

**INDEKS RELIABILITAS ITEM (iri) SEBAGAI PENYUSUN
STANDAR DEVIASI TOTAL (SD_t) PADA TES
KEMAMPUAN BERPIKIR MANTIK**

Oleh: I Gusti Ngurah Puger¹

Abstrak

Tujuan yang ingin disasar dalam pelaksanaan penelitian ini adalah: untuk mengetahui besarnya kontribusi skor-tampak masing-masing *item* dari tes kemampuan berpikir mantik di dalam menyusun standar deviasi total distribusi skor tampak, bila dikaji dari indeks reliabilitas *item*-nya. Penelitian yang dilaksanakan merupakan jenis penelitian evaluasi, yang subjeknya berupa 100 siswa kelas VIII SMP Negeri 1 Seririt. Pemilihan subjek penelitian ini menggunakan teknik *purposive sampling*. Instrumen yang digunakan berupa tes kemampuan berpikir mantik yang dikompilasi oleh peneliti. Untuk menentukan koefisien indeks reliabilitas *item* (iri) digunakan formula iri. Berdasarkan atas hasil analisis data dengan formula iri diperoleh koefisien iri pada masing-masing butir yang menyusun tes kemampuan berpikir mantik. Setelah ke-54 koefisien iri yang diperoleh dijumlahkan diperoleh nilai sebesar 9,178. Sedangkan nilai standar deviasi total (SD_t) dari tes kemampuan berpikir mantik sebesar 9,178. Oleh karena nilai koefisien iri secara total sama dengan nilai SD_t dari tes kemampuan berpikir mantik, maka H_1 diterima. Oleh karena itu dapat disimpulkan bahwa besarnya kontribusi skor-tampak masing-masing *item* tes kemampuan berpikir mantik di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak sama dengan besarnya indeks reliabilitas *item*-nya.

Kata kunci: *Kontribusi item, standar deviasi total, indeks reliabilitas item, dan tes kemampuan berpikir mantik.*

Abstract

The objectives to be targeted in the implementation of this study was to find out the magnitude of the contribution of the visible score of each item from the test of cognitive thinking ability in compiling a standard deviation of the total visible score distribution, when it was assessed from its item reliability index. The research conducted was a type of evaluation research, the subjects were 100 students in class VIII of SMP Negeri 1 Seririt. The selection of the subjects of this study was through a purposive sampling technique. The instrument used in the form of tests of cognitive thinking ability which was compiled by the researcher. In order to determine the Item Reliability Index (IRI) coefficient, IRI formula was used. Based on the results of data analysis using IRI formula, the IRI coefficient was obtained for each item that compiled the

¹ Dr. I Gusti Ngurah Puger, M.Pd. adalah staf edukatif pada Fakultas Keguruan dan Ilmu Pendidikan (FKIP) Universitas Panji Sakti Singaraja.

ability test to think logic. After the 54 coefficients of IRI obtained were summed up to a value of 9,178. Meanwhile the total standard deviation (SDt) value of the cognitive thinking ability test was 9,178. Due to the value of IRI coefficient was totally similar to SDt value from the logic thinking ability test, then H1 was accepted. Therefore, it can be concluded that the magnitude of the visible score contribution of each item of the logic thinking ability in compiling a total standard deviation (SDt) score distribution looked similar to the magnitude of its item reliability index.

Keywords: Contribution of items, total standard deviation, item reliability index, and tests of cognitive ability.

PENDAHULUAN

Dalam pengajaran sains, sekurang-kurangnya siswa diajak ke alam untuk mengenal alam. Terkait dengan hal ini, Wonorahardjo (2011) menyatakan pengenalan akan alam merupakan salah satu kegiatan manusia yang terjadi secara spontan. Pengenalan merupakan awal dari pokok bahasan mengenai pengetahuan, baik sebagai produk maupun sebagai proses. Produk pengenalan dan pengetahuan yang dikumpulkan secara sistematis akan membentuk wacana baru, yakni ilmu. Pengenalan manusia akan alam secara objektif dan ilmiah akan menghasilkan sains (*science*). Sains adalah pengetahuan alam yang terstruktur. Sains akan berkembang lebih mendalam karena manusia senantiasa mengetahui hal-hal baru di alam dan menganalisisnya dalam kaidah-kaidah ilmiah dan universal.

