

METODE-METODE STATISTIK UNTUK MENDETEKSI INDIKASI *COPYING ANSWER*: SUATU TINJAUAN TEORITIS

Nursalam

*) Dosen Jurusan Pendidikan Matematika UIN Alauddin Makassar

Email: nursalam_ftk@uin-alauddin.ac.id

ABSTRAK: *Asumsi yang digunakan untuk statistik uji yang dikembangkan dalam menjawab soal atau tes adalah bahwa seorang peserta tes akan menjawab suatu butir pertanyaan dengan 3 kemungkinan yaitu peserta menjawab karena mengetahui, peserta menjawab tes dengan menebak, dan peserta menjawab tes karena menyalin jawaban dari peserta lain. Akan tetapi jika peserta tes tidak memiliki akses untuk menyontek maka peserta hanya akan memberikan jawaban dengan asumsi pertama dan kedua. Artikel ini bertujuan untuk menjelaskan beberapa metode yang dapat digunakan untuk mendeteksi kecurangan dalam hal ini copying answer. Artikel ini memaparkan secara teoritis metode yang dapat digunakan untuk mendeteksi indikasi copying answer diantaranya B-index, H-index, K-index, ω -index, S_1 -index, dan S_2 index.*

Key Word: Metode Statistik, *Cheating*, *Copying Answer*

PENDAHULUAN

Idealnya selama pelaksanaan ujian berlangsung tidak terjadi kecurangan. Setiap peserta ujian diharapkan mampu memberikan respon terhadap setiap soal yang diberikan sesuai dengan kemampuan masing-masing. Hal ini diharapkan agar hasil ujian nantinya mampu memberikan informasi tentang kemampuan dari masing-masing peserta ujian. Namun, tidak bisa dipungkiri bahwa ketidaksiapan siswa menghadapi ujian dan tuntutan bahwa mereka harus lulus ujian (misalnya Ujian Nasional/UN) memaksa siswa, orang tua, guru, kepala sekolah, dan pejabat setempat melakukan kolaborasi untuk melakukan kecurangan. Hal yang menarik adalah bahwa kecurangan mungkin tidak hanya ditujukan untuk membantu siswa agar lulus UN, tetapi juga sebagai prestise agar ranking nilai UN sekolah atau daerah berada di papan atas tingkat nasional.

Ujian merupakan salah satu bagian dalam kegiatan proses pendidikan. Yaitu sebagai kegiatan yang digunakan untuk mengetahui atau mengukur keberhasilan dan kualitas suatu proses pembelajaran dalam lembaga pendidikan. Dari proses ujian itu akan dihasilkan angka-angka sebagai nilai yang menggambarkan prestasi siswa sebagai gambaran dari kualitas lembaga pendidikan tersebut. Selain itu, ujian bertujuan menilai pencapaian kompetensi lulusan pada mata pelajaran tertentu. Sebagai contoh adalah Ujian Nasional yang bertujuan untuk penilaian pencapaian kompetensi lulusan secara nasional mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan dan teknologi.

Pengujian atau penilaian memiliki peranan yang sangat penting dalam pendidikan saat ini. Hasil tes menjadi kekuatan utama dalam membentuk persepsi publik tentang kualitas sekolah, selain itu dapat dijadikan sebagai alat yang paling

penting dalam membuat kebijakan. Pendidikan menggunakan hasil penilaian untuk membantu meningkatkan pembelajaran dan untuk mengevaluasi program dan sekolah. Penilaian juga digunakan untuk menghasilkan data yang menjadi dasar keputusan atas kebijakan yang dibuat. Kerena penilaian memiliki peran yang sangat penting, maka penilaian pendidikan menjadi suatu kegiatan di setiap sekolah pada setiap daerah bahkan setiap negara karena menjadi komponen vital dalam inovasi, standar yang tinggi dan keunggulan pendidikan.

Hal yang penting berkaitan dengan skor tes adalah reliabel, akurat, dan valid dalam membuat suatu kesimpulan tentang pengetahuan seseorang, keterampilan atau kemampuan. Ada beberapa faktor yang dapat membatalkan skor tes, satu diantaranya adalah kecurangan pada saat tes. Kecurangan melanggar integritas akademik dan merupakan suatu ancaman yang sangat serius terhadap kualitas pendidikan.

