Examples of published studies relaxing the *p*< .05 statistical significance standard

Lee and Bryk (1989) used HLM on a subset of data from the High School and Beyond Survey (*df*=137), and retained an effect with *p* = .079 because there was substantial prior evidence in the literacy regarding this effect and the model was theory based.

Lee, V., & Bryk, A. (1989). A multilevel model of the social distribution of educational achievement. *Sociology of Education, 62*, 172-192.

Lee and Smith (1997) retained and discussed explanatory variables that achieved a significance level of *p*< .10, as they tested several theory based models and found consistency in effect size magnitudes across models with the specified set of explanatory variables. They focused on effect sizes for relevant variables rather than a significance standard for exploring the issues related to school size and gains in achievement. Their interpretation of effect sizes was as follows: 0.1 - 0.3 as small, 0.3 - 0.5 as moderate, and 0.5 SD or more as large; following Rosenthal and Rosnow (1984).

Lee, V.E., & Smith, J.B. (1997). High school size: Which works best and for whom? *Educational Evaluation and Policy Analysis, 19*, 205-227.

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

There are other examples of the inclusion of variables because they simultaneously explain additional variance in a particular level although they are not individually statistically significant at the .05 level (e.g., Simpson, Raudenbush, & Earls, 1997).

Simpson, R.J., Raudenbush, S.W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*, 918-924.

In his analysis of education spending and achievement, Wenglinsky (ETS), retained effects at the *p*< .10 level as significant when they contributed to the variance explained (e.g., his inclusion of the urban location indicator in his model of achievement achieved a *p*= .10 and was retained and supported). “Significance tests are not designed to confirm the lack of a relationship but, rather, to confirm the existence of one. In the case of a null finding, it may be that under slightly different circumstances (such as a larger sample size or different statistical controls) a relationship could be discerned” (p. 278).

Wenglinsky, H. (1998). Finance equalization and within-school equity: The relationship between education spending and the social distribution of achievement. *Educational Evaluation and Policy Analysis, 20*, 269-283.

Mayer (Mathematica Policy Research Inc.), in his study of implementation of the NCTM teaching approach and its effect on student performance in an Algebra growth model, retained variables that achieved *p*< .10 (e.g., NCTM teaching approach, the primary variable of interest; and students’ prior GPA) as “marginally significant” but since they were model and theory based they were important to retain and investigate further.

Mayer, D.P. (1998). Do new teaching standards undermine performance on old tests? *Educational Evaluation and Policy Analysis, 20*, 53-73.

Raudenbush and Bryk (2002) used HLM to study growth in Natural Science knowledge in a Head Start program (n=134) and discussed two nonsignificant effects as “plausible” because their magnitude and direction relayed useful and defensible information about the effects of native English language status and hours of instruction, as expected, and because they explained substantial variation (55%).

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rodriguez et al. (2005) participated in an editorial discussion with the editor and revieweres of *Reading Research Quarterly*, defending their use of the *p*< .10 as a meaningful and appropriate criterion for statistical significance in their 2005 article on the CIERA school change framework and the identification of teacher practices associated with reading improvement.

Rodriguez, M.C., Taylor, B.M., Pearson, P.D., & Peterson, D. (2005). Revealing an exchange between authors and reviewers about statistical significance [Editorial]. *Reading Research Quarterly*, *40*(1), 9-10.

Taylor, B.M., Pearson, P.D., Peterson, D., & Rodriguez, M.C. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly, 40*(1), 40-69.