

KETIDAKWAJARAN SKOR SELEKSI PENERIMAAN MAHASISWA BARU

Oleh:

Ahsanul Khair Asdar

STABN Sriwijaya Tangerang Banten

ahsanul.khair@stabn-sriwijaya.ac.id

ABSTRACT

In educational measurement, the inappropriate scores must be detected. The inappropriate scores can occur in respondents individually. The inappropriate scores become an interesting focus because of its the information can provide input to the tester about the condition of the test takers even more in decision making. This study focused on the description of the inappropriate scores of 2018 STABN Sriwijaya Tangerang Bantens' students candidate on the basics of Buddhism test. The population of this study was all of the 2018 STABN Sriwijaya Tangerang Bantens' students candidate and at the same time as the samples. The data of this study were secondary data. They were the responses of the studentscandidate on each item in the basics of Buddhism test. The results of data analysis by using Donlon-Fisher method indicated that there were as many as 20.93% of students candidate who had anappropriate score while 79.07% of students candidate had an inappropriate score.

Keywords: *inappropriate score, Donlon-Fisher method, basic of Buddhism test, index of difficulty*

PENDAHULUAN

Pendidikan merupakan hubungan antara pribadi seorang pendidik dengan peserta didik dalam kaitannya terjadi kontak atau komunikasi antara masing-masing pribadi yang pada akhirnya akan melahirkan tanggung jawab dan kewajiban pendidikan. Pasal 1 Undang-Undang Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional, dijelaskan bahwa pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara.

Dilanjutkan pada Pasal 3 bahwa pendidikan nasional berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab. Hal tersebut kemudian ditegaskan kembali di dalam Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 Bab II Pasal 4 bahwa Standar Nasional Pendidikan bertujuan menjamin mutu pendidikan

nasional dalam rangka mencerdaskan kehidupan bangsa dan membentuk watak serta peradaban bangsa yang bermartabat.

Pendidikan bagi seorang anak merupakan suatu proses yang tidak akan pernah putus mulai dari awal bahkan dari dalam kandungan sampai dengan pendidikan tertinggi pada perguruan tinggi, sehingga dalam rangka mendukung program pemerintah dalam pemerataan pendidikan nasional, STABN Sriwijaya Tangerang Banten sebagai salah satu lembaga pendidikan tinggi yang berada di bawah naungan Kementerian Agama Republik Indonesia mengambil bagian dalam penyelenggaraan pendidikan bagi anak bangsa. Sebagai upaya pemerolehan masukan (*input*) yang berkualitas, STABN Sriwijaya Tangerang Banten menyelenggarakan seleksi penerimaan mahasiswa baru dengan fungsi sebagai “proses untuk mengetahui dan memetakan kemampuan awal serta kemampuan dasar calon mahasiswa dalam rangka memprediksi potensi akademik calon mahasiswa” serta sebagai “proses tes/seleksi terhadap para calon mahasiswa yang diprediksi mampu menyelesaikan studi tepat waktu dengan prestasi yang baik”. Tentu saja kedua fokus tersebut dapat dicapai jika alat tes (instrumen) yang digunakan adalah instrumen yang valid dan reliabel. Sekalipun alat tes (instrumen) yang digunakan telah memenuhi persyaratan valid dan reliabel, skor calon mahasiswa tentu saja juga akan memengaruhi pengambilan keputusan terkait dua fungsi seleksi penerimaan mahasiswa yang telah dikemukakan. Tentu saja skor yang diharapkan dimiliki oleh masing-masing calon mahasiswa adalah skor yang wajar. Bagaimana jika seandainya skor yang dimiliki adalah skor yang tidak wajar?

Sebagaimana dikemukakan oleh Ansori (2017: 142) bahwa hasil tes yang diperoleh dari suatu tes yang telah dilaksanakan masih dapat dipertanyakan, apakah angka-angka tersebut murni hasil usaha siswa? Apakah angka tersebut terbebas dari ketidakwajaran pengukuran? Sehingga dengan dasar tersebut, angka-angka yang diperoleh calon mahasiswa STABN Sriwijaya Tangerang Banten pada proses seleksi penerimaan mahasiswa baru tidak serta-merta dapat dijadikan sebagai dasar untuk memenuhi kedua fungsi yang telah dikemukakan sebelumnya. Sebab hampir semua tes tidak pernah terlepas dari pengaruh ketidakwajaran skor atau bias.