Kecurangan dalam pelaksanaan tes dapat membatalkan keakuratan kesimpulan yang dibuat tentang pengetahuan dan kompetensi yang diukur (Cizak, 1999; Meijer, 1998). Hal ini menjadi masalah yang sangat serius dan perlu mendapatkan perhatian khusus dari sejumlah orang yang terlibat di dalamnya, baik itu siswa, orang tua, guru, lembaga pelaksanaan tes, dan pembuat kebijakan.

Ada berbagai cara untuk mengatasi masalah kecurangan, diantaranya adalah mengembangkan sikap yang benar di kalangan siswa atau mahasiswa, melakukan pencegahan atau dengan mendeteksi kecurangan yang dilakukan oleh peserta tes. Meskipun pencegahan dan pengembangan sikap yang benar di kalangan siswa atau mahasiswa dapat mengurangi kecurangan, namun hal yang penting adalah mendeteksi kecurangan untuk menjamin tes yang adil bagi pengambil tes.

Ada beberapa bentuk *cheating* diantaranya meniru pekerjaan teman (*copying answer*), bertanya langsung pada teman ketika sedang mengerjakan tes/ujian, membawa catatan pada kertas, pada anggota badan atau pada pakaian masuk ke ruang ujian, menerima kiriman jawaban dari pihak luar, mencari bocoran soal, arisan (saling tukar) mengerjakan tugas dengan teman, memberi lilin atau pelumas kepada lembaran jawaban komputer atau menebarkan atom magnet dengan maksud agar mesin scanner komputer dapat terkecoh ketika membaca lembar jawaban sehingga gagal mendeteksi jawaban yang salah atau menganggap semua jawaban benar.

Beberapa alasan siswa melakukan *cheating* adalah karena terpengaruh setelah melihat orang lain melakukan *cheating* meskipun pada awalnya tidak ada niat melakukannya. Terpaksa membuka buku karena pertanyaan ujian terlalu tekstual (buku sentris) sehingga memaksa peserta ujian harus menghafal kata demi kata dari buku teks, merasa dosen/guru kurang adil dan diskriminatif dalam pemberian nilai, adanya peluang karena pengawasan yang tidak ketat, takut gagal artinya yang bersangkutan tidak siap menghadapi ujian tetapi tidak mau menundanya dan tidak mau gagal, ingin mendapatkan nilai tinggi tetapi tidak bersedia mengimbangi dengan belajar keras atau serius, tidak percaya diri meskipun siswa bersangkutan sudah belajar teratur tetapi ada kekhawatiran akan lupa lalu akan menimbulkan kefatalan, sehingga perlu diantisipasi dengan membawa catatan kecil, terlalu cemas menghadapi ujian sehingga hilang ingatan sama sekali lalu terpaksa buka buku atau bertanya kepada teman yang duduk berdekatan, merasa sudah sulit menghafal atau mengingat karena faktor usia, sementara soal yang dibuat penguji sangat menekankan kepada kemampuan mengingat, mencari jalan pintas dengan pertimbangan daripada mempelajari sesuatu yang belum tentu keluar lebih baik mencari bocoran soal, menganggap sistem penilaian tidak objektif, sehingga pendekatan pribadi kepada dosen/guru lebih efektif daripada belajar serius,

penugasan guru/dosen yang tidak rasional yang mengakibatkan siswa/mahasiswa terdesak sehingga terpaksa menempuh segala macam cara, yakin bahwa guru tidak akan memeriksa tugas yang diberikan berdasarkan pengalaman sebelumnya sehingga bermaksud membalas dengan mengelabui guru yang bersangkutan.

Diantara berbagai metode kecurangan yang ada, dalam kajian ini difokuskan pada bagaimana mendeteksi indikasi menyalin jawaban (*copying answer*) dari suatu situasi bahwa seorang peserta tes menyalin jawaban peserta tes yang lain. Untuk mendeteksi menyalin jawaban pada tes pilihan ganda, ada dua metode yang dapat digunakan yaitu **metode observasi** dan **metode statistic**. Metode observasi menggunakan pengamatan manusia untuk mendeteksi menyalin jawaban yang dilakukan oleh peserta tes seperti peserta tes satu berbicara dengan peserta tes yang lain selama ujian berlangsung atau mengumpulkan bukti-bukti fisik misalnya adanya lembar jawaban yang dipertukarkan antara dua peserta tes. Sementara metode *statistic* dilakukan dengan melakukan pemodelan probabilitas respon dari peserta ujian dengan asumsi tidak ada kecurangan dan mencari pola jawaban yang sama antar peserta tes.

