

HUBUNGAN ANTARA KOEFISIEN EFEK SPURIOUS OVERLAP (ESO) DENGAN INDEKS RELIABILITAS ITEM (iri) TES KEMAMPUAN BERPIKIR MANTIK

I Gusti Ngurah Puger¹, I Nyoman Suprpta²

¹Fakultas Keguruan dan Ilmu Pendidikan, Universitas Panji Sakti, Singaraja

²Fakultas Ilmu Administrasi Universitas Panji Sakti, Singaraja

¹ngurah_puger@yahoo.co.id, ²nila_arianti@yahoo.com

ABSTRACT

The purpose of this study was to determine the significant relationship between the spurious overlap effect (ESO) coefficient and the item reliability index (IRI) of the logical thinking ability test. This research is a correlational study. Correlational research is a type of research that aims to find the relationship between one variable and another. The population of this study was all eighth-grade students of SMP Negeri 2 Seririt, while the sample consisted of the responses of 100 eighth-grade students of SMP Negeri 2 Seririt after being administered a logical thinking ability test, selected using purposive sampling. Purposive sampling is a sampling method in accordance with the previously established research objectives. The instrument needed to obtain data in the form of test responses was a logical thinking ability test. The logical thinking ability test used was deemed suitable for further use, as assessed by the content validity coefficient (VI), the inter-rater response reliability coefficient (rxx'), item validity, and test reliability (rxx'). To analyze the data regarding the relationship between the spurious overlap effect coefficient and the item reliability index (IRI) of the logical thinking ability test, the Pearson Product Moment correlation formula was used. The data analysis revealed that the spurious overlap effect (ESO) coefficient was inversely related to the item reliability index (IRI) of the logical thinking ability test. Based on the findings presented, it can be concluded that the spurious overlap effect (ESO) coefficient was inversely related to the item reliability index (IRI) of the logical thinking ability test.

Keywords: *spurious overlap effect, item reliability index, and logical thinking ability.*

ABSTRAK

Tujuan dari penelitian ini adalah untuk mengetahui hubungan secara signifikan antara koefisien efek *spurious overlap* (ESO) dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik. Penelitian ini termasuk jenis penelitian korelasional. Penelitian korelasional merupakan suatu jenis penelitian yang bertujuan untuk mencari hubungan antara suatu variabel dengan variabel yang lainnya. Populasi penelitian ini adalah seluruh siswa kelas VIII SMP Negeri 2 Seririt, sedangkan sampelnya berupa respon 100 siswa kelas VIII SMP Negeri 2 Seririt setelah diberikan tes kemampuan berpikir mantik yang dipilih secara *purposive sampling*. *Purposive sampling* merupakan cara pengambilan sampel sesuai dengan tujuan penelitian yang telah ditetapkan sebelumnya. Instrumen yang dibutuhkan untuk memperoleh data berupa jawaban tes adalah tes kemampuan berpikir mantik. Tes kemampuan berpikir mantik yang digunakan

sudah memiliki kelayakan untuk digunakan lebih lanjut, dikaji dari koefisien *content validity* (VI), koefisien reliabilitas ($r_{xx'}$) respon antar-rater, validitas butir, dan reliabilitas tes ($r_{xx'}$). Untuk menganalisis data mengenai hubungan antara koefisien efek *spurious overlap* dengan indeks reliabilitas butir tes (iri) kemampuan berpikir mantik digunakan formula korelasi *Product Moment* dari Pearson. Dari hasil analisis data diperoleh temuan bahwa koefisien efek *spurious overlap* (ESO) berhubungan secara terbalik dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik. Dari temuan yang sudah dikemukakan, dapat disimpulkan bahwa koefisien efek *spurious overlap* (ESO) berhubungan secara terbalik dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik.

Kata kunci: *Efek spurious overlap, indeks reliabilitas butir, dan kemampuan berpikir mantik.*

A. Pendahuluan

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.

Menurut Puger (2019), dalam menguji *validitas* butir tes kemampuan berpikir mantik biasanya digunakan formula korelasi *Product*

Moment (r_{ix}) yang tidak dikoreksi oleh efek *spurious overlap*. Koefisien r_{ix} ini langsung dibandingkan dengan nilai *r*-tabel. Bilamana koefisien $r_{ix} >$ nilai *r*-tabel, maka butir yang diuji *validitas*-nya dinyatakan *valid*, demikian juga sebaliknya. Pengujian dengan menggunakan formula ini sebetulnya masih mengandung efek *spurious overlap*, yakni suatu efek di mana dalam korelasi suatu butir tes, skor butir ikut dikorelasikan pada skor *total*-nya. Hal ini menyebabkan koefisien r_{ix} bersifat *overestimate* bila dibandingkan dengan koefisien r_{ix} yang sebenarnya. Koefisien r_{ix} yang sebenarnya harus dihitung dengan formula korelasi *Product Moment* yang dikoreksi oleh efek *spurious overlap* ($r_{i(x-i)}$). Selisih antara koefisien r_{ix} dan koefisien $r_{i(x-i)}$ merupakan koefisien efek *spurious overlap* (ESO). Koefisien ESO ini merupakan besarnya cecaran skor butir pada skor *total* suatu tes.

