

Abstract. This study aims to detect items reasoning measure that affected by differential item functioning (DIF). Gajah Mada Reasoning Test (GAMA-RT) was analyzed. Participants were 262, consisting of 179 males (68.3%) and 83 females (31.7%). Data were analyzed using the Mantel-Haenszel DIF test with item response theory (IRT) modeling. Results showed that 12 items of Gajah Mada Reasoning Test were indicated by DIF based on gender. Items which indicated DIF indicates that certain gender tend to find it easier to answer an item than the other. These findings of this study can contribute to the evaluation of the fairness of reasoning measure, help to understand how test items can function differently in each group, also serve as basis for reviewing and refining test items.

Keywords: *Differential Item Functioning, Reasoning Measure, Gender*

Abstrak. Penelitian ini bertujuan untuk mendeteksi butir-butir pengukuran penalaran yang terjangkit keberfungsian butir diferensial (DIF). Instrumen pengukuran yang dianalisis adalah Gajah Mada Reasoning Test (GAMA-RT). Partisipan penelitian berjumlah 262 peserta, terdiri atas 179 laki-laki (68,3%) dan 83 perempuan (31,7%). Data dianalisis dengan menggunakan uji DIF metode Mantel-Haenszel melalui pendekatan teori respons butir (Item Response Theory/IRT). Hasil analisis menunjukkan 12 butir Gajah Mada Reasoning Test terindikasi DIF berdasarkan jenis kelamin. Butir yang terindikasi DIF menunjukkan bahwa jenis kelamin tertentu cenderung lebih mudah mengerjakan suatu butir dibanding jenis kelamin lainnya. Temuan penelitian ini dapat memberikan kontribusi terhadap evaluasi keadilan pengukuran penalaran, membantu memahami bagaimana butir tes dapat berfungsi secara berbeda pada masing-masing kelompok, serta dasar dalam peninjauan dan penyempurnaan butir tes.

Kata kunci: *Keberfungsian Butir Diferensial, Pengukuran Penalaran, Jenis Kelamin*