Sebagai bahan penguat pendapat Wonorahardjo, ada baiknya juga dikaji pendapat Goethe. Goethe (dalam Bawa, 1992) seorang sastrawan dan penyelidik alam bangsa Jerman pernah membuat pernyataan yang artinya bahwa alam adalah satu-satunya buku, yang tiap-tiap lembarannya mengandung isi yang sangat luas. Dari segi biologi, alam dapat dipandang sebagai sebuah laboratorium yang besar, guna mempelajari fenomena biologi. Terkait dengan pernyataan Goethe ini, Galilio (dalam Rindjin, 2001) pernah mengemukakan *Universe –the grand book-stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in wich it is composed. It is written in the language of mathematics. Then mathematics is not science, but mathematics is the language of science.*

Dalam mempelajari sains, seorang siswa tidak bisa hanya mengandalkan teoretis saja, atau dalam domain filosofis sains dikenal sebagai mengandalkan rasionalitas saja melainkan harus juga bertumpu pada empiris (fakta-fakta ilmiah). Bukankah sains muncul diawali oleh observasi fenomena tertentu di alam, lalu maju menjadi teori-teori sains. Bilamana guru bidang studi sains di sekolah menengah hanya menjejali kajian teoretis saja, tanpa disangga oleh pengamatan langsung ke alam maka struktur kognitif siswa hanya dicemari oleh teori tanpa ada proses pembuktiannya. Untuk menanggulangi hal ini, hendaknya guru bidang studi sains di sekolah menengah membiasakan anak untuk berpikir ilmiah. Atau meminjam istilah yang dikemukakan oleh Nasoetion (1995), berpikir ilmiah ini dikenal sebagai berpikir mantik.

Berpikir dengan menggunakan kemampuan berpikir mantik berarti berpikir menggunakan penalaran. Terkait dengan berpikir menggunakan penalaran, Yanto (2001), menyatakan berpikir dengan penalaran maksudnya berpikir tepat dan jitu. Berpikir yang memerlukan kerja otak dan akal untuk merumuskan sesuatu yang sesuai dengan patokan logika. Melihat kenyataan pada diri kita lalu menggabungkan sebab dan akibat. Artinya setiap apa yang kita perbuat hendaknya disesuaikan dengan keadaan yang ada pada diri kita. Bila hal tersebut cocok dengan kenyataan dan andaikata dikerjakan bisa mengakibatkan keuntungan, maka laksanakanlah hal tersebut.

Bila dalam pembelajaran sains, guru bidang studi tersebut sudah biasa melatih siswa untuk berpikir deduktif ke induktif atau dari induktif ke deduktif pada hakikatnya sudah melatih siswa untuk mengasimilasi dan mengakomodasikan konsep-konsep dalam bidang sains. Pemahaman yang holistik dari suatu konsep dalam bidang sains akan dapat meningkatkan hasil belajar sains siswa.

Dalam mengukur kemampuan berpikir mantik siswa Sekolah Menengah (SM) diperlukan alat ukur yang ditentukan berdasarkan atas kisi-kisi tes kemampuan berpikir mantik. Tes kemampuan berpikir mantik yang digunakan dalam penelitian ini, menggunakan *grand theory* yang dikemukakan oleh Kusuma (2011). Dalam artian, penurunan butir-butir yang menyusun tes kemampuan berpikir mantik

menggunakan dimensi, sub-dimensi, dan indikator yang dikemukakan oleh Kusuma. Setelah tes kemampuan berpikir mantik selesai dikompilasi oleh perancang suatu kegiatan riset, tes yang bersangkutan perlu distandarisasi. Menurut Puger (2016), dalam proses standarisasi suatu tes, misalnya tes kemampuan berpikir mantik, sekurang-kurangnya dilakukan: (a) pengujian validitas isi, (b) menghitung reliabilitas respon antar-rater, (c) menguji validitas butir, dan (d) menghitung reliabilitas tes.

Dalam menghitung validitas butir tes kemampuan berpikir mantik, digunakan rumus korelasi *point biserial*. Untuk memperoleh koefisien r_{pbi} suatu butir, biasanya menggunakan standar deviasi total (SD_t) distribusi skor-tampak (Djaali dan Muljono, 2008). Standar deviasi total (SD_t) distribusi skor-tampak sebetulnya merupakan kontribusi skor-tampak dari masing-masing *item* yang menyusun tes kemampuan berpikir mantik.

