

บรรณานุกรม

- กาญจนา วัชรสุนทร. (2537). การพัฒนาเกณฑ์ตัดสินข้อสอบลำเอียงทางเพศ. วิทยานิพนธ์ปริญญาครุศาสตรดุษฎีบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.
- จิตติมา วรรณศรี. (2539). การเปรียบเทียบประสิทธิภาพในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบด้วยวิธีแมนเทิล - แฮนส์เซล กับวิธีชิปเทสต์ เมื่อความยาวของแบบสอบ ขนาดกลุ่มตัวอย่าง และอัตราส่วนระหว่างกลุ่มอ้างอิงและกลุ่มเปรียบเทียบต่างกัน. วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.
- ชุติมา แสงคารารัตน์. (2546). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของแบบทดสอบวัดความถนัดทางการเรียนตามความรู้สึกรู้สึกสนใจ ความรู้สึกสนใจ และความรู้สึกพอใจในข้อสอบ ด้วยวิธีตรวจสอบแตกต่างกัน. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- เชิญ สามารถ. (2547). สถิติสำหรับการวิจัยทางการศึกษา. สุรินทร์: สถาบันราชภัฏสุรินทร์.
- ญาณภัทร สีหะมงคล. (2540). การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Raju's Area Measures และวิธี Closed Interval Area. วิทยานิพนธ์ปริญญาการศึกษาดุษฎีบัณฑิต, สาขาวิชาการศึกษาและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- นพมาศ พิพัฒน์สุข. (2541). การเปรียบเทียบประสิทธิภาพระหว่างวิธีแมนเทิล-แฮนส์เซล กับวิธีถดถอยโลจิสติก ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ เมื่อใช้เกณฑ์จับคู่เปรียบเทียบแตกต่างกันในแบบสอบชนิดพหุมิติ. วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.
- นุชริน ใบโพธิ์. (2544). การเปรียบเทียบความสอดคล้องของผลการตรวจสอบข้อสอบที่ทำหน้าที่ต่างกันระหว่างวิธี Lord's χ^2 วิธี Mantel-Haenszel และวิธี SIBTEST. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

ปิยะทิพย์ ดินวร. (2549). การตรวจสอบการทำหน้าที่ต่างกันของข้อสอบในแบบสอบพหุมติ: การเปรียบเทียบประสิทธิภาพระหว่างวิธีการวิเคราะห์องค์ประกอบจำกัดกับวิธีถดถอยโลจิสติก. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาเทคโนโลยีการวัดทางการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยบูรพา.

พรรณี จินตมาศ. (2540). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบ โดยใช้ขนาดกลุ่มผู้สอบและวิธีวิเคราะห์ต่างกัน. วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต, สาขาการวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.

เขาวดี วิบูลย์ศรี. (2545). การวัดผลและการสร้างแบบทดสอบวัดผลสัมฤทธิ์ (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย

รักษนก ยี่สุนศรี. (2544). การวิเคราะห์การทำหน้าที่ต่างกันของข้อสอบและแบบสอบด้วยกระบวนการ ดีเอฟไอที สำหรับแบบสอบคัดเลือกบุคคลเข้าศึกษาในสถาบันอุดมศึกษา วิชาภาษาอังกฤษและวิชาคณิตศาสตร์. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาการศึกษาและประเมินผลการศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.

วลีมาศ แซ่เอ็ง. (2543). การเปรียบเทียบอำนาจทดสอบและอัตราความคลาดเคลื่อนประเภทที่ 1 ในการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบแบบอนกรูระหว่างวิธีซิทส์ปรับปรุงใหม่ วิธีแมนเทล – แชนส์เซลล์และวิธีการถดถอยโลจิสติก. วารสารวิธีวิทยาการวิจัย, 14(2), 216-230.

ศิริชัย กาญจนวาสี. (2545ข). ทฤษฎีการทดสอบแบบดั้งเดิม (พิมพ์ครั้งที่ 2). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_____. (2548). ทฤษฎีการทดสอบแบบดั้งเดิม (พิมพ์ครั้งที่ 5). กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

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

ศุภวัฒน์ มะลิเผือก. (2548). อิทธิพลของการทำหน้าที่ต่างกันของข้อสอบ ที่ส่งผลต่อคุณภาพของแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนระดับชาติ ชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิชาเทคโนโลยีการวัดทางการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยบูรพา.

