

PENDIDIKAN

LAPORAN PENELITIAN FUNDAMENTAL (Tahun I)



**Mengembangkan Metode Pendeteksian Keberfungsian
Butir Pembeda (*Differential Item Functioning, DIF*)
Berdasarkan Teori Respons Butir Multidimensi**

Penanggung Jawab Program
Dr. Badrun Karto Wagiran

Peneliti Anggota
Dr. Heri Retnawati

Dibiayai oleh
Direktur Jendral Pendidikan Tinggi Departemen Pendidikan Nasional
Sesuai Surat Perjanjian Pelaksanaan Penelitian
Nomor : 018/SP2H/PP/DP2M/III/2008, tanggal 6 Maret 2008

UNIVERSITAS NEGERI YOGYAKARTA
November 2008

Mengembangkan Metode Pendeteksian Keberfungsian Butir Pembeda (*Differential Item Functioning, DIF*) Berdasarkan Teori Respons Butir Multidimensi

Oleh:
Badrun Kartowagiran
Heri Retnawati

Abstrak

Idealnya, pelaksanaan tes seharusnya berdasarkan objektivitas, transparansi, akuntabilitas dan tidak diskriminatif. Jika pada suatu tes memuat suatu butir-butir soal yang memihak kelompok tertentu, maka tes tersebut dikatakan memuat bias atau mengandung keberfungsian butir diferensial (*differential item functioning, DIF*). Pada pendeteksian DIF, metode yang ada dan telah dikembangkan oleh ahli pengukuran/peneliti terdahulu masih berkisar pada analisis DIF dengan menggunakan teori respons butir unidimensi. Metode ini akan kurang tepat jika diterapkan pada sebagian besar tes prestasi dan psikologi yang memuat dimensi ganda atau bersifat multidimensi, karena akan menghasilkan informasi yang menyesatkan. Dengan memperhatikan konsep dasar teori respons butir unidimensi dan multidimensi, dan prinsip metode dalam pendeteksian keberfungsian butir diferensial, dapat dikembangkan metode pendeteksian DIF berdasarkan teori respons butir multidimensi. Penelitian ini bertujuan untuk : (1) mengembangkan metode pendeteksian DIF berdasarkan teori respons butir multidimensi, (2) menyusun program komputer untuk mendeteksi DIF berdasarkan teori respons butir multidimensi, (3) mengaplikasikan metode dan program untuk menganalisis data, (4) mengetahui sifat-sifat metode pendeteksian DIF berdasarkan teori respons butir multidimensi, dan (5) mengetahui sensitivitas metode-metode pendeteksian DIF berdasarkan teori respons butir multidimensi.

Untuk mencapai target tersebut, penelitian ini menggunakan pendekatan penelitian dan pengembangan (*research and development*), yang meterdiri dari dua tahap. Tahap I (tahun I) merupakan penelitian yang dilakukan secara analitis matematis mencari hubungan berbagai persamaan pada metode pendeteksian DIF dan teori respons butir multidimensi, membuat program computer (software) untuk melakukan analisis DIF pada tes yang multidimensi, dan mengaplikasikannya pada data Ujian Akhir Nasional. Pada tahap II (tahun II), dilakukan studi simulasi untuk mengetahui sifat-sifat metode pendeteksian DIF berdasarkan teori respons butir multidimensi dan perbandingan sensitivitas metode-metode pendeteksian DIF berdasarkan teori respons butir multidimensi.

Dengan berdasarkan sifat teori respons butir multidimensi dan metode mengidentifikasi DIF unidimensi, dapat dikembangkan 2 metode mengidentifikasi DIF multidimensi, yakni dengan metode indeks volume sederhana dan indeks perbedaan probabilitas, yang diuji signifikansinya dengan perbandingan likelihood dan distribusi sampling. Hasil lain yang diperoleh yakni manual mengidentifikasi DIF dengan pendekatan teori respons butir multidimensi yang berupa kumpulan sintaks (baris-baris perintah) dari *software*. Baris-baris perintah ini telah diaplikasikan pada data Ujian Nasional mata pelajaran Matematika tahun 2005 yang merupakan model data 2 dimensi untuk studi simulasi yang akan dilaksanakan pada tahun kedua.

Differential Item Functioning (DIF)
Based on multidimensional item response theory

Oleh:
Badrun Kartowagiran
Heri Retnawati

Abstract

The ideal condition of the administration of test is objective, transparency, fair and not discriminative. If a test contains any items favoring a group of testee, the test is unfair, containing item bias or differential item functioning, DIF. There are many methods for identifying DIF have been developed by psychometric researcher using item response theory that has two assumption that are local independent and unidimension. The definition of unidimensional test is a test is measuring only single ability. It can be shown by test only measures the dominant component of testees' ability. Practically the assumption is difficult to fulfill tightly. The most educational and psychological tests is multidimensional, because the tests are not only measuring the dominant component, but also other component. In this situation, the item analysis using one-dimensional approach has been inappropriate again, and will result a systematic mistake and the informations obtained will mislead. By paying attention to the characteristics of the unidimensional and multidimensional item response theory (MIRT) and the concepts of detecting DIF, can be developed the methods for detecting DIF using the multidimensional item response theory. The aims of research are: (1) developing the methods for detecting DIF using multidimensional item response theory, (2) arrange the syntax of software for detecting DIF using multidimensional item response theory, (3) apply the methods and syntax of software for data analyzing, (4) find out the characteristics of the methods for detecting DIF, and (5) find out the sensitivity of methods for detecting DIF using multidimensional item response theory.

The approach for this research is research and development, derived for two stages. The first stage (first year) is a research that it was held by mathematical analysis to find out the relations between many equations in methods for detecting DIF and the multidimensional item response theory, arrange syntax of software for detecting DIF using the multidimensional item response theory, and apply this methods and syntax for analysis of the national examination data. In the second stage (second year), this research will be held using simulation study to find out the characteristics of methods for detecting DIF using MIRT and the comparison of sensitivity of the methods.

The results of the first year are (1) there are two methods for detecting DIF using MIRT (the simple volume indices and probability different indices), and two methods for testing the significance of the indices (likelihood ratio test and sampling distribution), (2) the collection of syntax for detecting DIF using MIRT, and apply the methods and syntax for detecting DIF in the mathematics national examination data in 2005 academic year, and as an analysis model for simulation study that can be held in the second year.

