

= =

“ 가 가 ”

( )

1.

(1998) 가 [ 1]

가

Mislevy(1994)가

(intelligent measurement)

Bunderson, Inouye Olsen(1989)

가

Embretson(1983, 1985)

가

(multi-dimensional item response theory)

가

가

가

가

가

가

## 2. 가

가 (academic achievement test) 가

가 (instruction-driven measurement)

가 (measurement-driven instruction) Tyler 가 Popham (1987)

(norm-referenced test)가

(criterion-referenced test)가

[ ]

가

가

가

가

가

가,

가

가

가

가  
Z, T, SAT GRE , , Staint

(moderated score) Oxford  
. Mislevy(1992), Linn(1993, 1995), Wilson(1992),  
Keeves(1988)

가  
(Milsevy Stocking, 1992),  
가  
(social moderation), (consensus moderation; Bell,  
Burkhardt, & Swan, 1991)

가

가

가

가

(criterion, cut off, standard) 가  
(AERA/APA/NCME, 1985)

[ ] 가  
Angoff(1971)

Jaeger(1978) . Angoff 가 100  
가 가 가  
가 가 . Jaeger 가 가  
가 가

가  
가

가 ( ‘ ’, ‘ ’, ‘ ’  
’, ‘ ’)

(Meijer, 1996; Tatsuoka, 1983, 1996; , 1997; , 1998; , 1997)

가

가

가

가

### 3.

가

가

가

가

가

가

가

가

가

가

가

(Gronlund Linn, 1990; ,

1995).

가

Cureton(1951)

가

가

Cronbach(1971)

가 . Messick(1989) 1970 가  
 가  
 (consequential validity) .  
 Shepard(1997)  
 , 가  
 . Linn(1997) , Popham(1997)  
 가 가 ,  
 , 가  
 , , , 가 가  
 , 가 가 가

- (1998). 가 가 . 가
- (1997). : .  
가 . 10 2 . 139-176.
- (1991). , 29(2), 147-164.
- (1995). : .
- (1998, ). Rasch 가  
. 가 .
- (1997). .

AERA/APA/NCME(1985). Standard for Educational and Psychological Testing.  
Washington D. C. : American Psychological Association.

Angoff, W. H.(1971). Scale, norms and equivalent score. In R. L. Thorndike(Ed.)  
Educational Measurement(2nd ed.), Washington D. C. : American Council on  
Education, 508-600.

Bell, A., Burkhardt, H., & Swan, M.(1991). Balanced assessment and the  
mathematics curriculum. Nottingham, England: Shell Centre for Mathematical  
Education.

Berk, R. A.(1986). A consumer's guide to setting performance standards on  
criterion-referenced tests. Review of Educational Research, 56, 137-172.

Bunderson, C. V., Inouye, d. K., & Olsen, J. B.(1989). The four generations of  
computerized educational measurement. In R. L. Linn(Ed.), Educational  
Measurement. New York: American Council on Education and Macmillan.

Cronbach, L. J.(1971). Test Validation. In R. L. Thorndike(Ed.) Educational  
Measurement(2nd ed.), Washington D. C. : American Council on Education,  
443-507.

Cureton, E. E.(1951). Validity. In E. F, Lindquist(Ed.), Educational Measurement,  
Washington D. C. : American Council on Education. 621-695.

- de Gruijter, D. N. M.(1985). Compromise models for establishing standards. *Journal of Educational Measurement*, 22, 263-269.
- Ebel, R. L.(1979). *Essentials of educational measurement*(3rd ed). Englewood Cliffs, NJ: Prentice-Hall, 337-342.
- Embretson, S. E.(1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93, 179-197.
- Embreston, S. E.(1985). Multicomponent latent trait models for test design. In S. E. Embretson(Ed.), *Test Design*, New York, N.Y.: Academic Press, Inc, 195-218.
- Gronlund, N. E. & Linn, R. L. (1990). *Measurement and Evaluation in teaching*(6th ed.). New York: Macmillan, 48-50.
- Hofstee, W. K. B.(1980). Policies of educational selection and grading: The case for compromise models. Paper presented at the Fourth International Symposium in Educational Testing, Antwerp, Belgium.
- Hofstee, W. K. B.(1983). The case for compromise in educational selection and grading. In S. B. Anderson & J. S. Helmick(Eds.), *On educational testing*, Sanfransisco: Jossey-Bas, 109-127.
- Jaeger, R. M.(1978). A proposal for setting a standard on the North Carolina High School Competency test. Paper presented at the spring meeting of the North Carolina Association for Research in Education, Chapel Hill.
- Keeves, J. P.(1988). Scaling Achievement Test Scores. In T. Husen & T. N. Postlethwaite(Eds.) *International Encyclopedia of Education*, Oxford: Pergamon Press, 403-419.
- Linn, R. L.(1993). Linking Results of Distinct Assessments. *Applied Measurement in Education*, 6(1), 83-102.
- Linn, R. L.(1995). Linking Assessments, In M. B. Kane & R. Mitshcell(Eds.), *Implementing Performance Assessment.*, New Jersey: Lawrence Erlbaum Associates, 91-105.
- Linn, R. L.(1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement : Issue and Practice*, 16(2), 14-15.
- Livingston, S. A. & Zieky, M. J.(1982) *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Princeton,

- NJ: Educational Testing Service.
- Meijer, R. R.(1996). Person-fit research: An Introduction Applied Measurement in Education, 9(1). 3-8.
- Messick, S.(1989). Validity, In R. L. Linn(Ed.) Educational Measurement(3rd ed.), Washington D. C. : American Council on Education & National Council on Measurement in Education, 13- 103.
- Mislevy, R. M.(1992). Linking Educational Assessments: Concepts, Issues, Methods, and Prospects. Princeton, NJ: Educational Testing Service.
- Mislevy, R. M.(1994). Evidence and Inference in educational assessment. Psychometrika, 59, 4, 439-483.
- Mislevy, R. M. & Stocking, M. S.(1992). Memorandum for the record, subject calibration. Internal ETS memorandum. Princeton, NJ: Educational Testing Service(cited with permission)
- Nedelsky, L.(1954). Absolute grading standards for objective test. Educational and Psychological Measurement, 14, 3- 19.
- Popham, W. J.(1987). The merits of measurement-driven instruction. Phi Delta Kappan, 68(9), 679-682.
- Popham, W. J.(1997). Consequential validity: right concern-wrong concept. Educational Measurement : Issue and Practice, 16(2), 9- 13.
- Shepard, L. A.(1997). The Centrality of test use of consequences for test validity. Educational Measurement : Issue and Practice, 16(2). 5-8.
- Tatsuoka, K. K.(1983). Rule space : Approach for dealing with misconceptions based on item response theory. Journal of Educational Measurement, 20(4). 345-354.
- Tatsuoka, K. K.(1996). Use of Generalized Person-Fit Indexed, Zetas for Statistical Pattern Classification. Applied Measurement In Education, 9(1), 65-75.
- Wilson, M.(1992). The integration of school based assessments into a state-wide assessment system: Historical perspectives and contemporary issues. Unpublished manuscript, University of California, Berkeley.
- Zieky, M. J. & Livingston, S. A.(1977). Manual for setting standards in the basic skill assessment test. Princeton, NJ: Educational Testing Service.