บรรณานุกรม

กาญจนา วัธนสุนทร. (2537). *การพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงทางเพศ.* วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

เกสร หว่างจิตร. (2539). *การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบสำหรับแบบสอบคัดเลือกระดับบัณฑิตศึกษาวิชาภาษาไทยและภาษาอังกฤษด้วยวิธีแมนเทล เฮลส์เซล.* วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ชนะศึก นิชานนท์. (2553). *ประสิทธิภาพของการประมาณค่าพารามิเตอร์แบบเบส์โดยใช้การสรุปอ้างอิงความน่าเชื่อถือของโมเดลการตอบสนองข้อสอบ.*วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ชัยวัช โชวเจริญสุข. (2552). *การวิเคราะห์ความเสี่ยงและผลตอบแทนของหลักทรัพย์ไทยด้วยวิธีการ Bayesian approach risk and return analysis in the stock exchange of Thailand: Bayesian approach.* เชียงใหม่: คณะเศรษฐศาสตร์มหาวิทยาลัยเชียงใหม่.

ชุติมา แสงดารารัตน์. (2545*). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนตามความรู้สึกคุ้นเคย ความรู้สึกสนใจและความรู้สึกพอใจ ในข้อสอบด้วยวิธีการตรวจสอบต่างกัน.* ปริญญานิพนธ์การศึกษามหาบัณฑิต, สาขาวิชาการวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

ญาณภัทร สีหะมงคล. (2540). *การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's^2 วิธี Raju's area measures และวิธี Closed interval area.* ปริญญานิพนธ์การศึกษาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลทางการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

นงลักษณ์ วิรัชชัย.(2542). *โมเดลลิสเรล: สถิติวิเคราะห์สำหรับการวิจัย.* กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

มหาวิทยาลัยบูรพา Burapha University

นพดล มีชั้นช่วง. (2544).*การเปรียบเทียบผลของการประมาณค่าพารามิเตอร์ ตามทฤษฎีการตอบสนองข้อสอบ ระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีฮิวริสติก และวิธีของเบย์ของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1.* วิทยานิพนธ์การศึกษามหาบัณฑิต, สาขาวิชาวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยมหาสารคาม.

นิพล พลกลาง. (2549). *ปัจจัยที่ส่งผลต่อผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์ เขต 1 ปีการศึกษา 2549.* รายงานการวิจัย กลุ่มงานวัดและประเมินผลการศึกษา กลุ่มนิเทศ ติดตามและประเมินผล การจัดการศึกษาสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์ เขต 1.

พรรณี จิตมาศ.(2540). *การเปรียบเทียบผลการวิเคราะห์ความลำเอียงข้อสอบ โดยใช้ขนาดกลุ่มผู้สอบและวิธีวิเคราะห์ต่างกัน.* ปริญญานิพนธ์การศึกษามหาบัณฑิต, สาขาวิชาการวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

รักชนก ยี่สุ่นศรี. (2544). *การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดี เอฟ ไอ ที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษาวิชาภาษาอังกฤษและวิชาคณิตศาสตร์.* วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

รัตนา ศรีเหรัญ. (2539). *การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติก3 พารามิเตอร์ ระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีฮิวริสติคและวิธีของเบส์ เมื่อขนาดของกลุ่มตัวอย่างและจำนวนข้อสอบต่างกันด้วยเทคนิคมอนติคาร์โล.* ปริญญานิพนธ์การศึกษาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

เรวดี อินทะสะระ. (2539).*ผลการตรวจสอบความลำเอียงของข้อสอบต่อการศึกษาความเที่ยงตรงเชิงพยากรณ์ของแบบทดสอบคัดเลือกที่คิดคะแนนต่างกัน.* ดุษฎีนิพนธ์การศึกษาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

วรนุช แหยมแสง. (2536). *การพัฒนากระบวนการตรวจสอบความเป็นเอกมิติของแบบสอบ.* วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