Metode mendeteksi *copying* yang berdasarkan pada teori peluang adalah G_2 index (Frary, Tideman, & Watts, 1977) dan ω -index (Wollack, 1997). Kedua metode ini dikembangkan berdasarkan asumsi bahwa setiap peserta tes mempunyai peluang menjawab benar atau salah suatu soal yang diberikan berdasarkan tingkat kemampuan peserta tes dan tingkat kesukaran soal tersebut. Dalam kondisi tidak terjadi kecurangan, peluang peserta tes menjawab benar pada soal-soal mudah dan menjawab salah pada soal-soal sukar akan lebih besar. Kedua metode ini sama, perbedaannya yang pertama berdasarkan teori tes klasik dan yang satu lagi berdasarkan teori tes modern (*Item Response Theory*). Contoh dari kategori kedua adalah K-index (Holland, 1996). Dasar dari metode ini adalah membandingkan pola jawaban pasangan-pasangan peserta tes yang dicurigai dengan distribusi jawaban dari pasangan-pasangan peserta tes yang disebut *benchmark group*. *Benchmark group* adalah kelompok peserta tes yang tidak mungkin (sangat kecil kemungkinannya) menyontek satu sama lain (melakukan kecurangan) karena mereka berada di lokasi tes yang berbeda.

Metode lain yang dikembangkan untuk mendeteksi indikasi adanya *copying answer* selain yang disebutkan di atas adalah B indeks dan H indeks (Angoff, 1974), K_2 index (Sotaridona & Meijer, 2002), S_1 indeks dan S_2 indeks (Sotaridona & Meijer, 2003), Kappa Statistic (Sotaridona, van der Linden, & Meijer, 2006). Metode yang dikembangkan dalam mendeteksi *copying answer* berdasarkan pada teori tes klasik, dengan membandingkan jawaban yang sama antara satu peserta tes dengan peserta tes yang lain.

Salah satu hal yang perlu dicatat bahwa bukti *statistic* tidak boleh digunakan sebagai bukti hanya untuk menuduh peserta tes melakukan suatu kecurangan, tetapi perlu di dukung oleh bukti-bukti yang lain. Salah satu hal yang menarik dalam hal ini adalah pernyataan yang dikemukakan oleh Teamster v U.S.: Goog (2001) bahwa:

“Statistics are not irrefutable; they come in infinite variety and, like any other kind of evidence, they may be rebutted. In short, their usefulness depends on all of the surrounding facts and circumstances (Teamster v. U.S.: Good, 2001, p. 13).”

Fakta-fakta dan keadaan dalam situasi pengujian mungkin adalah pengamatan dari kecurangan atau perbedaan yang sangat besar antara skor tes yang diujikan

dalam dua administrasi tes yang serupa. Sehingga beberapa pemicu yang mendahului penggunaan metode statistika perlu dilakukan untuk meminimalkan perilaku kecurangan.

Cheating dan Metode Mendeteksi Indikasi *Copying Answer*

Kecurangan adalah bertindak tidak jujur atau tidak adil untuk mendapatkan keuntungan (Ehrlich, Flexner, Carruth, & Hawkins, 1980, hal 141). Sedangkan Cizek (2003) mengemukakan bahwa perilaku curang dibagi dalam tiga kategori yaitu (1) member, mengambil, atau menerima informasi tertentu, (2) menggunakan suatu alat yang dilarang, (3) memanfaatkan kelemahan orang, prosedur, proses untuk mendapatkan keuntungan. Berdasarkan kedua definisi di atas, dapat disimpulkan bahwa kecurangan adalah suatu perbuatan yang dilakukan dengan cara-cara yang tidak baik untuk mendapatkan keuntungan.

