยาวชนได้แหล่งหนึ่งด้วย

6. การพัฒนาแบบประเมินในการคัดกรองเด็กและเยาวชนนั้น เป็นรูปแบบการประเมินใหม่ซึ่งนำมาใช้กับระบบงานในการบำบัด ฟื้นฟูเด็กและเยาวชน เพราะฉะนั้นเพื่อเป็นการมั่นใจว่าแบบประเมินดังกล่าวมีความเหมาะสมกับระบบงานยุติธรรมในสังคมไทย นักวิชาการหรือผู้ที่มีส่วนเกี่ยวข้องควรที่จะจัดทำโครงการ

ในการประเมินผลการนำเครื่องมือดังกล่าวไปใช้ในมิติของความสอดคล้องเชิงยุทธศาสตร์ขององค์กร เพื่อที่จะทำให้ทราบส่วนดี ส่วนที่ควรปรับปรุง หรืออุปสรรคในการใช้งาน รวมทั้งโอกาสในการพัฒนาให้เหมาะสมต่อไป ดังนั้นการดำเนินโครงการพัฒนาเครื่องมือดังกล่าวจะมีความสมบูรณ์มากยิ่งขึ้น และเป็นประโยชน์ที่แท้จริงของหน่วยงานต่อไป

▣ บรรณานุกรม

กัลยา วานิชย์บัญชา. (2552). **การวิเคราะห์ข้อมูลหลายตัวแปร**. กรุงเทพมหานคร: ศูนย์หนังสือ จุฬาลงกรณ์มหาวิทยาลัย.

ณัฐวุฒิ อรินทร์. (2553). **รายงานการศึกษาติดตามผลการประเมินความเสี่ยงต่อการเกิดพฤติกรรมรุนแรง ของผู้ป่วยนิติเวชหลังการจำหน่าย** (รายงานการวิจัย). สถาบันกัลยาณ์ราชนครินทร์ กรมสุขภาพจิต.

บุญธรรม กิจปรีดาบริสุทธิ์. (2553). **คู่มือการวิจัย การเขียนรายงาน การวิจัยและวิทยานิพนธ์**. กรุงเทพฯ : โรงพิมพ์เรือนแก้ว.

บุญธรรม กิจปรีดาบริสุทธิ์. (2553). **เทคนิคการสร้างเครื่องมือรวบรวมข้อมูลสำหรับการวิจัย**. กรุงเทพฯ : ศรีอนันต์การพิมพ์.

บุญเรียง ขจรศิลป์. (2543). **วิธีวิจัยทางการศึกษา**. กรุงเทพฯ : หจก. พีเอ็นการพิมพ์.

ปัญญา ธีระวิทยเลิศ. (2548). การวิเคราะห์แบบสอบถามรายข้อ. **วารสารสมาคมนักวิจัย**, 10(2): 42-44.

ศิริชัย กาญจนวาสี. (2554). **การวิเคราะห์พหุระดับ**.กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สาลิกา โค้วบุญงาน. (2551). **ความเที่ยงและความตรงของแบบประเมินความสามารถรู้ผิดชอบของ ผู้กระทำผิดนิติเวช** (รายงานการวิจัย). สถาบันกัลยาณ์ราชนครินทร์ กรมสุขภาพจิต.

สุชีรา ภัทรายุตวรรตน์. (2545). **คู่มือการวัดทางจิตวิทยา**. กรุงเทพฯ : สำนักพิมพ์เมดิคัลมีเดีย.

สุวิมล ติรกานันท์. (2553). **การวิเคราะห์ตัวแปรพหุในงานวิจัยทางด้านสังคมศาสตร์**. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย

สุวิมล ติรกานันท์. (2551). **การสร้างเครื่องมือวัดตัวแปรในการวิจัยทางสังคมศาสตร์: แนวทางสู่การปฏิบัติ**. กรุงเทพฯ : โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย

Bryman, A. (2008). **Social Research Methods** (3 ed.): Oxford: Oxford University Press.

Kline, R.B., (2005). **Principles and Practice of Structural Equation Modeling**. New York: Guilfford Press.

Nunnally, J. C. (1978). **Psychometric Theory**. New York, McGraw-Hill.

Streiner, D. L. and G. R. Norman (1991). **Health Measurement Scale**. Oxford, Oxford University Press.

Thomas, A. (2007). Self-report data in cross-cultural research: issues of construct validity in questionnaires for quantitative research in educational leadership. **International Journal of Leadership in Education**, 10(2), 211-226