

## Computerized Adaptive Testing for Ipsative Tests with Multidimensional Pairwise-Comparison Items

**Authors :** Wen-Chung Wang, Xue-Lan Qiu

**Abstract :** Ipsative tests have been widely used in vocational and career counseling (e.g., the Jackson Vocational Interest Survey). Pairwise-comparison items are a typical item format of ipsative tests. When the two statements in a pairwise-comparison item measure two different constructs, the item is referred to as a multidimensional pairwise-comparison (MPC) item. A typical MPC item would be: Which activity do you prefer? (A) playing with young children, or (B) working with tools and machines. These two statements aim at the constructs of social interest and investigative interest, respectively. Recently, new item response theory (IRT) models for ipsative tests with MPC items have been developed. Among them, the Rasch ipsative model (RIM) deserves special attention because it has good measurement properties, in which the log-odds of preferring statement A to statement B are defined as a competition between two parts: the sum of a person's latent trait to which statement A is measuring and statement A's utility, and the sum of a person's latent trait to which statement B is measuring and statement B's utility. The RIM has been extended to polytomous responses, such as preferring statement A strongly, preferring statement A, preferring statement B, and preferring statement B strongly. To promote the new initiatives, in this study we developed computerized adaptive testing algorithms for MPC items and evaluated their performance using simulations and two real tests. Both the RIM and its polytomous extension are multidimensional, which calls for multidimensional computerized adaptive testing (MCAT). A particular issue in MCAT for MPC items is the within-person statement exposure (WPSE); that is, a respondent may keep seeing the same statement (e.g., my life is empty) for many times, which is certainly annoying. In this study, we implemented two methods to control the WPSE rate. In the first control method, items would be frozen when their statements had been administered more than a prespecified times. In the second control method, a random component was added to control the contribution of the information at different stages of MCAT. The second control method was found to outperform the first control method in our simulation studies. In addition, we investigated four item selection methods: (a) random selection (as a baseline), (b) maximum Fisher information method without WPSE control, (c) maximum Fisher information method with the first control method, and (d) maximum Fisher information method with the second control method. These four methods were applied to two real tests: one was a work survey with dichotomous MPC items and the other is a career interests survey with polytomous MPC items. There were three dependent variables: the bias and root mean square error across person measures, and measurement efficiency which was defined as the number of items needed to achieve the same degree of test reliability. Both applications indicated that the proposed MCAT algorithms were successful and there was no loss in measurement proficiency when the control methods were implemented, and among the four methods, the last method performed the best.

**Keywords :** computerized adaptive testing, ipsative tests, item response theory, pairwise comparison

**Conference Title :** ICSRD 2020 : International Conference on Scientific Research and Development

**Conference Location :** Chicago, United States

**Conference Dates :** December 12-13, 2020