Masalah kecurangan peserta tes dalam menjawab soal-soal pilihan ganda bertahun-tahun telah menjadi perhatian para ahli di bidang penilaian pendidikan (*educational measurement*). Mereka mengembangkan metode statistik untuk mendeteksi siapa-siapa saja yang diduga melakukannya. Metode ini dibagi menjadi dua kategori. Masing-masing metode berdasarkan teori peluang menjawab soal dan metode berdasarkan perbandingan pola jawaban antarpeserta tes.

a. Metode B-index dan H-index

William Angoff (1974) mengembangkan suatu prosedur untuk mendeteksi indikasi *copying answer* dengan membandingkan statistic uji dari dua pasangan peserta tes yang bertentangan dengan distribusi sampling dari peserta tes yang tidak terlibat dalam menyalin yaitu “pasangan yang jujur” setelah menstratifikasi peserta tes berdasarkan variabel-variabel independen tertentu. Angoff memulai dengan mendefinisikan variabel-variabel yang digunakan sebagai berikut:

- R_jR_m : jumlah item yang dijawab benar oleh masing-masing peserta tes
- R_{jm} : jumlah item yang dijawab benar dari dua peserta tes
- w_jw_m : jumlah item yang dijawab salah oleh masing-masing peserta tes
- w_{jm} : jumlah item yang dijawab salah dari dua peserta tes
- $h(w)_{jm}$: jumlah item dengan jawaban salah yang sama dari dua peserta tes
- O_jO_m : jumlah jawaban yang omit dari setiap peserta tes
- O_{jm} : jumlah item yang omit dari dua peserta tes
- w_j : $\min(w_j, w_m)$
- O_j : jumlah item yang omit untuk peserta tes dengan skor yang tinggi
- S_j : $w_j + O_j$
- S_{jm} : $Q_{jm} + O_{jm}$
- K_{jm} : panjang jawaban yang salah berturut-turut yang identik dan ommit

Pengembangan untuk mendeteksi indikasi *copying answer* dapat dilihat pada tabel berikut:

Tabel 1. Angoff's Copying Indices

Copying Index	Covariate	Dependent Variabel
A	R_jR_m	R_{jm}
B	w_jw_m	$h(w)_{jm}$
C	w_{jm}	$h(w)_{jm}$
D	O_jO_m	O_{jm}
E	w_j	$h(w)_{jm}$

F	O_j	O_{jm}
G	S_j	S_{jm}
H	S_j	K_{jm}

B-index (Anngoff, 1974) membahas tentang deteksi adanya copying answer. Prinsip kerja dari metode ini adalah berdasarkan pada banyaknya item salah yang identik atau sama (Q_{ij}). Setelah itu mengecek banyaknya jawaban yang salah antara copier dengan source (W_iW_j). Setelah itu membuat stratifikasi peserta agar lebih homogen. Nilai B indeks diperoleh dengan membandingkan antara Q_{ij} untuk pasangan peserta tes dengan rerata Q semua pasangan peserta tes dalam setiap strata ke-ij ($\bar{Q}_{W_iW_j}$) dan membagi dengan standar deviasi pada Q untuk setiap strata ke-ij ($S_{Q_{W_iW_j}}$).

Persamaan B index dapat dituliskan sebagai berikut:

$$B = \frac{Q_{ij} - \bar{Q}_{W_iW_j}}{S_{Q_{W_iW_j}}}$$

Asumsi B berdistribusi normal. Nilai B yang besar mengindikasikan adanya *copying answer*.

H index (Angoff, 1974) menyelidiki tentang jumlah maksimum jawaban salah yang identik atau menghilangkan item-item pada setiap respon yang sama (K_{CS}) setelah pengkondisian pada banyaknya jawaban yang salah dari C atau S.

H index dapat dihitung dengan menggunakan persamaan berikut:

$$H = \frac{K_{CS} - \bar{K}_+}{S_+}$$

Dimana \bar{K}_+ adalah rata-rata sampel dan S_+ adalah standard deviasi. H diasumsikan berdistribusi normal. Nilai H yang besar mengindikasikan adanya *copying answer*.

b. Metode K-Index

K index adalah suatu metode statistik yang dapat digunakan untuk menilai tingkat kesepakatan antara jawaban-jawaban yang salah dari dua orang peserta tes untuk tes pilihan ganda. Satu orang disebut *source* (s) dan lainnya disebut yang mengkopi atau *copier* (c). Dimana copier menyalin jawaban dari sumber (Holland, 1996). K index hanya mengambil jawaban yang salah dari peserta ujian. Berikut beberapa notasi yang digunakan oleh Satoridona dan Meijer (2002), yaitu:

