

ABSTRAK

Analisis Tes Berpikir Verbal A. IIIA1 Dengan Pendekatan Item Response Theory

Oleh: Ananto Febrian

Tes Berpikir Verbal A. IIIA1 adalah subtes Tes Bakat yang bertujuan untuk mengukur kemampuan penalaran siswa SMA. Tes ini adalah hasil pengembangan P2TKP yang telah melalui proses standarisasi dan telah diterapkan secara nasional. Namun demikian telah sekian lama, tes ini tidak pernah dikaji ulang, sehingga dikhawatirkan tes ini memiliki beberapa kelemahan.

Penelitian ini adalah penelitian deskriptif yang bertujuan untuk menjabarkan karakteristik item dan tes dengan pendekatan *item response theory* indikator yang digunakan adalah parameter ambang, parameter diskriminasi, *standard error*, fungsi informasi, dan reliabilitas. Subjek pada tes ini adalah siswa lulusan SMP dan siswa SMA dengan rentang umur antara 14 – 18 tahun.

Hasil analisis parameter ambang (b) menunjukkan 26 dari 50 item masuk kategori rata – rata ($-0,6 < SD < 0,6$), 4 item mudah, 16 item sulit, dan 4 item sangat sulit, yaitu item 36, 39, 42, dan 47, 3 diantaranya (item 36 ($b=4,338$), item 42 ($b=4,667$), dan item 47 ($b=12,634$)) berada di luar rentang skala. Item – item yang mempunyai *standard error* estimasi nilai b yang signifikan ($p = 0,05$) adalah 2 ($t=1,630$), 6 ($t= 1,048$), 15 ($t= -0,358$), 23 ($t= 1,105$), 26 ($t= -0,979$), dan 37 ($t= 1,520$). *Mean* nilai b sebesar 0,830 dan $SD= 0,241$. Hasil analisis parameter diskriminasi (a) menunjukkan terdapat 3 item yang tidak diskriminatif (2 ($a = 0,298$), 10 ($a = 0,262$), 47 ($a = 0,247$)), serta tiga item yang diskriminasinya sangat rendah, yaitu 15 ($a = 0,452$), 21 ($a = 0,588$), 33 ($a = 0,509$). Nilai a tidak mempunyai nilai *standard error* yang signifikan ($p = 0,05$). *Mean* nilai a sebesar 1,101 dan $SD= 0,440$. Hasil analisis informasi item menunjukkan 5 item, yaitu item 1, 3, 10, 11, dan 49 mempunyai ketelitian maksimal untuk tingkat *ability* yang sangat rendah ($SD < -1,8$). 15 item, yaitu item 4, 12, 13, 17, 24, 25, 28, 29, 31, 33, 34, 35, 43, 46, dan 50 teliti untuk tingkat *ability* tinggi ($0,6 < SD < 1,8 =$ tinggi), 5 item, yaitu item 21, 36, 39, 42, dan 47 teliti untuk item yang sangat tinggi ($SD > 1,8$). Sementara 25 item sangat teliti pada tingkat *ability* rata – rata ($-0,6 < SD < 0,6$). Dari analisis informasi menunjukkan terdapat 6 item dengan informasi terendah, yaitu item 2 ($I=0,0212$), 10 ($I= 0,016$), 15 ($I= 0,049$), 21 ($I= 0,085$), 33 ($I= 0,0632$), dan 47 ($I= 0,063$). Hasil *test information curve* (TIC) menunjukkan bahwa puncak kurva (14,5175) jatuh pada nilai *ability* 0,25. Ini menunjukkan bahwa tes ini mempunyai kemampuan mengukur yang paling rinci pada subjek dengan tingkatan *ability* rata – rata. Analisis reliabilitas dengan pendekatan empiris menunjukkan reliabilitas sebesar 0,9122 dan teoretis sebesar 0,921. Hal ini menunjukkan bahwa skor yang dihasilkan oleh tes sangat konsisten.

Keywords: *Item response theory*, Tes Berpikir Verbal A. IIIA1, parameter ambang, parameter diskriminasi, informasi, *standard error*, reliabilitas, tingkat *ability*.

ABSTRACT

Tes Berpikir Verbal A. IIIA1 Analysis With Item Response Theory Approach

By: Ananto Febrian

Tes Berpikir Verbal A. IIIA1 is a subtest from Tes Bakat (aptitude test) aimed to measure high school student's reasoning ability. The test is the product of P2TKP's development that already pass the standardization procedure and nationally applied. Nevertheless for a long time the test never been re-analyzed, so it's feared that this test has several weaknesses.

This research is descriptive research aimed to explain the item and test characteristic. This research using item response theory approach with indicators such as threshold parameter, discrimination parameter, standard error, information function, and reliability. The subject of this research is middle school graduate and highschool student from 14 to 18 years old.

The result from threshold parameter (b) show that 26 of 50 items falls on average category ($-0.6 < SD < 0.6$), 4 items falls on easy category, 16 items is hard, 4 items is very hard, that is item 36, 39, 42, 47 and three of them are outside the scale area (36 ($b=4,338$), 42 ($b=4,667$), 47 ($b=12,634$)). Items that have significant b value estimate *standard error* ($p=0,05$) is, item 2 ($t=1,630$), 6 ($t=1,048$), 15 ($t=-0,358$), 23 ($t=1,105$), 26 ($t=-0,979$), and 37 ($t=1,520$). b value have mean= $0,830$ and $SD= 0,241$. The result of discrimination parameter analysis show that there are 3 items that not discriminative, that are item 2 ($a=0,298$), 10 ($a=0,262$), and 47 ($a=0,247$), very low discrimination, that are item 15 ($a=0,452$), 21 ($a=0,588$), 33 ($a=0,509$). Estimate a value doesn't have significant standard error ($p=0,05$). a value have mean= $1,101$ and $SD=-0,440$. The result of item information analysis show that 5 items, that are item 1, 3, 10, 11, and 49, That have maximum precision for very low ability level ($SD < -1,8$). 15 items, that is item 4, 12, 13, 17, 24, 25, 28, 29, 31, 33, 34, 35, 43, 46, and 50, have maximum precision for high ability level ($0,6 < SD < 1,8 =$ tinggi), 5 items, that is item 21, 36, 39, 42, and 47, have maximum precision for very high ability level ($SD > 1,8$). 25 items left have maximum precision for average ability level ($-0,6 < SD < 0,6$). This analysis also shows that there is 6 items with the lowest information, that is item 2 ($I=0,0212$), 10 ($I= 0,016$), 15 ($I= 0,049$), 21 ($I= 0,085$), 33 ($I= 0,0632$), dan 47 ($I= 0,063$). The result from test information function show that the top of the curve ($14,5175$) falls on point $0,25$ in the ability scale. This result show that the test have the maximum precision to measure subjects with the average ability level. The result from reliability analysis show that from empirical approach, the test have reliability $0,9122$ and from theoretical approach $0,921$. This result show that the score that produced by this test is very consistent.

Keywords: Item response theory, Tes Berpikir Verbal A. IIIA1, Threshold parameter, discrimination parameter diskriminasi, informastion, standard error, reliability, ability