Demikian juga halnya di dalam menghitung reliabilitas tes kemampuan berpikir mantik, biasanya digunakan formula K-R 20. Untuk menentukan koefisien $r_{xx'}$ tes kemampuan berpikir mantik, diperlukan varians total (SD_t^2) distribusi skor tampak (Maba, 2002). Varians total (SD_t^2) merupakan standar deviasi total distribusi skor-tampak dikuadratkan. Dengan kata lain, akar kuadrat varians total distribusi skor-tampak merupakan standar deviasi total. Standar deviasi total (SD_t) distribusi skor-tampak yang digunakan untuk menghitung koefisien $r_{xx'}$, juga merupakan kontribusi skor-tampak dari masing-masing *item* yang menyusun tes kemampuan berpikir mantik.

Hanya saja, tidak banyak orang yang menggunakan standar deviasi dalam kegiatan risetnya, yang merupakan hasil dari pengukuran menggunakan tes kemampuan tahu mengenai asal-usul standar deviasi total distribusi skor-tampak. Dengan kata lain, tidak banyak orang yang tahu mengenai skor deviasi total distribusi skor-tampak yang digunakan dalam penelitiannya, merupakan hasil dari kontribusi skor-tampak pada masing-masing *item* yang menyusun tes kemampuan, misalnya tes kemampuan berpikir mantik.

Pada sisi lain, periset di dalam menentukan koefisien $r_{xx'}$ sering melupakan peran dari penentuan indeks reliabilitas *item* (iri). Bahkan konsep indeks reliabilitas

item (iri) dalam standarisasi tes kemampuan tidak banyak, bahkan hampir tidak pernah tersentuh oleh periset yang menggunakan tes kemampuan. Menurut Azwar (2003), indeks reliabilitas *item* (iri) merupakan karakteristik *item* yang diderivasi dari konsistensi satu *item* dengan keseluruhan item yang lainnya dalam tes yang bersangkutan. Bahkan Puger (2017) menyatakan keseluruhan indeks reliabilitas *item* ini apabila dijumlahkan maka besarnya harus sama dengan standar deviasi total (SD_t) distribusi skor-tampak.

Hal ini sesuai dengan pendapat Azwar (2003) yang pada hakikatnya menyatakan karena standar deviasi *item* dikotomi (*biased estimation*) tidak lebih besar daripada 0,50 dan korelasi *item*-total tidak dapat lebih besar daripada 1,00, maka indeks reliabilitas butir juga tidak akan dapat lebih besar daripada 0,50. Keseluruhan indeks reliabilitas *item* ini apabila dijumlahkan maka besarnya akan sama dengan standar deviasi total (SD_t) skor tes yang bersangkutan.

Berdasarkan atas kenyataan-kenyataan yang sudah disampaikan, dalam artikel ini akan dipertelakan masalah kontribusi *item* di dalam menyusun standar deviasi total (SD_t) tes kemampuan berpikir mantik.

METODOLOGI PENELITIAN

Penelitian ini termasuk jenis penelitian evaluasi. Evaluasi adalah proses sistematis dari pengumpulan dan analisis data dalam urutan untuk mengambil keputusan. Evaluasi melibatkan pertanyaan seperti berikut. (1) Apakah suatu instrumen pengukur itu berharga?, (2) mengapa respon terhadap instrumen lebih penting daripada bentuk instrumen?, dan (3) apakah keputusan Anda terhadap instrumen yang telah teruji validitas isi, validitas butir, dan terhitung reliabilitasnya? Jawaban untuk pertanyaan tersebut menghendaki pengumpulan, analisis, dan interpretasi dari data itu berkenaan dengan satu atau beberapa kriteria. Kriteria yang lebih objektif, lebih baik, walaupun beberapa derajat subjektivitas tidak diperhatikan waktu orang menentukan kriteria (Gay, 1987). Dalam penelitian ini dilakukan proses analisis eksternal tentang instrumen penelitian berupa ‘tes kemampuan berpikir mantik.’ Analisis eksternal dilakukan melalui penentuan tentang indeks reliabilitas

item (iri) dari tes kemampuan berpikir mantik. Dari indeks reliabilitas *item* (iri) yang diketahui, selanjutnya digunakan sebagai bagian untuk menentukan kontribusi *item* terhadap standar deviasi total (SD_i) distribusi skor-tampak.