Misalkan:

$j = (1, 2, \dots, J)$ melambangkan jumlah peserta tes (examinees)

$i = (1, 2, \dots, I)$ melambangkan butir tes (items)

$v = (1, 2, \dots, V)$ melambangkan kategori respon item (item response categories)

s = melambangkan peserta tes yang teridentifikasi sebagai sumber (source)

c = melambangkan peserta tes yang dicurigai melakukan *copying answer* dari s

W_j = melambangkan jumlah jawaban salah pada peserta tes j

M (dengan direalisasikan m) = melambangkan banyaknya jawaban yang sama tetapi jawaban yang salah antara peserta tes j dan s

$r = 1, 2, \dots, c', \dots, R$ melambangkan subgroup peserta tes, dimana setiap grup memiliki perbedaan jumlah jawaban yang salah, dan c' adalah subgroup dimana peserta tes c berada.

$j' = 1, 2, \dots, n_r$ melambangkan peserta tes dalam subgroup r , dimana setiap subgroup memiliki minimal 1 peserta tes dan $\sum_{r=1}^n n_r$.

$M_r = (M_{r1}, M_{r2}, \dots, M_{rn})$ melambangkan vector jumlah jawaban salah dan sama dalam setiap subgroup r

$M_{c'} = (M_{c'1}, M_{c'2}, \dots, M_{c'n})$ melambangkan vector jumlah pencocokan jawaban yang salah pada peserta tes $n_{c'}$ dalam subgroup c' dimana c' terdiri dari peserta tes yang memiliki jumlah jawaban sama yang salah sebagai hasil dari proses menyalin jawaban.

$Q_r = \frac{w_r}{I}$ melambangkan proporsi jawaban yang salah pada subgroup r dan I adalah banyaknya jumlah item dalam tes.

Untuk $j' = 1, 2, n_{c'}$, andaikan:

$$I = \begin{cases} 1 & m_{c'j'} \geq m_{c'c} \\ 0 & \text{lainnya} \end{cases}$$

dimana:

$m_{c'c}$ adalah pencocokan jumlah jawaban salah antara penyontek dan sumber.

$m_{c'j'}$ adalah pencocokan jumlah jawaban salah antara penyontek dan peserta tes j' dalam subgroup

maka K index didefinisikan sebagai berikut:

$$K = \frac{\sum_{j'=1}^{n_{c'}} I_{c'j'}}{n_{c'}}$$

Jika nilai K sangat kecil, maka ada indikasi adanya copying answer yaitu c menyalin jawaban dari s . Secara umum, jumlah pencocokan nilai yang salah bergantung pada tingkat kemampuan (*ability*) s dan c . Jumlah jawaban yang sama dan salah akan sangat kecil ketika s atau c keduanya memiliki skor nilai yang benar (kemampuan tinggi), sedangkan keduanya akan memiliki banyak jumlah jawaban yang salah (kemampuan rendah).

c. Metode ω indeks

Salah satu pendekatan teori respon butir untuk mendeteksi adanya copying answer adalah dengan menggunakan pendekatan ω indeks. Metode ω indeks (Wollack, 1997) menggambarkan perbedaan standard antara jumlah jawaban yang sama (baik yang benar dan salah) antara sepasang peserta tes (h_{CS}), dan banyaknya prediksi secara kebetulan mengkondisikan jawaban S , C adalah level kemampuan, dan parameter item dapat dituliskan sebagai berikut:

$$\omega = \frac{h_{CS} - \sum_{i=1}^n P_C(u_{iS})}{\sigma_{hCS}}$$

dengan u_{iS} adalah respons item S ke item i , dan

$$\sigma_{hCS} = \sqrt{\sum_{i=1}^n P_C(u_{iS})[1 - P_C(u_{iS})]}$$

$P_C(u_{iS})$ adalah peluang peserta tes dengan kemampuan θ_C melakukan contek (copier). Pemilihan respon S ke item i (u_{iS}), dan ω berbeda dengan index lain yang dalam hal ini berdasarkan pada Item Response Theory (Lord, 1980; Lord & Novick, 1968) dan estimasi probabilitas menggunakan nominal response model (Bock, 1972) yaitu:

$$P_C(u_{iS}) = \frac{\exp(\zeta_{iS} + \lambda_{iS}\theta_C)}{\sum_{m=1}^K \exp(\zeta_{im} + \lambda_{im}\theta_C)}$$

dimana λ dan ζ adalah item slope dan parameter intercept, dan m ($m = 1, 2, \dots, k$) yang melambangkan pemilihan kategori item pada item i . ω diasumsikan berdistribusi normal. Jika nilai ω besar, maka indikasi adanya cheating.