สมนึก ภัททิยธนี. (2544). การวัดผลการศึกษา (พิมพ์ครั้งที่ 3). กรุงเทพฯ: ประสานการพิมพ์.

สมศักดิ์ จันทอง. (2541). จากความลำเอียงของข้อสอบ(Item Bias) มาเป็นการทำหน้าที่ต่างกันของข้อสอบ(DIF). วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร, 3(1), 81 – 93.

- สุมาลี แก้วทนต์. (2547). สาเหตุของการทำหน้าที่ต่างกันของข้อสอบสาระการเรียนรู้
ภาษาไทยและสาระการเรียนรู้สังคมศึกษา ศาสนา และวัฒนธรรม. วันที่ค้นข้อมูล
11 ตุลาคม 2552, เข้าถึงได้จาก <http://cuir.car.chula.ac.th/handle/123456789/3912>
- สุรศักดิ์ อมรรदनศักดิ์. (2531). การเปรียบเทียบผลการวิเคราะห์ความลำเอียงของข้อสอบที่แตกต่างกัน
4 วิธี. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต, สาขาวิชาการวัดผลและประเมินผล
การศึกษา, บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.
- สิริรัตน์ วิชาศิลป์. (2545). การเปรียบเทียบวิธีชิปเทสต์และดีเอฟไอที ในการตรวจสอบการทำ
หน้าที่เบี่ยงเบนของข้อสอบ หมวดข้อสอบและแบบทดสอบ จากข้อมูลการตอบข้อสอบ
ที่ใช้ความสามารถหลายมิติ. วิทยานิพนธ์ปริญญาการศึกษาดุษฎีบัณฑิต, สาขาวิชา
การทดสอบและวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- สำเร็จ บุญเรืองรัตน์. (2548). ความลำเอียงของเครื่องมือทดสอบ. สารานุกรมศึกษาศาสตร์ ฉบับที่
34. มหาวิทยาลัยศรีนครินทรวิโรฒ .
- เสรี ชัดเข้ม. (2540). วิธีการทางสถิติที่ใช้ตรวจสอบข้อสอบทำหน้าที่ต่างกัน. วารสารมหาวิทยาลัย
บูรพา, 2(1), 41-53 .
- _____. (2544). เอกสารประกอบการสอนรายวิชาการวัดผลการศึกษา. ชลบุรี: ภาควิชาวิจัยและ
วัดผลการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา.
- อารี วัชรโสติกุล. (2543). การเปรียบเทียบผลการตรวจสอบการทำหน้าที่ต่างกันของข้อสอบ
โดยใช้รูปแบบและวิธีการต่าง . วิทยานิพนธ์ปริญญาการศึกษามหาบัณฑิต, สาขาวิชา
การวัดผลการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- Angoff, W. H. (1993). Perspective on differential item functioning methodology. In P. W.
Holland, and H. Wainer (Eds), *Differential Item Functioning*, 3-23. Hillsdale, NJ:
Lawrence Erlbaum Associates, Inc.
- Angoff, W. H., & Ford, S.F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal
of Educational Measurement*, 10, 95-105.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. London:
Sage Publication.
- Chang, H. H., Mazzeo, J., & Roussos, L. (1996). Detecting DIF for polytomously scored items:
An adaptation of the SIBTEST procedure. *Journal of Educational Measurement*, 33(3),
333-353.