Wilayah generalisasi hasil penelitian ini adalah semua instrumen daya (*power instrument*), yaitu instrumen yang bisa diselesaikan oleh sebagian besar peserta dalam satuan waktu tertentu. Sehingga waktu bukan merupakan faktor yang menentukan besarnya skor instrumen (Mardapi, 1994).

Populasi penelitian ini adalah jawaban terhadap tes kemampuan berpikir mantik yang dikembangkan oleh peneliti. Sedangkan cuplikannya adalah data jawaban 100 siswa kelas VIII SMP Negeri 1 Seririt terhadap tes kemampuan berpikir mantik yang berupa jawaban terhadap tiap butir tes kemampuan berpikir mantik. Penentuan subjek sebanyak 100 siswa kelas VIII dalam penelitian ini menggunakan *purposive sampling*. Menurut Dantes (1991), *purposive sampling* merupakan cara pengambilan sampel sesuai dengan tujuan penelitian yang telah ditetapkan sebelumnya. Anggota sampel yang diseleksi dari sub-sub kelompok dalam populasi yang sesuai dengan tujuan penelitian.

Instrumen yang dibutuhkan untuk memperoleh data berupa jawaban tes adalah tes kemampuan berpikir mantik yang dikembangkan oleh peneliti. Tes tersebut terdiri atas 54 butir dan harus diberikan jawaban oleh 100 siswa dalam waktu 60 menit. Sebelum tes ini disusun, terlebih dahulu dibuat definisi operasional kemampuan berpikir mantik dan tabel spesifikasi tes kemampuan berpikir mantik, yang memuat nama variabel, dimensi (aspek), indikator, dan nomor butir.

Kemampuan berpikir mantik adalah kemampuan berpikir dengan jalan menggabungkan antara berpikir deduktif dan berpikir induktif. Kemampuan berpikir mantik ini diukur dengan tes kemampuan berpikir mantik (*achievement test of mantiq mind*), dan data yang diperoleh berupa data interval. Data skor-tampak pada setiap butir dan data skor-tampak total, selanjutnya digunakan untuk menentukan nilai SD_i dan koefisien korelasi *product moment* dari Pearson (r_{ix}) pada setiap butir yang menyusun tes kemampuan berpikir mantik. Selanjutnya nilai SD_i dan koefisien r_{ix} digunakan untuk menentukan indeks reliabilitas *item* (iri).

Tes kemampuan berpikir mantik yang digunakan untuk menjaring data dalam penelitian ini sudah diuji validitas isi dan dihitung reliabilitas respon antar-rater-nya. Dengan menggunakan formula Gregory, diperoleh koefisien validitas isi (VI) sebesar 0,98. Sedangkan reliabilitas respon antar-rater-nya dihitung dengan menggunakan formula Anava Hoyt, yang selanjutnya diperoleh koefisien $r_{xx'}$ sebesar 0,81. Berdasarkan atas koefisien VI dan koefisien $r_{xx'}$ yang diperoleh, maka tes kemampuan berpikir mantik ini dapat digunakan lebih lanjut.

Untuk menentukan besarnya iri digunakan formula yang dikemukakan oleh Azwar (2003). Adapun formula iri tersebut adalah: $iri = S_i \times r_{ix}$. Di mana, iri merupakan indeks reliabilitas butir, S_i sebagai standar deviasi butir, dan r_{ix} sebagai koefisien korelasi *product moment* antara butir i dengan skor total dari tes kemampuan berpikir mantik.

Kriteria pengujian: (a) bila semua nilai iri dijumlahkan dan membentuk iri-total (iri_t) jumlahnya sama dengan standar deviasi total (SD_t), maka hipotesis nol (H_0) ditolak dan hipotesis alternatif (H_1) diterima, dan (b) bila semua nilai iri dijumlahkan dan membentuk iri-total (iri_t) jumlahnya tidak sama dengan standar deviasi total (SD_t), maka hipotesis nol (H_0) diterima dan hipotesis alternatif (H_1) ditolak. Dari penerimaan dan penolakan hipotesis nol (H_0) akan dapat ditarik suatu simpulan.