มหาวิทยาลัยบูรพา Burapha University

วลีมาศ แซ่อึ้ง. (2543). *การเปรียบเทียบอำนาจการทดสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอเนกรูประหว่างวิธีซิปเทสท์ปรับใหม่วิธีซิปเทสท์ วิธีแมนเทล – แฮนส์เซล และวิธีการถดถอยโลจิสติก*. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

วิชุดา บัวคง. (2533). *การเปรียบเทียบประสิทธิผลของวิธีประมาณค่าพารามิเตอร์ของแบบจำลองโลจิสติค 3 พารามิเตอร์ ระหว่างวิธีแมกซิมัมไลค์ลิฮูด วิธีฮิวริสติค และวิธีของเบส์ในแบบสอบวัดผลสัมฤทธิ์และแบบสอบความถนัด*. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

ศศิธร สุริยา. (2551). *ปัจจัยที่มีผลต่อการทดสอบวัดผลสัมฤทธิ์ระดับชาติของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3*. วิทยานิพนธ์ศึกษาศาสตรมหาบัณฑิต, สาขาวิชาการวัดและประเมินผลทางการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยเชียงใหม่.

ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการทดสอบแนวใหม่ (Modern test theories)* (พิมพ์ครั้งที่ 2). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_____. (2548ก). *การวิเคราะห์พหุระดับ (Multilevel analysis)* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_____. (2548ข). *ทฤษฎีการทดสอบแบบดั้งเดิม (Classical test theories)* (พิมพ์ครั้งที่ 5). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_____. (2550). *ทฤษฎีการทดสอบแนวใหม่* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สำนักงานสถิติแห่งชาติ. (2543). *สำมะโนประชากรและเคหะ พ.ศ. 2533*. กรุงเทพ: สำนักนายกรัฐมนตรี.

สำนักทดสอบการศึกษา. (2549). *แนวกำกับการสอบสำหรับสนามสอบ โครงการประเมินคุณภาพการศึกษาขั้นพื้นฐาน เพื่อประกันคุณภาพของผู้เรียน*. กรุงเทพฯ: สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ.

สำเริง บุญเรืองรัตน์. (2548). การทำหน้าที่ต่างกันของข้อสอบและแบบสอบ.*วารสารครุศาสตร์มหาวิทยาลัยราชภัฏอุตรดิตถ์*, *6*(1), 1 – 28.

มหาวิทยาลัยบูรพา Burapha University

สิริรัตน์ วิภาสศิลป์. (2545). *การเปรียบเทียบวิธีซิบเทสท์และดีเอฟไอทีในการตรวจสอบการทำหน้าที่เบี่ยงเบนของข้อสอบหมวดข้อสอบและแบบทดสอบจากข้อมูลการตอบข้อสอบที่ใช้ความสามารถหลายมิติ*. ดุษฎีนิพนธ์การศึกษาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

สุนทร เทียนงาม. (2551). *ผลของความไม่เป็นอิสระของข้อสอบที่มีต่อค่าความเที่ยง ค่าพารามิเตอร์ของข้อสอบค่าความสามารถของผู้สอบและค่าสารสนเทศของแบบสอบเมื่อมีเงื่อนไขการทดสอบที่แตกต่างกัน*. วิทยานิพนธ์ครุศาสตร์ดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

สุภมาส อังศุโชติ, สมถวิล วิจิตรวรรณา และ รัชนีกูล ภิญโญภานุวัฒน์. (2554). สถิติวิเคราะห์สำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์: เทคนิคการใช้โปรแกรม LISREL (พิมพ์ครั้งที่ 3). กรุงเทพฯ: เจริญดีมั่นคงการพิมพ์.