Salah satu bentuk ujian yang digunakan dalam proses seleksi penerimaan calon mahasiswa baru STABN Sriwijaya Tangerang Banten adalah ujian tertulis berbentuk pilihan ganda (*multiple choice*) dengan materi ujian meliputi dasar-dasar agama Buddha, pengetahuan umum, bahasa Indonesia, bahasa Inggris, matematika dasar, dan wawasan kebangsaan. Hal ini mengacu pada pendapat Santoso (2011: 134) bahwa salah satu bentuk tes tipe objektif yang dikenal dan sering digunakan adalah tes objektif berbentuk pilihan ganda. Bentuk tes tersebut dapat mengukur lebih efektif jika dibandingkan dengan bentuk tes jawaban singkat dan bentuk tes betul salah maupun bentuk tes menjodohkan. Sejalan dengan hal tersebut, Ebel dan Frisbie (1991: 154) juga berpendapat bahwa jenis tes yang paling sering dan umum digunakan adalah tes pilihan ganda. Hal ini dikarenakan tes pilihan ganda mudah untuk mengukur

beberapa aspek seperti pengetahuan, pemahaman, dan keputusan tentang kemampuan pemecahan masalah, merekomendasikan tindakan yang sesuai atau membuat prediksi.

Robb dan Ercanbrack (1999: 20) bahwa keragaman skor tes yang dicapai siswa disebabkan oleh sejumlah faktor, yaitu: (1) pengalaman atau latihan tes, sehingga siswa merasa terbiasa, senang, dan percaya diri dalam menyelesaikan tes; (2) ketidakwajaran pengukuran, sering ditemukan siswa yang menunjukkan peningkatan skor pada pengukuran berikutnya kendatipun tidak terdapat peningkatan dalam penguasaan sasaran ukur tes, dan pada pengukuran lainnya mungkin menunjukkan penurunan skor secara drastis. Dengan demikian, ketidakwajaran skor pada tes merupakan skor yang tidak memberikan informasi yang benar tentang hal-hal yang dimaksud untuk diukur. Ketidakwajaran skor tes merupakan ketidakwajaran skor yang bersumber dari responden yang tidak memberikan informasi yang sebenarnya, mungkin lebih rendah atau mungkin pula lebih tinggi dari yang seharusnya dicapai tentang hal-hal yang diukur dalam suatu tes.

Sekalipun secara teoretis telah dikemukakan bagaimana ketidakwajaran skor bisa terjadi, dalam kenyataannya berdasarkan hasil observasi dan wawancara terbatas yang telah dilakukan diperoleh informasi bahwa penetapan calon mahasiswa yang dinyatakan lulus hanya didasarkan pada skor yang diperoleh pada masing-masing ujian (tertulis, wawancara, dan praktik). Tidak dilakukan penelusuran lebih lanjut terhadap skor mereka dalam ujian tulis yang memuat keenam materi ujian yang telah dikemukakan sebelumnya. Padahal dalam kenyataannya skor pada suatu tes ada kalanya tidak memberikan informasi yang benar tentang peserta tes. Mungkin saja informasi tersebut tidak menjangkau sampai kepada besaran atau dimensi yang hendak diukur oleh tes tersebut. Mungkin pula, hasil tes itu tercampur dengan besaran atau dimensi lain yang tidak dimaksudkan untuk diukur oleh tes itu sehingga hasil tes menjadi rancu.