HASIL PENELITIAN DAN PEMBAHASAN

Sebelum dilakukan analisis data dengan menggunakan formula iri, terlebih dahulu harus dikemukakan hipotesis alternatif (H_1) dan hipotesis nol (H_0). Adapun bunyi rumusan H_1 adalah: besarnya kontribusi skor-tampak masing-masing *item* tes kemampuan berpikir mantik di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak sama dengan besarnya indeks reliabilitas *item*-nya. Sedangkan H_0 berbunyi: besarnya kontribusi skor-tampak masing-masing *item* tes kemampuan berpikir mantik di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak tidak sama dengan besarnya indeks reliabilitas *item*-nya.

Hasil analisis data dengan formula iri, diperoleh koefisien iri pada butir: 1 = 0,256, butir 2 = 0,268, butir 3 = 0,187, butir 4 = 0,190, butir 5 = 0,170, butir 6 =

0,156, butir 7 = 0,136, butir 8 = 0,246, butir 9 = 0,173, butir 10 = 0,202, butir 11 = 0,048, butir 12 = 0,149, butir 13 = 0,158, butir 14 = 0,065, butir 15 = 0,079, butir 16 = 0,120, butir 17 = 0,121, butir 18 = 0,145, butir 19 = 0,181, butir 20 = 0,197, butir 21 = 0,169, butir 22 = 0,160, butir 23 = 0,179, butir 24 = 0,184, butir 25 = -0,049, butir 26 = 0,157, butir 27 = 0,242, butir 28 = 0,173, butir 29 = 0,185, butir 30 = 0,200, butir 31 = 0,144, butir 32 = 0,193, butir 33 = 0,183, butir 34 = 0,157, butir 35 = 0,219, butir 36 = 0,151, butir 37 = 0,203, butir 38 = 0,262, butir 39 = 0,199, butir 40 = 0,184, butir 41 = 0,202, butir 42 = 0,189, butir 43 = 0,174, butir 44 = 0,160, butir 45 = 0,243, butir 46 = 0,184, butir 47 = 0,148, butir 48 = 0,168, butir 49 = 0,170, butir 50 = 0,158, butir 51 = 0,263, butir 52 = 0,181, butir 53 = 0,027, dan butir 54 = 0,166 dari tes kemampuan berpikir mantik.

Apabila nilai indeks reliabilitas butir (iri) yang diperoleh dijumlahkan maka akan diperoleh nilai iri total (-atau Σiri). Nilai dari $iri_t = 9,178$. Sedangkan nilai SD_t (yang dihitung dengan Microsoft Excel 2003) sebesar 9,178. Oleh karena nilai iri total (iri_t) sebesar 9,178 sama dengan nilai SD_t sebesar 9,178, maka H_1 yang berbunyi: besarnya kontribusi skor-tampak masing-masing *item* tes kemampuan berpikir mantik di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak sama dengan besarnya indeks reliabilitas *item*-nya, diterima.

Hal ini disebabkan oleh butir-butir yang menyusun tes kemampuan berpikir mantik sudah diturunkan dari teori dan indikator yang jelas strukturnya bila dikaji dari kajian teoretis. Perlu diketahui bahwa teori dan indikator-indikator yang digunakan sebagai dasar untuk menurunkan butir-butir tes kemampuan berpikir mantik diadopsi dari Kusuma (2011).

Tes kemampuan berpikir mantik yang sudah selesai dikompilasi, sebelum diujicobakan kepada 100 siswa kelas VIII SMP Negeri 1 Seririt sudah diuji validitas isi (VI) dan reliabilitas respon antar-penilainya. Dari hasil pengujian validitas isi, diperoleh koefisien VI sebesar 0,98 dan dari hasil penghitungan koefisien $r_{xx'}$ diperoleh nilai sebesar 0,81. Berdasarkan atas hal ini, tes kemampuan berpikir mantik sudah boleh diujicobakan pada siswa SMP.

Skor yang menyusun suatu butir tes kemampuan berpikir mantik bisa digunakan untuk menentukan nilai SD_i dan korelasi antara skor butir dengan skor totalnya (r_{ix}). Hasil kali dari SD_i dengan r_{ix} akan menghasilkan indeks reliabilitas butir (iri). Indeks reliabilitas butir (iri) ini dapat digunakan sebagai dasar untuk menentukan kontribusi skor-tampak masing-masing butir di dalam menyusun standar deviasi total (SD_t) distribusi skor tampak.