สุมาลี แกล้วทนงค์. (2547). *สาเหตุการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้ภาษาไทยและสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม*. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

สุวิมล ติรกานันท์. (2553). การวิเคราะห์ตัวแปรพหุในงานวิจัยทางสังคมศาสตร์. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

เสรี ชัดแช้ม. (2539). *การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันแบบไม่สม่ำเสมอของข้อสอบระหว่างวิธีแมนเทล – แฮนส์เซลแบบปกติ กับวิธีแมนเทล – แฮนส์เซลแบบแบ่งกลุ่มความสามารถของผู้สอบและความยากของข้อสอบ*. วิทยานิพนธ์ครุศาสตร์ดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

อิทธิฤทธิ์ พงษ์ปิยะรัตน์. (2551). *การวิเคราะห์ข้อสอบและการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ: การวิเคราะห์พหุระดับ*. วิทยานิพนธ์ครุศาสตร์ดุษฎีบัณฑิต, สาขาวิชาวิจัยการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

อุทัยวรรณ สายพัฒนะ. (2547). *การเปรียบเทียบประสิทธิภาพของผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ ในแบบทดสอบที่มีการให้คะแนนแบบหลายค่าระหว่างวิธี GMH และวิธี Polytomous SIBTEST*. ดุษฎีนิพนธ์การศึกษาดุษฎีบัณฑิต, สาขาวิชาการทดสอบและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

มหาวิทยาลัยบูรพา Burapha University

อุมาภรณ์ ภัทรวาณิชย์ และ ปัทมา อมรสิริสมบูรณ์. (2550). *ความไม่เท่าเทียมด้านการศึกษา: เมืองและชนบท*. นครปฐม: ประชากรและสังคม.

Adams, R. J., Wilson, M., & Wu, M. (1997). Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and Behavioral Statistics, 22*(1), 47 – 76.

Aitkin, M. A., & Longford, N. (1986). Statistical modeling in school effectiveness studies (with discussion). *Journal of the Royal Statistical Society, 149*(1), 1 – 43.

Andrich, D. (1978). Relationships between the Thurstone and Rasch approaches to item scaling. *Applied Psychological Measurement, 2*, 449 – 460.

Angoff, W. H., & Ford, S. F. (1973). Item – race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 10*(2), 95 – 106.

Angoff, W. H., & Sharon, A. L. (1972). Patterns of test and item difficulty for six foreign language groups on the test of English as a foreign language. *Research Bulletin, 17*(2), 45 – 78.

Bafurmi, J., Gelman, A., Park, D. K., & Kaplan, N. (2005). Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis, 13*, 171 – 187.

Barton, M. A., & Lord, F. M. (1981).*An upper asymptote for the three – parameter logistic item – response model.* New Jersey: Princeton.

Bock, R. D. (1997). A brief history of item response theory. *Educational Measurement: Issues and Practice*, *4*(3), 21 – 33.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika, 46*, 443 – 459.

Briggs, C. D., & Wilson, M. (2007). Generalizability in item response modeling. *Journal of Educational Measurement, 44* (2), 131 – 155.

Burstein, L. (1980). The analysis of multi – level data in educational research and evaluation. *Review of Research in Education, 8*, 158 – 233.

Camilli, A., & Shapard, L. A. (1994). *Methods for identifying biased test items.* California: SAGE.

มหาวิทยาลัยบูรพา Burapha University

Cleary, T. A., & Hillton, T. L. (1968). Comparison of logistic regression and analysis of variance differential item functioning detection methods. *Educational and Psychological Measurement, 2*(4), 29 – 35.

Cronbach, L. J. (1976). *Research on classrooms and schools: Formulation of questions, design and analysis*. New York: Wiley.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberal – ization of reliability theory. *The British Journal of Statistical Psychology, 16*, 137–163.

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the scholastic aptitude test, *Journal of Educational Measurement*, *23*(4), 355 – 368.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Erlbaum.

Feinstein, Z. S. (1995). Effect of differing item parameters on closed – interval DIF statistics. *Applied Psychological Measurement*, *19*(2), 131 – 142.

Finch, H. (2005). The MIMIC model as a method for detecting DIF: Comparison with Mantel – haenszel, SIBTEST and the IRT likelihood ratio. *Applied Psychological Measurement*, *29*(5), 278 – 295.