- Cleary, T. A., & Hilton, T. L. (1968). An investigation of item bias. *Education and Psychological Measurement*, 28, 61-75.
- Cronbach, L., & Furby, L. (1970). How should we measure change – Or should we?. *Psychological Bulletin*, 74, 68-80.
- David Scrams, J., & Lori Mcleod, D.. (2000). An expected response function approach to graphical differentiation item function. *Journal of Educational Measurement*, 37(3), 263-280.
- Dimensionality-Based DIF Package (Version 1.7)* [Computer software]. The Roussos-Stout Software Development Group.
- Dorans, N. J., & Kulick, E. M. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the scholastic aptitude test. *Journal of Educational Measurement*, 23, 355-368.
- Feinstein, Z. S. (1995). Effect of differing item parameters on closed-interval DIF statistics. *Applied Psychological Measurement*, 19(2), 131-142.
- Holland, P.W., Longford, N.T., & Thayer, D.T. (1989). *Stability of the MH D-DIF statistics across populations* (PRPC Report). Princeton, NJ: Educational Testing Service.
- Holland, P. W., & Thayer, D.T. (1988). Differential item performance and the Mantel-Haenszel procedure. In Wainer, H., & Braun, H. I. (eds.), *Test Validity*, 129-145. NJ: Lawrence Erlbaum Associates.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. Homewood, I.L.: Dow Jones – Irwin.
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement?. *Harvard Educational Review*, 39, 1-123.
- _____. (1974). How biased are culture-loaded tests ?. *Genetic Psychology Monographs*, 90, 185-244.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Kederman, H. (1990). Item bias detection using loglinear IRT. *Psychometrika*, 54, 681-697.

- Kim, S. H., & Cohen, A.S. (1991). A comparison of two area measures for detecting differential item functioning. *Applied Psychological Measurement*, 15(3), 269-278.
- Linn, R. L., Levine, M.V., Hastings, C.N., & Wardrop, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement*, 5, 159-173.
- Linn, R. L. (1993). The use of differential item functioning statistics: A discussion of current practice and future implementation. In P. W., Holland, & H., Wainer, (Eds), *Differential item functioning* (pp. 349 - 364). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Master, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Mellenbergh, G.J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7, 105-118.
- Millsap, R. E., & Everson, H. T.. (1993, February). Methodology Review: Statistical Approaches for Assessing Measurement Bias. *Applied Psychological Measurement*, 17 (1), 297-334.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16, 159-176.
- _____. (1993). Information functions of the generalized partial credit model. *Applied Psychological Measurement*, 17, 351-363.
- Nandakumar, R. (1993). Assessing essential dimension of real data. *Applied Psychological Measurement*, 17, 29-38.
- Narayanan, P., & Swaminathan, H. (1994). Performance of the Mantel-Haenszel and simultaneous item bias procedure for detecting differential item functioning. *Applied Psychological Measurement*, 18(4), 315-328.
- _____. (1996). Identification of item that show nonuniform DIF. *Applied Psychological Measurement*, 20(3), 257-274.
- Popham, W. J. (1981). *Modern Educational Measurement*. Englewood Cliffs, NJ: Prentice – Hall.

- Potenza, M. T., & Dorans, N. J. (1995). DIF assessment for polytomous score items: A framework for classification and evaluation. *Applied Psychological Measurement*, 19(1), 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., Vander, Linden, W. J., & Fleer, P.F. (1995). IRT – based internal measures of differential functioning of item and tests. *Applied Psychological Measurement*, 19(4), 353-368.
- Rudner, L. M., Getson, P. R., & Knight, D. L. (1980). Biased item detection techniques. *Journal of Educational Statistics*, 5, 213-233.
- Scheuneman, J. D. (1979). A method for assessing bias in test items. *Journal of Educational Measurement*, 16, 143-152.
- Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability difference and detect test / DIF as well as item bias/ DIF. *Psychometrika*, 58(2)(June 1993), 159-194.
- Shepard, L. A., Camilli, G., & Williams, D. M. (1984). Accounting for statistical artifacts in item bias research. *Journal of Educational Statistics*, 9, 93-128.
- Stout, W. F., Hsin-Hung, L. Nandakumar, R., & Bolt, D. (1997). MULTISIB: A procedure to investigate DIF when a test is intentionally two-dimensional. *Applied Psychological Measurement*, 21, 195-213.
- Swaminathan, H., & Roger, H. J. (1990). Detecting differential item functioning using logistic regression procedure. *Journal of Education Measurement*, 27(4), 361-370.
- Dimensionality-Based DIF Package (Version 1.7)* [Computer software]. The Roussos-Stout Software Development Group.
- Thissen, D. M., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P.W. Holland & H. wainer (Eds). *Differential item functioning: Theory and practice*, (pp.67-113). Hillsdale, NJ: Lawrence Erlbaum.