Seorang peneliti yang sedang melakukan penelitian tentang standarisasi suatu tes sering keliru, bahwasannya koefisien r_{ix} bisa digunakan secara langsung untuk

menentukan status *valid* atau *drop* suatu butir tes. Padahal koefisien r_{ix} sebetulnya hanya berfungsi sebagai *starting point* dari penentuan koefisien ESO. Koefisien r_{ix} digunakan untuk menentukan status *valid* dan *drop* suatu butir tes dengan tujuan agar butir tes yang sedang distandarisasi lebih banyak yang berstatus *valid*. Sebetulnya, yang absah digunakan untuk menentukan status *valid* dan *drop* suatu butir tes adalah koefisien $r_{i(x-i)}$.

Dalam menghitung *validitas* butir tes kemampuan berpikir mantik, digunakan rumus korelasi *Product Moment* yang dikoreksi oleh efek *spurious overlap* ($r_{i(x-i)}$). Untuk memperoleh koefisien $r_{i(x-i)}$ suatu butir, biasanya menggunakan standar deviasi *total* (SD_t) distribusi skor-tampak (Azwar, 2003). Standar deviasi *total* (SD_t) distribusi skor-tampak sebetulnya merupakan kontribusi skor-tampak dari masing-masing *item* yang menyusun tes kemampuan berpikir mantik.

Standar deviasi butir (SD_i) sering digunakan sebagai unsur pada formula indeks reliabilitas butir. Demikian juga koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* digunakan sebagai unsur pengompilasi formula indeks reliabilitas butir (*iri*). Secara utuh formula dari indeks reliabilitas butir (*iri*) dirumuskan sebagai: $iri = SD_i \times r_{ix}$. Di mana, *iri* adalah indeks reliabilitas item, SD_i merupakan standar deviasi butir ke-*i*, dan r_{ix} merupakan koefisien korelasi

sebelum dikoreksi oleh efek *spurious overlap*.

Oleh karena koefisien efek *spurious overlap* dan indeks reliabilitas butir bila dikaji dari asal-usulnya mengandung unsur koefisien korelasi *Product Moment* sebelum dan sesudah dikoreksi oleh efek *spurious overlap*, dan indeks reliabilitas butir (*iri*) juga ada unsur pengompilasi koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap*, maka secara derivatif dapat dikatakan memiliki hubungan dengan indeks reliabilitas butir (*iri*).

Bahkan Puger (2014) juga menyatakan indeks reliabilitas butir (*iri*) di samping sebagai karakteristik butir yang menunjukkan konsistensi butir tersebut dengan seluruh *item* yang menyusun tes yang bersangkutan, juga berperan sebagai dasar penentuan kontribusi suatu butir di dalam menyusun standar deviasi *total* (SD_t) distribusi skor-tampak. Hal ini disebabkan oleh pada dasarnya semua nilai *iri* yang menyusun suatu tes, misalnya tes Y besarnya sama dengan skor deviasi *total* (SD_t) distribusi skor-tampak.

Lebih lanjut Ernawati (2018), menyatakan sebetulnya indeks reliabilitas butir (*iri*) bukan hanya bisa digunakan sebagai dasar untuk memilih butir-butir tes yang bisa digunakan sebagai penyusun tes itu sendiri, tetapi juga bisa dikaitkan dengan koefisien ESO. Dari koefisien ESO dapat diramalkan mengenai besar kecilnya indeks reliabilitas butir. Keterkaitan antara ESO dengan

indeks reliabilitas butir (iri) merupakan hubungan secara terbalik. Dari bagian ini dapat dipaparkan bahwa makin kecil koefisien ESO akan diikuti oleh indeks reliabilitas butir yang makin besar, demikian juga sebaliknya. Pendapat Ernawati ini, didukung oleh pernyataan Gayatri (2020), bahwasannya indeks reliabilitas butir (iri) sebetulnya menggunakan koefisien r_{ix} sebagai salah satu penyusun formulanya. Oleh karena r_{ix} berkorelasi secara terbalik dengan koefisien efek *spurious overlap* (ESO), maka dapat diprediksikan indeks reliabilitas butir juga berkorelasi secara berbalik dengan koefisien ESO.

