GADO, ISSAOU, Ph.D., December 2004 Educational Foundations and Special Services

PSYCHOMETRIC PROPERTIES OF TEACHER-MADE SCIENCE TESTS USED IN NATIONAL EXAMINATIONS FOR MIDDLE GRADE STUDENTS IN BENIN (WEST AFRICA): A LONGITUDINAL STUDY (223 PP.)

Co-Directors of Dissertation: Shawn Fitzgerarld, Ph.D.
Rafa Kasim, Ph.D.

The purpose of this study was to determine the psychometric properties (item difficulty, item discrimination, internal consistency reliability, content validity and construct validity) of teacher-made science tests used in national examinations for Middle Grade students in Benin (West Africa) for three consecutive years. The study also described the assessment methods used in science classrooms. Research data were collected from two sources: first, a survey questionnaire administered to 250 secondary school teachers purposively selected; second, a total of 630 graded physical science copies for three consecutive years of national examinations randomly selected from the Service of Examination and Testing data sources. Descriptive statistics were used to analyze the survey questionnaire. Factor analysis was used to explore construct validity of the measurements. Classical test theory methods were used to explore item difficulty, item discrimination, and reliability of examination scores. Generalizability theory provided estimate of variance components and proportions of total variance accounted for by sources of error related to persons, items, and person-by-item interaction. The result of this study shows that teacher-made tests used in large scale high-stakes examination for three consecutive years are highly reliable and have a satisfactory level of difficulty and discrimination. However, even though the items of teacher-made tests are associated

with the objectives of the national curriculum standards, the proportion of objectives tested in national examinations and the number of items across three consecutive years show a non uniform and inconsistent distribution of items across years, content domains, and within the fields of chemistry and physics. Therefore, teacher-made tests used in national examinations for three consecutive years lack content validity. Discussion of the results and suggestions about constructing exam items with high validity are provided.