Perlu diketahui bahwa metode K-index dan ω -index keduanya memanfaatkan kesamaan respon item, namun antara metode K-indeks dan ω -index memiliki perbedaan yaitu ω -index membandingkan respon *copier* (c) dengan seluruh tanggapan *source* (s), sedangkan K-indexx membandingkan respon yang salah antara *copier* (c) dengan respon yang salah dari *source* (s). Model *Item Response Theory*, khususnya untuk model respon nominal, memerlukan ukuran sampel yang cukup besar untuk mengestimasi parameter model (De Ayaala & Sava-Bolesta, 1999; Wollack, Bolt, Cohen, & Lee, 2002). Hal ini dimaksudkan agar nantinya diperoleh nilai ω yang *robust*. (Wollack, 2003; Wollack & Cohen, 1998). Wollack (1997) mengemukakan bahwa *power* statistic yang tidak memperhitungkan informasi dari respon dari item yang benar kemungkinan akan menurun karena pengurangan jumlah item yang digunakan dalam melakukan analisis.

d. Metode S₁ index

Metode S₁ index (Satoridona & Meijer, 2003) secara konseptual sama dengan metode K₂ index. Ada dua hal penting yang membedakan antara metode S₁ dengan K₂. Metode K₂ menggunakan distribusi binomial, sedangkan metode S₁ menggunakan distribusi Poisson. Selain itu, metode K₂ mengestimasi parameter p dengan regresi linear, sedangkan metode S₁ mengestimasi parameter μ menggunakan model loglinear yaitu $\log(\mu_r) = \beta_0 + \beta_1 w_r$, untuk setiap r , dimana w_r adalah banyaknya jawaban yang salah untuk setiap peserta dengan r adalah jawaban benar. Estimasi persamaan loglinear untuk mengevaluasi w_C yaitu banyaknya jawaban yang salah pada peserta yang menyontek yang akan memberikan suatu hasil taksiran parameter Poisson yaitu $\hat{\mu}$. Oleh karena itu S₁ index diberikan oleh persamaan berikut:

$$S_1 = \sum_{g=W_{CS}}^{W_S} \frac{e^{-\hat{\mu}} \mu^g}{g!}$$

Jika nilai S₁ kecil, maka indikasi adanya *cheating*.

e. Metode S₂ index

Metode S₂ (Sotaridona & Meijer, 2003) merupakan perluasan dari metode S₁ untuk memasukkan informasi dari jawaban benar yang identik. Secara statistic S₂ memuat m_{CS} yaitu banyaknya jawaban salah yang identik ditambah dengan jumlah jawaban benar yang identik. Sotaridona dan Meijer (2003) menyatakan mendukung penggunaan informasi dari kesamaan jawaban yang benar sesuai dengan kemungkinan jawaban yang sama. Statistic m_{CS} dituliskan dalam bentuk persamaan berikut:

$$m_{CS} = w_{CS} + \sum_i \delta_{ijr} = w_{CS} + \sum_i (d_1) e^{d_2 P_{ijr}} I_{(u_{iS}=u_{i^*})}$$

Dimana i^* melambangkan jawaban benar untuk item ke i ; $I_{(u_{iS}=u_{i^*})}$ adalah fungsi indicator yang sama dengan 1 jika S menjawab item ke i dengan benar dan 0 untuk yang lainnya; P_{ijr} adalah persentase peserta tes dalam grup r

Persamaan S₂ index adalah:

$$S_2 = \sum_{g=m_{CS}}^I \frac{e^{-\hat{\mu}} \mu^{\hat{g}}}{g!}$$

μ adalah parameter poisson yang diestimasi dengan menggunakan persamaan loglinear. Jika nilai S_2 kecil, maka indikasi adanya *cheating*.