Lebih lanjut Lindawati (2019) menyatakan dalam menghitung efek *spurious overlap* (ESO) suatu butir tes, salah satu unsurnya menggunakan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* (r_{ix}) sebagai pengompilasi formulanya, sedangkan pada indeks reliabilitas butir (iri) juga menggunakan r_{ix} sebagai pembentuk formulanya. Hal ini berarti, baik ESO dan iri sama-sama menggunakan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* sebagai penyusun formulanya. Dari sini dapat diduga bahwa antara koefisien ESO terdapat hubungan dengan indeks reliabilitas butir (iri) suatu tes.

Mengenai hubungan suatu variabel dengan variabel lainnya, keputusan akhir bisa berhubungan dan tidak-berhubungan. Terkait dengan

keputusan ini, ada baiknya dikaji pernyataan Kusumawati dan Heryani. Kusumawati (2019) menyatakan bilamana dua variabel dikorelasikan, dan menghasilkan koefisien korelasi (r_{xy}) yang lebih kecil dari nilai r -tabel dan bernilai negatif, maka hubungan tersebut dikatakan sebagai hubungan yang terbalik. Jika penentuan koefisien korelasi (r_{xy}) ini digunakan sebagai dasar uji asumsi dari analisis jalur (*path analysis*), maka kedua variabel tersebut sebaiknya tidak dipasangkan sebagai pengaruh langsung maupun tidak langsung pada analisis jalur tersebut. Terkait dengan pendapat Kusumawati (2019), lebih lanjut Heryani (2019) menyatakan sebelum seorang peneliti memasang suatu variabel memiliki pengaruh langsung maupun pengaruh tidak langsung terhadap variabel yang lain, sebaiknya kedua variabel tersebut diuji dulu dengan mengorelasikan kedua variabel tersebut. Jika kedua variabel tersebut berkorelasi secara signifikan, sebaiknya kedua variabel tersebut dipasangkan dalam analisis jalur, sebaliknya jika koefisien korelasi dari kedua variabel tersebut tidak signifikan dan bernilai negatif, sebaiknya kedua variabel tersebut tidak dipasangkan dalam analisis jalur.

B. Metode Penelitian

Pada hakikatnya, penelitian ini merupakan jenis penelitian korelasional. Penelitian korelasional merupakan suatu jenis penelitian yang bertujuan untuk mengetahui

hubungan antara variabel yang satu dengan variabel yang lainnya. Hal ini dapat disimbolkan dalam bentuk koefisien r_{xy} (Puger, 2026). Hal ini disebabkan oleh tujuan dari penelitian ini adalah untuk mengetahui hubungan antara koefisien efek *spurious overlap* (ESO) dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik.

Lokasi penelitian ini adalah Sekolah Menengah Pertama (SMP) Negeri 2 Seririt. SMP Negeri 2 Seririt ini terletak di Desa Banjarasem, tepatnya di Jalan Menuju Desa Pangkung Paruk, Kecamatan Seririt, Kabupaten Buleleng. Dipilihnya lokasi penelitian ini, berdasarkan atas berbagai alasan, yakni: sekolah ini letaknya dekat dengan rumah peneliti, kemampuan siswanya termasuk kategori sedang, saat ini tidak ada peneliti yang menggunakan siswa kelas VIII sebagai subjek penelitian, kepala sekolahnya mengizinkan siswanya digunakan sebagai subjek penelitian, dan kemampuan siswanya tergolong sedang yang cocok untuk mengerjakan tes kemampuan berpikir mantik yang digunakan dalam penelitian ini.

Subjek dari penelitian ini adalah siswa kelas VIII SMP Negeri 2 Seririt yang berjumlah 100 siswa. Siswa tersebut yang akan memberikan respon tes kemampuan berpikir mantik, berupa jawaban terhadap tiap butir tes kemampuan berpikir mantik. Penentuan subjek sebanyak 100 siswa kelas VIII dalam penelitian ini menggunakan teknik *purposive sampling*.

Objek penelitian ini adalah jawaban dari tes kemampuan berpikir mantik. Jawaban tes kemampuan berpikir mantik ini digunakan sebagai dasar untuk menentukan koefisien korelasi *Product Moment* sebelum dilakukan koreksi oleh efek *spurious overlap* (r_{ix}), koefisien korelasi *Product Moment* setelah dilakukan koreksi oleh efek *spurious overlap* ($r_{i(x-i)}$), koefisien efek *spurious overlap* (ESO), dan indeks reliabilitas butir (iri).

