



The LAS: Validity and reliability

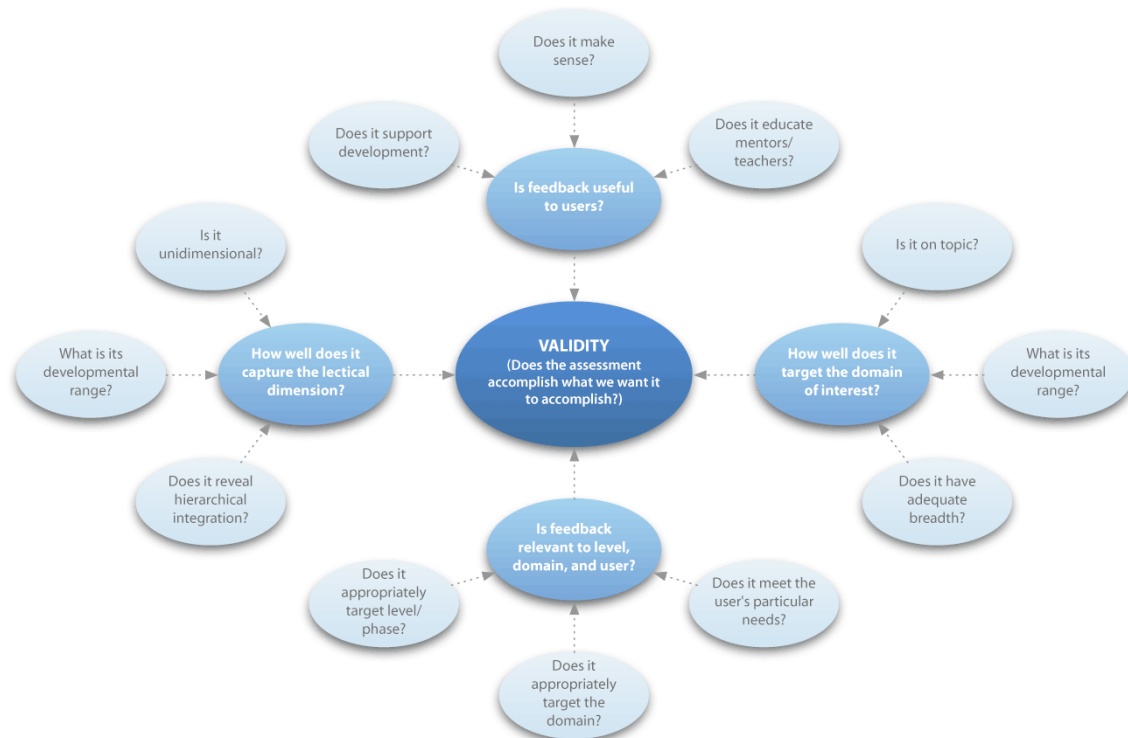
Prepared by Dr. Theo L. Dawson¹, Developmental Testing Service, LLC

The LAS: Validity and reliability

DTS meets or exceeds the validity and reliability standards for educational and psychological testing set jointly by the APA, AERA, and NCME.¹

Validity

We focus on two sources of validity for the Lectual™ Assessment System (LAS), both of which represent forms of construct validity²—the degree to which the LAS measures what it is supposed to measure. These are (1) the extent to which the LAS measures the same dimension as other cognitive developmental assessment systems, and (2) the degree to which patterns of performance on the LAS reflect expectations for unidimensionality and wave-like growth that have been postulated by cognitive developmental theorists.



¹ <http://www.apa.org/science/standards.html>

² Construct validity is considered by many to be the most important form of validity. See: Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35(11), 1012-1027.

Reliability

We evaluate the reliability of the LAS in two ways: (1) with a psychometric tool called Rasch analysis, which produces an index of internal consistency that is similar to Cronbach's alpha, and (2) by examining inter-rater reliability.