Fisher, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika, 48*, 3 – 26.

Fleishman, J. (2003). Using MIMIC models to assess the influence of differential item functioning. *Center for Financing, Access, and Cost Trends Agency for Healthcare Research and Quality*, *1*(2), 6.

มหาวิทยาลัยบูรพา Burapha University

French, A. W., & Miller, T. R. (1996). Logistic Regression and its use in detection differential item functioning in polytomous items. *Journal of Education Measurement, 33*, 315 – 332.

Fukuhara, H., & Kamata, A. (2007). *DIF detection in a presence of locally dependent items*. Florida: Tampa.

Geman, T., & Geman, D. (1984). Stochastic relaxation Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *3*(6), 721 – 741.

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, *8*(4), 369 – 395.

Hambleton, R. K., & Cook, L. L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, *14*(1), 75 – 96.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.

Hill, M. A., & King, E. M. (1993). Women's education in developing countries: An overview. *Women's Education in Developing*, *3*(5), 1 – 50.

Holland, P. W., & Thayer, D. T. (1988). *Differential item performance and the Mantel – haenszel procedure*. New Jersey: Erlbaum.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. New Jersey: Erlbaum.

Hulin, C. L., Drasgow, F., & Komocar, J. (1982). Applications of item response theory to analysis of attitude scale translations. *Journal of Applied Psychology*, *67*, 818 – 825.

Jensen, A. R. (1974). How biased are culture – loaded tests? *Genetic Psychology Monographs, 40*, 185 – 244.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. *Journal of Educational Measurement*, 38(1), 79 – 93.

Kim, S. H., & Cohen, A. S. (1994). A comparison of two area measures for detecting differential item functioning. *Applied Psychological Measurement*, *15*, 269 – 278.

มหาวิทยาลัยบูรพา Burapha University

Kim, W. (2003). *Development of a differential item functioning (DIF) procedure using the hierarchical generalized linear model: A comparison study with logistic regression procedure*. Doctoral Dissertation in Education Psychology, The Pennsylvania State University.

Knodel, J., & Jones G. W. (1996). Post – cairo population policy: Does promoting girls' schooling miss the mark? *Population and Development Review, 22*(4), 683 – 702.

Linacre, J. M. (1989). *Many – facet rasch measurement*. Chicago: IL MESA.

Linn, R. L., Levine, M. V., Hastings, C. N., & Wardrup, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement, 5*, 159 – 173.

Lord, F. M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement, 23*(2), 157 – 162.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Massachusetts: Addison – Wesley.

Maier, K. S. (2001). A Rasch Hierarchical Measurement Model. *Journal of Educational and Behavioral Statistics, 26*(3), 307 – 330.

Maier, N. R. F., & Casselman, G. G. (1970). Locating the difficulty in insight problems: Individual and sex differences. *Psychological Rep, 26*, 103 – 107.

Mantel, N., & Haenszel, W. (1959). Statistical procedures of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*, 719 – 748.

McCullagh, P., & Nelder, J. (1989). *Generalized linear models* (2nd Ed.). Boca Raton: Chapman and Hall.

McDonald, R. P. (1967). *Nonlinear Factor Analysis*. Virginia: Psychometric Corporation.

Mehrens, W. A., & Lehman, I. J. (1984). *Measurement and Evaluation in Education and Psychology* (3rd ed.). Tokyo: Holt Rinehart and Winston.

Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Education Statistics, 7*(2), 105 – 118.

มหาวิทยาลัยบูรพา Burapha University

Murphy, R., Charles, K., & Davidshofer, O. (2001). *Psychological testing principles and applications*. New Jersey: Prentice Hall.

Muthen, B. O. (1988). *Some uses of structural equation modeling in validity studies: Extending IRT to external variables*. New Jersey: Erlbaum.