Metode pengumpulan data yang digunakan adalah metode tes. Instrumen yang dibutuhkan untuk memperoleh data berupa jawaban tes adalah tes kemampuan berpikir mantik yang sudah distandarisasi. Tes tersebut terdiri atas 54 butir dan setelah distandarisasi berdasarkan validitas butirnya, ternyata hanya 49 butir yang bisa digunakan sebagai penyusun tes. Ke-49 butir tes kemampuan berpikir mantik tersebut harus diberikan jawaban oleh 100 siswa dalam waktu 60 menit.

Tes kemampuan berpikir mantik yang digunakan sudah dilakukan pengujian validitas isi dan reliabilitas respon antar-rater. Validitas isi dari tes kemampuan berpikir mantik ini sebesar 0,98, dan reliabilitas respon antar-raternya sebesar 0,81. Dalam ujicoba diperoleh 49 butir tes kemampuan berpikir mantik berkategori *valid*, dan reliabilitas tesnya sebesar 0,81. Oleh karena keempat persyaratan dari tes kemampuan berpikir mantik ini sudah terpenuhi, maka 49 butir tes kemampuan berpikir mantik sudah

feasibel digunakan dalam penelitian ini.

Data mengenai koefisien efek *spurious overlap* (ESO) dan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik yang berhasil dikumpulkan, selanjutnya dianalisis dengan analisis korelasi *Product Moment* (r_{xy}).

C. Hasil Penelitian dan Pembahasan

1. Hasil Penelitian

Data mengenai koefisien efek *spurious overlap* (ESO) dapat diperoleh dengan jalan mencari selisih antara koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* (r_{ix}) dengan koefisien korelasi *Product Moment* setelah dikoreksi oleh efek *spurious overlap* ($r_{i(x-i)}$), sedangkan data mengenai indeks reliabilitas butir (iri) diperoleh dengan jalan mengalikan standar deviasi butir (SD_i) dengan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* (r_{ix}).

Setelah data koefisien efek *spurious overlap* (ESO) dan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik disusun ulang pada microsoft-Excel selanjutnya dilakukan analisis data dengan formula korelasi *Product Moment* dari Pearson. Adapun formula yang diaplikasikan pada program microsoft-Excel adalah =CORREL(area ESO;area iri) atau =CORREL(C3:C51;\$D\$3:\$D\$51) Enter. Dari hasil analisis data pada

program microsoft-Excel dengan formula korelasi *Product Moment* diperoleh koefisien r_{xy} -hitung sebesar -0,233.

Koefisien r_{xy} -tabel pada jumlah sampel (n) sebanyak 49 dan taraf signifikansi 5% sebesar 0,281. Bila dibandingkan koefisien r_{ix} -hitung (-0,233) dengan koefisien r_{ix} -tabel sebesar 0,281, ternyata koefisien r_{ix} -hitung < koefisien r_{ix} -tabel dan bernilai negatif. Hal ini berarti koefisien r_{ix} -hitung yang diperoleh signifikan. Dengan kata lain, hipotesis nol (H_0) yang diajukan ditolak, atau menerima hipotesis alternatif (H_1). Oleh karena itu, dapat disimpulkan bahwa koefisien efek *spurious overlap* (ESO) berhubungan secara terbalik dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik.

2. Pembahasan

Temuan dalam penelitian ini adalah koefisien efek *spurious overlap* (ESO) berhubungan secara terbalik dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik.

Sudah dikemukakan bahwa kemampuan berpikir mantik siswa SMP dapat diukur dengan tes Kemampuan berpikir mantik. Skor-skor dari respon siswa terhadap tes kemampuan berpikir mantik dapat digunakan untuk menentukan koefisien efek *spurious overlap* (ESO), dan indeks reliabilitas butir (iri). Kajian pada bagian ini adalah mencari hubungan antara koefisien efek *spurious overlap* (ESO) dengan indeks reliabilitas butir (iri).

Untuk menghitung koefisien korelasi antara skor butir dengan skor total instrumen digunakan rumus statistika yang sesuai dengan jenis skor butir dari instrumen tersebut. Menurut Djaali dan Muljono (2008), untuk menghitung koefisien korelasi antara skor butir dengan skor *total* instrumen digunakan rumus statistika yang sesuai dengan jenis skor butir dari instrumen tersebut. Jika skor butir kontinum, maka untuk menghitung koefisien korelasi antara skor butir dengan skor *total* instrumen digunakan koefisien korelasi *Product Moment* (r_{ix}) dari Pearson.