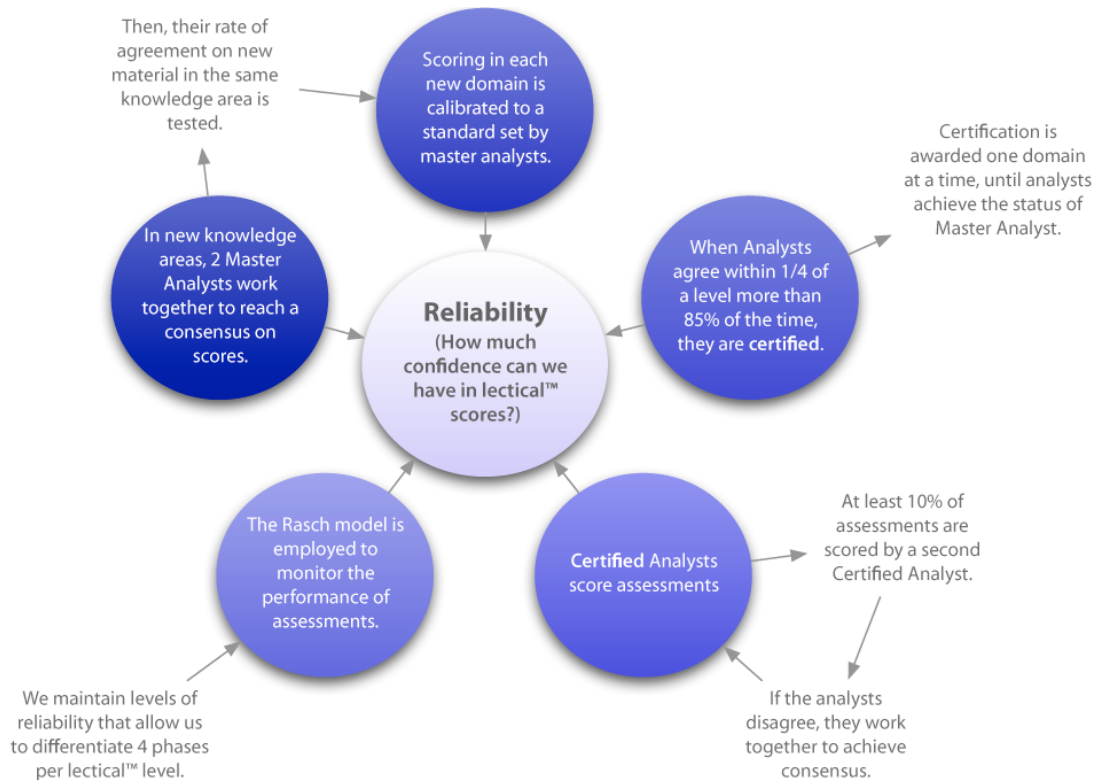


Table 1: Reliability and validity of the LAS

Findings	References
Psychometric modeling and qualitative analyses reveal that the LAS and several longitudinally validated domain specific cognitive developmental assessment systems assess the same dimension of performance.	Dawson, T. L. (2000). Moral reasoning and evaluative reasoning about the good life. <i>Journal of Applied Measurement</i>, 1, 372-397. Dawson, T. L. (2002). A comparison of three developmental stage scoring systems. <i>Journal of Applied Measurement</i>, 3, 146-189. Dawson, T. L. (2003). A stage is a stage is a stage: A direct comparison of two scoring systems. <i>Journal of Genetic Psychology</i>, 164, 335-364. Dawson, T. L., & Gabrielian, S. (2003). Developing conceptions of authority and contract across the life-span: Two perspectives. <i>Developmental Review</i>, 23, 162-218. Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? <i>Cognitive Development</i>, 18, 61-78. Dawson, T. L. (2004). Assessing intellectual development: Three approaches, one sequence. <i>Journal of Adult Development</i>, 11, 71-85.
Rasch modeling shows that the LAS measures a unidimensional "latent dimension."	Dawson-Tunik, T. L. (2004). "A good education is..." The development of evaluative thought across the life-span. <i>Genetic, Social, and General Psychology Monographs</i>, 130, 4-112. Dawson-Tunik, T. L., Commons, M. L., Wilson, M., & Fischer, K. W. (2005). The shape of development. <i>The European Journal of Developmental Psychology</i>, 2, 163-196.
Rasch modeling shows that development along the latent dimension measured by the LAS is wave-like, a pattern that is consistent with the cognitive developmental postulate that development is characterized by a series of nested, hierarchical reorganizations of knowledge structures (rather than the simple accumulation of knowledge).	Xie, Y., & Dawson, T. L. (2006). Multidimensional models in a developmental context. In M. Garner, G. Engelhard, M. Wilson & W. Fisher (Eds.), <i>Advances in Rasch Measurement</i>: JAM Press. Dawson-Tunik, T. L. (2004). "A good education is..." The development of evaluative thought across the life-span. <i>Genetic, Social, and General Psychology Monographs</i>, 130, 4-112. Dawson-Tunik, T. L., Commons, M. L., Wilson, M., & Fischer, K. W. (2005). The shape of development. <i>The European Journal of Developmental Psychology</i>, 2, 163-196. Dawson-Tunik, T. L. (2005, June). <i>Cognitive change is stage-like: The cumulative evidence from a decade of Rasch modeling</i>. Paper presented at the Annual Meeting of the Jean Piaget Society, Vancouver. Dawson, T. L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W. Boone (Eds.), <i>Applications of Rasch measurement in science education</i> (pp. 111-136). Maple Grove, MN: JAM Press.

