Random effects growth curve models provide a powerful method for estimating individual trajectories to test hypotheses about the development of psychopathology over time. However, a salient challenge encountered in many studies of high-risk development is the reliance on categorical repeated measures that have been assessed over an extended developmental period. A common analytical approach in this situation is to fit a growth model to repeated measures scored as the proportion of endorsed items. Although sometimes advantageous, the proportion score approach imposes several strict and untestable assumptions that likely do not hold in practice. Two alternative and underutilized categorical measurement models are item response theory (IRT) and categorical confirmatory factor analysis (CCFA). Despite several promising advantages, little is currently known about the experimental conditions under which each method performs optimally. To gain a better understanding of these important issues, we first examine the analytical model and associated assumptions underlying the proportion score, IRT, and CCFA frameworks. We then apply each of the three methods to empirical data drawn from a large longitudinal study of problem behavior over a 10 year period. We conclude with initial recommendations for the use of these methods in practice and propose potential directions for future quantitative research.