Apabila koefisien korelasi *item-total* itu dihitung pada suatu skala yang berisi hanya sedikit *item*, maka sangat mungkin akan diperoleh koefisien korelasi *item-total* yang *overestimated* (lebih tinggi daripada yang sebenarnya) dikarenakan adanya *overlap* antara skor *item* dengan skor skala (Suwarni, 2007). *Overestimasi* ini dapat terjadi dikarenakan pengaruh kontribusi skor masing-masing *item* ikut menentukan besarnya skala.

Lebih lanjut Sutiman (2009) menyatakan makin sedikit *item* yang ada dalam skala akan makin besar *overlap* yang terjadi. Sebaliknya, makin banyak jumlah *item* dalam skala maka akibat yang ditimbulkan oleh *spurious overlap* makin kecil dan tidak signifikan. Sebagai pegangan kasar, bila jumlah *item* dalam skala lebih dari 30 buah maka umumnya efek

spurious overlap tidak begitu besar dan karenanya dapat diabaikan, sedangkan bila jumlah *item* dalam skala kurang dari 30 buah maka pengaruhnya menjadi substansial, sehingga perlu diperhitungkan. Untuk itu, agar kita memperoleh informasi yang lebih akurat mengenai korelasi antara *item* dengan skala, diperlukan suatu rumusan koreksi terhadap efek *spurious overlap*. Untuk menentukan koreksi dari efek *spurious overlap* digunakan formula koreksi terhadap efek *spurious overlap* ($r_{i(x-i)}$).

Proses menghilangkan efek *spurious overlap* pasti akan menurunkan nilai r_{ix} . Hal ini senada dengan saran dari Aryani (2009), bahwasannya jika pengujian validitas butir tes secara unidimensi tanpa dikoreksi oleh efek *spurious overlap*, sebaiknya menggunakan analisis korelasi *point biserial* (r_{pbi}). Namun, bila pengujian validitas butir tes secara unidimensi yang dikoreksi oleh efek *spurious overlap*, sebaiknya menggunakan analisis korelasi *Product Moment* (r_{ix}). Koefisien ESO dapat diperoleh melalui proses pengurangan koefisien r_{ix} dengan koefisien $r_{i(x-i)}$. Di mana, r_{ix} merupakan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap*, dan $r_{i(x-i)}$ merupakan koefisien korelasi *Product Moment* setelah dikoreksi oleh efek *spurious overlap*.

Puger (2018) menyatakan makin tinggi koefisien $r_{i(x-i)}$ yang diperoleh akan menyebabkan

koefisien ESO makin kecil, demikian juga sebaliknya. Hal ini didukung oleh pendapat Kresnawati (2019), yang pada hakikatnya menyatakan jika seseorang ingin menghitung ESO pada suatu tes, harus menghitung dulu koefisien r_{ix} dan $r_{i(x-i)}$. Selisih antara r_{ix} dan $r_{i(x-i)}$ merupakan koefisien ESO. Jika koefisien ESO dikorelasikan dengan koefisien r_{ix} akan diperoleh koefisien r_{xy} yang terbalik (atau koefisien r_{xy} dengan harga negatif). Hal ini berarti makin besar koefisien ESO akan diikuti oleh koefisien r_{ix} yang makin kecil, demikian juga sebaliknya. Lebih lanjut Suarni (2020), menyatakan nilai efek *spurious overlap* (ESO) sangat berkaitan secara terbalik dengan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* (r_{ix}) melalui koefisien korelasi *Product Moment* sesudah dikoreksi oleh efek *spurious overlap* ($r_{i(x-i)}$).

Standar deviasi butir (SD_i) sangat penting kedudukannya di dalam menentukan indeks reliabilitas butir (iri). Dalam kasus-kasus tertentu, misalnya penentuan kontribusi *item* dalam menyusun standar deviasi total (SD_t), indeks reliabilitas *item* mempunyai peran yang sangat vital. Bahkan dalam seleksi *item*, terutama untuk tes yang mengukur kemampuan yang disajikan dalam format *item* pilihan ganda, seringkali dilakukan dengan menggunakan kriteria indeks reliabilitas *item*.

Bahkan Azwar (2003) menyatakan indeks reliabilitas *item* harus dibedakan dari koefisien reliabilitas tes ($r_{xx'}$). Koefisien reliabilitas tes ($r_{xx'}$) merupakan indikator konsistensi *item-item* tes dalam menjalankan fungsi ukurnya secara bersama-sama, sedangkan indeks reliabilitas *item* (iri) merupakan karakteristik *item* yang diderivasi dari konsistensi satu *item* dengan keseluruhan *item* lainnya dalam tes yang bersangkutan.