PENUTUP

Untuk menghasilkan kualitas hasil tes yang baik, maka security tes perlu mendapat perhatian khusus. Hal ini dimaksudkan agar hasil tes yang dihasilkan mampu mengungkap dan mengemukakan kemampuan yang dimiliki oleh masing-masing peserta tes. Untuk mendapatkan informasi yang akurat maka diasumsikan tidak terjadi kecurangan selama pelaksanaan tes. Namun kenyataannya masih banyak terjadi kecurangan selama pelaksanaan tes khususnya copying answer. Oleh karenanya dikembangkan suatu metode pendeteksian untuk mengetahui indikasi *copying answer* tersebut. Metode itu menggunakan pendekatan teori tes klasik dan teori respon butir.

Metode yang dikemukakan diatas semuanya cenderung pada mendeteksi indikasi copying answer antara dua peserta tes dengan melihat respon siswa yang salah dan sama. Metode yang ada belum mampu mendeteksi adanya kecurangan copying answer dalam kelompok peserta tes (dalam arti kelas, sekolah, kotamadya/kabupaten, dan/atau provinsi). Oleh karena perlu mengembangkan suatu teknik khusus untuk mendeteksi hal tersebut, misalnya dengan menerapkan Computerized Adaptive Test (CAT) untuk melakukan *screening* atas respon peserta tes.

REFERENSI

- Belov, D. I. & Armstrong, R. D. (2010). Automatic Detection of Answer Copying via Kullback-Leibler Divergence and K-Index. *Applied Psychological Measurement*, 34(6), 379 – 392.
- Nath, L. & Lovaglia, M. (2009). Chating on Multiple-Choice Exams. *College Teaching*, 57(1), 3 – 8.
- Satoridona, L. & Meijer, R.R. (2002). Statistical Properties of the K-Index for Detecting Answer Copying. *Journal of Educational Measurement*, 39(2), 115 – 132.
- Satoridona, L. & Meijer, R.R. (2003). Two New Statistics to Detect Answer Copying. *Journal of Educational Measurement*, 40(1), 53 – 69.
- Satoridona, L., Van der Linden, W.J., & Meijer, R.R. (2006). Detecting Answer Copying Using The Kappa Statistic. *Applied Psychological Measurement*, 30(5), 412 – 431.
- van der Ark, L.A., Emons, W.H.M., & Sijtsma, K. (2008). Detecting Answer Copying Using Alternate Test Forms and Seat Locations in Small-Scale Examinations. *Journal of Educational Measurement*, 45(2), 99 – 117.
- van der Linden, W.J. & Satoridona, L. (2004). A Statistical Test for Detecting Answer Copying on Multiple-Choice Tests. *Journal of Educational Measurement*, 41(4), 361 – 377.
- van der Linden, W.J. & Satoridona, L. (2006). Detecting Answer Copying When The Regular Response Process Follows a Known Response Model. *Journal of Educational and Behavioral Statistics*, 31(3), 283 – 304.

- van der Linden, W. J. (2009). A Bivariate Lognormal Response-Time Model for the Detecting of Collusion Between Test Takers. *Journal of Educational and Behavioral Statistics*, 34(3), 378 – 394.
- Wesolowsky, G.O. (2000). Detecting Excessive Similarity in Answer on Multiple Choice Exams. *Journal of Applied Statistics*, 27(7), 909 – 921.
- Wollack, J. A., & Cohen, A. S. (1998). Detection of answer copying with unknown item and trait parameters. *Applied Psychological Measurement*, 22, 144–152.
- Wollack, J. A., Cohen, A.S., & Serlin, R.C. (2001). Defining Error Rates and Power For Detecting Answer Copying. *Applied Psychological Measurement*, 25(4), 385 – 404.
- Wollack, J. A., Bolt, D. M., Cohen, A. S., & Lee, Y.-S. (2002). Recovery of item parameters in the nominal response model: A comparison of marginal maximum likelihood estimation and Markov chain Monte Carlo estimation. *Applied Psychological Measurement*, 26(3), 339–352.
- Wollack, J. A. (2003). Comparison of Answer Copying Indices With Real Data. *Journal of Educational Measurement*, 40(3), 189 – 205.
- Wollack, J. A. (2006). Simultaneous Use of Multiple Answer Copying Indexes to Improve Detection Rates. *Applied Measurement in Education*, 19(4), 265 – 288.