After learning to read in primary school, students progressively use their reading skills for the acquisition of new knowledge. However, textbooks as well as written material in everyday life contain besides verbal information also pictorial information, which requires besides verbal literacy also visual literacy. Although the combination of visual and verbal literacy is a widespread requirement in schooling and in society, little is known about the development of these skills. The BITE (Bild-Text-Integration) project supported by the German Science Foundation, carried out by the University of Koblenz-Landau and the Max-Planck-Institute for Human Development, investigates this development through two longitudinal studies. Based on a cognitive theoretical model, test items for different hierarchical levels of text-picture integration were developed and evaluated with students from grade 5 to 8 of three different levels of schools in order to develop Rasch-homogenous performance scales. In order to identify structural characteristics and procedural requirements influencing item difficulty, rational task analyses were performed for all test units and used to predict item parameters. Results are now subject to further cognitive studies on items processing with the help of eye-movement analyses. A point is made that assessment studies should combine psychometric and experimental methods.