Indeks reliabilitas *item* dirumuskan sebagai: $iri = SD_i \times r_{ix}$ (Suryani, 2017). Di mana, iri merupakan indeks reliabilitas *item*, SD_i sebagai standar deviasi *item* i , dan r_{ix} sebagai koefisien korelasi *Product Moment item* i dan skor total tes X . Sedangkan Azwar (2003) mengenalkan formula iri sebagai: $iri = s_i \times r_{ix}$. Di mana, iri = indeks reliabilitas *item*, s_i = standar deviasi *item* i , dan r_{ix} = koefisien korelasi *Product Moment item* i dan skor total tes X . Dari dua rumus yang dikemukakan oleh Suryani, dan Azwar, sebetulnya merupakan rumus yang sama, hanya saja Azwar menggunakan simbol standar deviasi *item* i dengan s_i .

Misalnya, diketahui standar deviasi untuk butir 3, 2, dan 1 adalah $s_3 = 0,2667$, $s_2 = 0,2667$, dan $s_1 = 0,1581$. Koefisien korelasi *Product Moment* untuk butir 3, 2, dan 1 adalah $r_{3x} = -0,2503$, $r_{2x} = -0,0970$, dan $r_{1x} = 0,3785$. Dengan demikian, iri nomor 1 = $0,1581 \times 0,3785 = 0,060$, iri untuk nomor 2 =

$0,2667 \times -0,0970 = -0,026$, dan iri untuk nomor 3 = $0,2667 \times -0,2503 = -0,067$. Iri nomor 1 sebesar 0,060 artinya konsistensi butir nomor 1 dengan keseluruhan butir lainnya dalam tes X sebesar 0,060. Iri nomor 2 sebesar -0,026 artinya konsistensi butir nomor 2 dengan keseluruhan butir lainnya dalam tes X sebesar -0,026. Iri nomor 3 sebesar -0,067 artinya konsistensi butir nomor 3 dengan keseluruhan butir lainnya dalam tes X sebesar -0,067.

Karena indeks reliabilitas *item* merupakan hasil perkalian antara standar deviasi *item* dengan korelasi *item*-total, dan karena standar deviasi *item* selalu berharga positif, maka indeks reliabilitas *item* akan negatif manakala koefisien korelasi *item*-totalnya negatif. Lebih lanjut, dikarenakan 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 pun tidak akan dapat lebih besar daripada 0,50. Keseluruhan indeks reliabilitas *item* ini apabila dijumlahkan maka besarnya akan sama dengan standar deviasi skor tes yang bersangkutan (Azwar, 2003).

Bahkan Puger (2025) juga menyatakan indeks reliabilitas butir (iri) di samping sebagai karakteristik butir yang menunjukkan konsistensi butir tersebut dengan seluruh *item* yang menyusun tes yang bersangkutan, juga berperan sebagai dasar

penentuan kontribusi suatu butir di dalam menyusun standar deviasi total (SD_t) distribusi skor tampak. Hal ini disebabkan oleh pada dasarnya semua nilai iri yang menyusun suatu tes, misalnya tes X besarnya sama dengan skor deviasi total (SD_t) distribusi skor-tampak.

Indeks reliabilitas *item* (iri) yang dihitung dengan formula $iri = SD_i \times r_{ix}$, merupakan sumbangan efektif di dalam mengompilasi standar deviasi total (SD_t) distribusi skor tampak tes kemampuan berpikir mantik. Hal ini disebabkan oleh total indeks reliabilitas *item* (iri) harus sama dengan skor deviasi total (SD_t) distribusi skor tampak. Hal ini sebagai bukti langsung bahwa indeks reliabilitas *item* (iri) sebetulnya masih mengandung cecar berupa *efek spurious overlap* (ESO). Efek *spurious overlap* (ESO) merupakan efek autokorelasi skor butir yang terdapat pada skor total. Hal ini menyebabkan koefisien korelasi (r_{ix}) bersifat *overestimate*, artinya menampakkan nilai koefisien yang lebih besar bila dibandingkan dengan yang sebenarnya. Mestinya, koefisien korelasi yang digunakan adalah koefisien korelasi yang sudah dikoreksi oleh efek *spurious overlap* ($r_{i(x-i)}$).