_____. (2004). *Mplus user's guide*. Los Angeles: Author.

Muthen, B. O., Kao, C. F., & Burstein, L. (1991).Instructionally sensitive psychometrics: Application of a new IRT – based detection technique to mathematics achievement test items. *Journal of Educational Measurement*, *28*(1), 1 – 22.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus User's Guide* (5th Ed.). Los Angeles: Author.

_____. (2010*). Mplus User's Guide* (6th Ed.). Los Angeles: Author.

Narayanan, P., & Swaminathan, H. (1996). Identification of items that show nonuniform DIF. *Applied Psychological Measurement, 20*(3), 257 – 274.

Potenza, M. T., & Dorans, N. J. (1995). DIF assessment for polytomously scored items: A framework for classification and evaluation. *Applied Psychological Measurement, 19*(3), 23 – 37.

Raju, N. S. (1990). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement, 14*(2), 197 – 207.

Raju, N. S., Drasgow, F., & Slinde, J. A. (1993). An empirical comparison of the area methods, Lord's chi – square test, and the Mantel – haenzel technique for assessing differential item functioning. *Educational and Psychological Measurement, 53*(2), 301 – 314.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Chicago: MESA.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, *59*, 1 – 17.

_____. (2002). *Hierarchical linear models: An applications and data analysis methods* (2th ed.). California: Sage.

มหาวิทยาลัยบูรพา Burapha University

Rijman, T., Meulders, S., & Balazs, D. (2005). Mixed model estimation methods for the Rasch model. *Journal of Applied Measurement, 6*(3), 237 – 288.

Rudner, L. M. (1977). Biased item detecting technique. *Journal of Educational Statistics, 5*, 213 – 233.

Rudner, L. M., Getton, T., & Knight, D. L. (1980). Biased item detection techniques. *Journal of Educational Statistics, 5*(5), 213 – 233.

Saengla Chaimongkol, Fred, W. H., & Akihito Kamata. (2007). An explanatory differential item functioning (DIF) model by the WinBUG 1.4. *Songklanakarin J. Sci. Technol, 29*(2), 449 – 458.

Schultz, T. P. (1993). *Returns to Women's Education.* Baltimore: John's Hopkins University.

Shealy, R., & Stout, W. F. (1993). *An item response theory model for test bias.* New Jersey: Erlbaum.

Shepard, L. A., Camilli, G., & Williams, D. M. (1984). Accounting for statistical artifacts in item bias research. *Journal of Educational Statistics, 9*, 93 – 128.

Swaminathan, H., & Gifford, J. (1985). Bayesian estimation in the two – parameter logistic model. *Psychometrika, 50*, 349 – 364.

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*(4), 361 – 370.

Teresi, J. A. (2006). Overview of quantitative measurement methods: Equivalence, invariance and differential item functioning in health applications. *Medical Care, 44*(11), 39 – 49.

Thissen, D. (1991). *MULTILOG user's guide: Multiple categorical item analysis and test coring using item response theory.* Chicago: Scientific Software.

Thissen, D., Steinberg, L., & Gerrand, M. (1986). Beyond group – mean differences: The concept of item bias. *Psychological Bulletin, 99*, 118 – 128.

Thissen, D., Steinberg, L., & Wainer, H. (1993). *Detection of differential item functioning using the parameters of item response models.* New Jersey: Erlbaum.

มหาวิทยาลัยบูรพา Burapha University

Thurstone, L. L. (1927). A mental unit of measurement. *Psychological Review*, *34*, 415 – 423.

Vaughn, B. (2006). *A hierarchical generalized linear model of random differential item functioning for polytomous items: A bayesian multilevel approach.* Tallahassee: Florida State University,

Welch, C. J., & Miller, T. R. (1995). Assessing differential item functioning in direct writing assessment: Problem and example. *Journal of Evolutional Measurement, 32*(2), 163 – 178.

Zumbo, B. D. (1999). *A Handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert – type (ordinal) item scores.* Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Zwinderman, A. H. (1991). A generalized Rasch model for manifest predictors. *Psychometrika*, *56*, 589 – 600.

มหาวิทยาลัยบูรพา Burapha University