Findings	References
The internal consistency of the LAS consistently has been above .90. (As of 2009, we are maintaining alphas of .95 and above.)	Dawson, T. L. (2000). Moral reasoning and evaluative reasoning about the good life. <i>Journal of Applied Measurement</i>, 1, 372-397. Dawson, T. L. (2002). A comparison of three developmental stage scoring systems. <i>Journal of Applied Measurement</i>, 3, 146-189. Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? <i>Cognitive Development</i>, 18, 61-78. Dawson-Tunik, T. L. (2004). "A good education is..."The development of evaluative thought across the life-span. <i>Genetic, Social, and General Psychology Monographs</i>, 130, 4-112. Dawson-Tunik, T. L., Commons, M. L., Wilson, M., & Fischer, K. W. (2005). The shape of development. <i>The European Journal of Developmental Psychology</i>, 2, 163-196.
Inter-rater reliability for the LAS consistently has been above 85% agreement within 1/3 of a lectical level. (As of 2007, we maintain an inter-rater agreement rate of 85% within 1/4 of a level.)	Dawson-Tunik, T. L. (2004). "A good education is..."The development of evaluative thought across the life-span. <i>Genetic, Social, and General Psychology Monographs</i>, 130, 4-112. Dawson-Tunik, T. L., Commons, M. L., Wilson, M., & Fischer, K. W. (2005). The shape of development. <i>The European Journal of Developmental Psychology</i>, 2, 163-196. Dawson, T. L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W. Boone (Eds.), <i>Applications of Rasch measurement in science education</i> (pp. 111-136). Maple Grove, MN: JAM Press.

Table 2: Ongoing evaluation of the reliability of the LAS and Lectical™ Assessments

Goal	Approach
Maintaining inter-rater reliability rates at or above 85% within 1/4 of a lectical level.	Approximately 20% of all Lectical Assessments are scored by at least two Certified Analysts, whose scores are regularly compared. If agreement rates drop below 85% within 1/4 of a lectical level, we conduct a scoring review, and analysts work together to reconcile scores that differ by more than 1/4 of a level.
Maintaining internal consistency (reliability) levels above .90.	All Lectical Assessments require analysts to make 5 or more different scoring decisions. This number is usually adequate for maintaining the desired level of reliability.
Continuously upgrading descriptions of the conceptual content associated with particular developmental phases.	The scoring interface we employ to score Lectical Assessments requires analysts to continuously monitor and comment upon the accuracy of phase descriptions, which are modified as required. We periodically conduct formal studies of the relation between developmental level and conceptual content, employing the methods of <i>developmental maieutics</i>.
Ensuring that new Lectical Assessments meet our requirements for accuracy, reliability, and validity.	To develop a new Lectical Assessment, we study development in the targeted domain (usually in a life-span sample—ages 5 through adulthood), and use the methods of <i>developmental maieutics</i> to construct detailed developmental pathways for the domain. These form the basis for report feedback and suggestions. We also evaluate the psychometrics of the assessment with the Rasch model. We endeavor to submit the results of all our research and psychometric evaluations to peer review. This not only establishes the credibility of our assessments, it “open sources” some of what we learn about conceptual development.