Lindawati (2019) menyatakan dalam menghitung efek *spurious overlap* (ESO) suatu butir tes, salah satu unsurnya menggunakan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek

spurious overlap (r_{ix}) sebagai pengompilasi formulanya, sedangkan pada indeks reliabilitas butir (iri) juga menggunakan r_{ix} sebagai pembentuk formulanya. Hal ini berarti, baik ESO dan iri sama-sama menggunakan koefisien korelasi *Product Moment* sebelum dikoreksi oleh efek *spurious overlap* sebagai penyusun formulanya. Dari sini dapat diduga bahwa antara koefisien ESO terdapat hubungan dengan indeks reliabilitas butir (iri) suatu tes.

Hal ini didukung oleh pendapat Puger (2014), yang menyatakan indeks reliabilitas butir (iri) di samping sebagai karakteristik butir yang menunjukkan konsistensi butir tersebut dengan seluruh *item* yang menyusun tes yang bersangkutan, juga berperan sebagai dasar penentuan kontribusi suatu butir di dalam menyusun standar deviasi total (SD_t) distribusi skor-tampak. Hal ini disebabkan oleh pada dasarnya semua nilai iri yang menyusun suatu tes, misalnya tes Y besarnya sama dengan skor deviasi total (SD_t) distribusi skor-tampak.

Bahkan Ernawati (2018), menyatakan sebetulnya indeks reliabilitas butir (iri) bukan hanya bisa digunakan sebagai dasar untuk memilih butir-butir tes yang bisa digunakan sebagai penyusun tes itu sendiri, tetapi juga bisa dikaitkan dengan koefisien ESO. Dari koefisien ESO dapat diramalkan mengenai besar kecilnya indeks reliabilitas butir.

Keterkaitan antara ESO dengan indeks reliabilitas butir (iri) merupakan hubungan secara terbalik. Dari hal ini dapat dipaparkan bahwa makin kecil koefisien ESO akan diikuti oleh indeks reliabilitas butir yang makin besar, demikian juga sebaliknya. Pendapat Ernawati ini, didukung oleh pernyataan Gayatri (2020), bahwasannya indeks reliabilitas butir (iri) sebetulnya menggunakan koefisien r_{ix} sebagai salah satu penyusun formulanya. Oleh karena r_{ix} berkorelasi secara terbalik dengan koefisien efek *spurious overlap* (ESO), maka dapat diprediksikan indeks reliabilitas butir juga berkorelasi secara berbalik dengan koefisien ESO.

Pendapat Ernawati dan Gayatri di atas, sebetulnya dapat dikonfirmasi melalui temuan empiris Puger (2021), bahwasannya: (1) koefisien efek *spurious overlap* (ESO) tidak berpengaruh secara langsung terhadap indeks reliabilitas butir (iri) tes kemampuan berpikir mantik, dan (2) koefisien efek *spurious overlap* (ESO) tidak berpengaruh secara tidak langsung terhadap indeks reliabilitas butir (iri) melalui standar deviasi butir (SD_i) tes kemampuan berpikir mantik.

Mengenai hubungan yang terbalik antara nilai efek *spurious overlap* (ESO) dengan indeks reliabilitas butir (iri), sebetulnya sehaluan dengan pernyataan Kusumawati (2019) dan Heryani (2019). Kusumawati (2019)

menyatakan bilamana koefisien korelasi (r_{xy}) ini digunakan sebagai uji asumsi dari analisis jalur (*path analysis*), maka variabel ESO dan iri sebaiknya tidak dipasangkan sebagai pengaruh langsung maupun tidak langsung pada analisis jalur tersebut. Sedangkan Heryani (2019), menyatakan oleh karena koefisien korelasi dari variabel ESO dan iri diprediksikan tidak signifikan dan bernilai negatif, sebaiknya kedua variabel tersebut tidak dipasangkan dalam analisis jalur.

D. Kesimpulan

Berdasarkan atas hasil analisis data dan pembahasan yang sudah dikemukakan dapat disimpulkan bahwa koefisien efek *spurious overlap* (ESO) berhubungan secara terbalik dengan indeks reliabilitas butir (iri) tes kemampuan berpikir mantik.

Temuan ini dapat digunakan sebagai dasar uji asumsi dalam pengujian model dengan analisis jalur (*path analysis*). Oleh karena korelasi antara koefisien efek *spurious overlap* (ESO) dengan indeks reliabilitas *item* (iri) tes kemampuan berpikir mantik menunjukkan nilai negatif (-0,233), maka kedua variabel tersebut tidak bisa dipasangkan dalam uji model, baik dalam hal pengaruh langsung maupun pengaruh tidak langsung.