Hal ini sesuai dengan pendapat Azwar (2003) yang pada hakikatnya menyatakan karena standar deviasi *item* dikotomi (*biased estimation*) tidak lebih besar daripada 0,50 dan korelasi *item*-total tidak dapat lebih besar daripada 1,00, maka indeks reliabilitas butir juga tidak akan dapat lebih besar daripada 0,50. Keseluruhan indeks reliabilitas *item* ini apabila dijumlahkan maka besarnya akan sama dengan standar deviasi total (SD_t) skor tes yang bersangkutan.

Temuan dalam penelitian ini juga didukung oleh kajian Puger (2017) yang pada hakikatnya menyatakan indeks kekonsistenan butir suatu tes kemampuan dapat digunakan sebagai dasar untuk menentukan kontribusi masing-masing butir tes yang bersangkutan di dalam menyusun standar deviasi total (SD_t). Hal ini disebabkan oleh pada hakikatnya masing-masing indeks reliabilitas *item* (iri) bila dijumlahkan besarnya harus sama dengan SD_t .

Dengan demikian, bisa juga dikatakan bahwa pada hakikatnya standar deviasi total (SD_t) dari tes kemampuan berpikir mantik disusun oleh indeks reliabilitas butir (iri) dari tes yang bersangkutan. Kesesatan berpikir sering terjadi atau miskonsepsi dalam hal penyusunan standar deviasi total (SD_t) terdiri atas standar deviasi masing-masing butir (SD_i) dari tes kemampuan berpikir mantik. Bilamana hal ini yang terjadi, maka Anda diharapkan untuk mengkaji kembali konsep standar deviasi butir (SD_i) dan standar deviasi total (SD_t). Setelah Anda paham dengan kedua konsep standar deviasi yang dikemukakan, cobalah dicek, apakah $\sum SD_i$ sama besarnya dengan SD_t ? Atau $\sum iri$ (indeks reliabilitas butir secara keseluruhan) yang sama besarnya dengan SD_t .

Dari pendapat Azwar dan kajian Puger di atas, dapat dikatakan bahwa dalam menentukan kontribusi masing-masing *item* suatu tes kemampuan di dalam

menyusun standar deviasi total (SD_t) distribusi skor- tampak harus menasar tentang besarnya nilai indeks reliabilitas butir (iri) dan mengabaikan besarnya standar deviasi butir (SD_i). Hal ini disebabkan oleh pada hakikatnya yang membentuk kontribusi masing-masing butir dari tes kemampuan di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak bukanlah standar deviasi butir (SD_i), melainkan ditentukan oleh indeks reliabilitas butir (iri).

Informasi yang lain, juga menyatakan dalam menentukan indeks iri, bisa digunakan formula: $iri = \sqrt{p_i q_i} * r_{ix}$. Di mana iri = indeks reliabilitas butir, p_i = proporsi siswa yang menjawab benar butir ke-i, q_i = proporsi siswa yang menjawab salah butir ke-i, dan r_{ix} = koefisien korelasi *product moment* antara skor butir ke-i dengan skor totalnya. Penjumlahan nilai absolut dari iri sama dengan standar deviasi total (SD_t) skor-tampak.