DAFTAR PUSTAKA

- Aryani, Made. 2009. "Pengujian Validitas Butir dengan Mengoreksi Efek *Spurious Overlap*." *Makalah* yang Disampaikan dalam Seminar Ilmiah dengan Tema 'Pengembangan Instrumen Penelitian Menjadi Instrumen Standar,' yang Diselenggarakan Oleh K3S Kabupaten Jembrana, Tanggal 26 Juni 2009.
- Azwar, Saifuddin. 2003. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Azwar, Saifuddin. 2003. *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.
- Djaali, H. dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: PT Gramedia Widiasarana Indonesia.
- Ernawati, Ni Luh. 2018. *Penentuan Efek Spurious Overlap (ESO) dan Implikasinya pada Tes Hasil Belajar*. Surabaya: Paramita.
- Gayatri, Ni Luh. 2020. *Indeks Reliabilitas Item (iri) dan Pemilihan Butir Tes Hasil Belajar*. Melaya: CV Jimbarwangi.
- Heryani, Nyoman. 2019. *Structural Equation Modeling (SEM) dan Path Analysis*. Surabaya: Paramita.
- Kresnawati, I Nyoman. 2019. *Konstruksi dan Analisis Tes Hasil Belajar*. Denpasar: Bali Post Offset.
- Kusumawati, Ni Ketut. 2019. *Uji Asumsi dalam Analisis Data*. Jakarta: Kencana Prenada Media Group.
- Lindawati, Nila. 2019. *Efek Spurious Overlap (ESO) Butir pada Skor Skala*. Surabaya: Paramita.
- Nasoetion, Andi Hakim. 1995. "Berpikir Mantik." *Makalah* yang

- Disampaikan dalam Penataran Metodologi Penelitian Bagi Dosen Muda IPB, Agustus 1995.
- Puger, I Gusti Ngurah. 2026. *Metode Penelitian Pendidikan (Suatu Kajian Rasional dan Empiris)*. Padang: CV Dunia Penerbitan Buku.
- Puger, I Gusti Ngurah. 2025. "Sumbangan Efektif Indeks Reliabilitas Item (iri) di dalam Mengompilasi Standar Deviasi Total (SD_t) Distribusi Skor Tampak." *Orasi Ilmiah Pengukuhan Guru Besar dalam Bidang Evaluasi Pendidikan Bimbingan dan Konseling*, Sidang Senat Terbuka Universitas Panji Sakti Singaraja, 4 September 2025.
- Puger, I Gusti Ngurah. 2021. "Pengaruh Koefisien Efek Spurious Overlap (ESO), Standar Deviasi Butir (SD_i), dan Standar Deviasi Total (SD_t) Terhadap Indeks Reliabilitas Butir (iri) Tes Hasil Belajar Biologi pada Siswa Kelas VIII SMP Negeri 1 Singaraja." *Hasil Penelitian Hibah Kerjasama Antara Disdikpora Kabupaten Jembrana dengan FKIP Unipas Singaraja*.
- Puger, I Gusti Ngurah. 2019. *Validitas Butir dan Efek Spurious Overlap (ESO) Secara Unidimensi dan Multidimensi (Suatu Kajian Efek Spurious Overlap pada Tes Kemampuan Berpikir Mantik)*. Laporan Hasil Penelitian Hibah Kerjasama antara K3S Kabupaten Jembrana dengan FKIP Unipas Singaraja.
- Puger, I Gusti Ngurah. 2018. "ESO- r_{ix} Paradoks pada Tes Hasil Belajar." Makalah yang Disampaikan Kepada Guru-Guru SMP se-Kecamatan Melaya, dalam Seminar Ilmiah yang Diselenggarakan Oleh CV Jimbarwangi, Tanggal 19 Oktober 2018.
- Puger, I Gusti Ngurah. 2014. *Keterkaitan Indeks Reliabilitas Item (iri) dan Standar Deviasi Total (SD_t)*. Melaya: CV Jimbarwangi.
- Suarni, Ketut. 2020. *Pengujian Validitas Butir Tes Hasil Belajar dengan Formula Korelasi Product Moment yang Dikoreksi Oleh Efek Spurious Overlap*. Surabaya: Paramita.
- Suryani, Nina. 2017. *Analisis Skor Tes Hasil Belajar*. Surabaya: Paramita.
- Sutiman, Bambang. 2009. *Validitas dan Reliabilitas Tes Hasil Belajar*. Singaraja: Percetakan Mutiara.
- Suwarni, Wayan. 2007. *Standarisasi Tes Hasil Belajar*. Singaraja: Percetakan Mutiara.