SIMPULAN

Berdasarkan atas hasil analisis data dan pembahasan, dapat disimpulkan hal-hal sebagai berikut. Besarnya kontribusi skor-tampak masing-masing *item* tes kemampuan berpikir mantik di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak sama dengan besarnya indeks reliabilitas *item*-nya. Dengan kata lain, kontribusi butir 1 = 0,256, butir 2 = 0,268, butir 3 = 0,187, butir 4 = 0,190, butir 5 = 0,170, butir 6 = 0,156, butir 7 = 0,136, butir 8 = 0,246, butir 9 = 0,173, butir 10 = 0,202, butir 11 = 0,048, butir 12 = 0,149, butir 13 = 0,158, butir 14 = 0,065, butir 15 = 0,079, butir 16 = 0,120, butir 17 = 0,121, butir 18 = 0,145, butir 19 = 0,181, butir 20 = 0,197, butir 21 = 0,169, butir 22 = 0,160, butir 23 = 0,179, butir 24 = 0,184, butir 25 = -0,049, butir 26 = 0,157, butir 27 = 0,242, butir 28 = 0,173, butir 29 = 0,185, butir 30 = 0,200, butir 31 = 0,144, butir 32 = 0,193, butir 33 = 0,183, butir 34 = 0,157, butir 35 = 0,219, butir 36 = 0,151, butir 37 = 0,203, butir 38 = 0,262, butir 39 = 0,199, butir 40 = 0,184, butir 41 = 0,202, butir 42 = 0,189, butir 43 = 0,174, butir 44 = 0,160, butir 45 = 0,243, butir 46 = 0,184, butir 47 = 0,148, butir 48 = 0,168, butir 49 = 0,170, butir 50 = 0,158, butir 51 = 0,263, butir 52 = 0,181, butir 53

= 0,027, dan butir 54 = 0,166 di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak dari tes kemampuan berpikir mantik.

DAFTAR PUSTAKA

- Azwar, Saifuddin. 2003. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Bawa, Wayan. 1992. "Peranan Biologi dalam Beberapa Sektor Pembangunan." *Orasi Pengukuhan Jabatan Guru Besar dalam Ilmu Pendidikan Biologi pada FKIP Universitas Udayana*, Sabtu, 22 Pebruari 1992.
- Dantes, Nyoman. 1991. "Teknik Sampling." *Makalah* Disajikan dalam Penataran Dosen-Dosen Muda Kopertis Wilayah VIII, pada Tanggal 23 Maret 1991.
- Djaali, H. dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: PT Gramedia Widiasarana Indonesia.
- Gay, L.R. 1987. *Educational Research: Competencies for Analysis and Application*. Columbus, Ohio: Merrill Publishing Company.
- Kusuma, I Nyoman Wisnu. 2011. "Penerapan Kemampuan Berpikir Mantik dalam Pembelajaran Sains." *Makalah* yang Disampaikan dalam Seminar Nasional dengan Tema 'Meningkatkan Kemampuan Berpikir Kritis Siswa dalam Proses Pembelajaran Sains,' yang Diselenggarakan Oleh Dinas Pendidikan Pemuda dan Olahraga (Disdikpora) Provinsi Bali. Denpasar, 6 s.d. 8 Oktober 2011.
- Maba, Wayan. 2002. "Evaluasi Pembelajaran." *Makalah* yang Disajikan dalam Penataran PBM Dosen Kopertis Wilayah VIII, Tanggal 27-30 Oktober 2002.
- Mardapi, Djemari. 1994. *Analisis Butir dengan Teori Tes Klasik dan Teori Tes Respons Butir*. Laporan Penelitian yang Dibiayai Oleh Dana DPP IKIP Yogyakarta, No. 046/PT27H9/N.03.DPP/93.
- Puger, I Gusti Ngurah. 2016. "Konstruksi dan Analisis Hasil Ujicoba Tes Hasil Belajar Biologi." *Laporan Hasil Karya Inovatif* Program Hibah Mandiri FKIP Unipas Singaraja.
- , 2017. "Menentukan Indeks Reliabilitas Butir (iri) dari Tes Hasil Belajar." *Makalah* yang Disampaikan dalam Seminar Ilmiah dengan Tema 'Menganalisis Hasil Ujicoba Tes Hasil Belajar Sebagai Langkah untuk Menyusun Instrumen Pembelajaran yang Berkualitas,' yang Diselenggarakan Oleh K3S Kabupaten Banyuwangi, Tanggal 16 Juli 2017.
- Nasoetion, Andi Hakim. 1995. "Berpikir Mantik." *Makalah* yang Disampaikan dalam Penataran Metodologi Penelitian Bagi Dosen Muda IPB, Agustus 1995.
- Rindjin, Ketut. 2001. *Suplemen Materi Filsafat Ilmu Pengetahuan*. Singaraja: PPs IKIP Negeri Singaraja.
- Wonorahardjo, Surjani. 2011. *Dasar-Dasar Sains: Menciptakan Masyarakat Sadar Sains*. Jakarta: Indeks.
- Yanto, B.M. 2001. *Mengembangkan Pola Berpikir yang Baik*. Surabaya: Putra Pelajar.