Ketidakwajaran skor menjadi satu fokus kajian yang menarik dikarenakan informasi yang diberikan dapat memberikan masukan kepada pelaksana tes terkait kondisi peserta tes terlebih lagi dalam pengambilan keputusan. Beberapa penelitian sebelumnya terkait ketidakwajaran skor yang pernah dilakukan antara lain penelitian yang telah dilakukan oleh Nur Hayati Setyaningsih pada tahun 2017 yang mana dalam penelitiannya diperoleh hasil bahwa perbedaan jenis kelamin/gender tidak memberikan pengaruh pada ketidakwajaran skor pemahaman Matematika dan metode Donlon Fisher lebih sensitif jika dibandingkan dengan metode Jacob dalam pendeteksian ketidakwajaran skor peserta didik laki-laki dan perempuan pada instrumen pemahaman Matematika ruang lingkup Aljabar dan Geometri di Sekolah Menengah Atas (SMA). Hal ini sejalan dengan hasil penelitian yang telah dilakukan oleh Suciati Rahayu Widyastuti pada tahun 2015 yang memperoleh hasil bahwa tidak terdapat perbedaan antara metode SHL dan metode Donlon-Fisher dalam hal indeks ketidakwajaran skor pada hasil belajar matematika SMP kelas VII.

Sementara penelitian yang telah dilakukan oleh Ari Hasan Ansori pada tahun 2017 memberikan hasil bahwa untuk meningkatkan indeks ketidakwajaran skor tes matematika metode penyekorannya dilakukan dengan menggunakan metode *penalty score*, untuk meningkatkan indeks ketidakwajaran skor tes matematika pada siswa yang memiliki kecenderungan kecemasan tes kategori ringan dapat dilakukan dengan menggunakan metode penyekoran *penalty score*, dan untuk meningkatkan indeks ketidakwajaran skor tes matematika pada siswa yang memiliki kecenderungan kecemasan tes kategori berat dapat dilakukan dengan menggunakan metode penyekoran *correct score*.

Hasil penelitian yang telah dilakukan oleh Budi Santoso pada tahun 2011 memberikan gambaran bahwa pada tingkat kemampuan rendah, penggunaan model penskoran *correct score* pada soal dengan lima pilihan menghasilkan skor yang lebih wajar dibandingkan soal dengan tiga pilihan jawaban serta pada tingkat kemampuan rendah penggunaan ketiga model penskoran pada soal dengan lima pilihan tidak menghasilkan perbedaan skor wajar. Tetapi pada tingkat kemampuan tinggi ada perbedaan skor wajar yang terjadi, khususnya model penskoran *penalty score* dan *reward score* menghasilkan skor yang lebih wajar dibandingkan model penskoran *correct score*. Selain digunakan untuk memberikan informasi terkait kondisi peserta tes, identifikasi ketidakwajaran skor juga dapat digunakan untuk melakukan standarisasi penilaian sebagaimana hasil penelitian yang telah dilakukan oleh Dadan Rosana dan Sukardiyono pada tahun 2015.

Dengan memperhatikan struktur materi ujian pada proses seleksi penerimaan mahasiswa baru STABN Sriwijaya Tangerang Banten dan karakteristik STABN Sriwijaya Tangerang Banten yang diketahui sarat akan keagamaan Buddha, maka peneliti tertarik untuk menganalisis indeks ketidakwajaran skor mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha dengan menggunakan metode Donlon-Fisher. Pemilihan materi ujian dasar-dasar agama Buddha sebagai bahan kajian di dalam penelitian ini didasarkan pada persepsi bahwa pemahaman calon mahasiswa terkait dasar-dasar agama Buddha akan mengambil peran dengan porsi yang sangat besar dalam menentukan keberhasilan calon mahasiswa selama menempuh pendidikan di STABN Sriwijaya Tangerang Banten. Harapan yang muncul dari proses seleksi tersebut adalah diperolehnya gambaran yang objektif terkait kondisi yang dimiliki oleh calon mahasiswa, sehingga pada akhirnya pengambil kebijakan dapat memetakan kemampuan yang dimiliki serta menentukan calon mahasiswa mana saja yang diprediksi dapat menyelesaikan pendidikan tepat waktu. Sehingga fokus permasalahan di dalam penelitian ini adalah bagaimanakah indeks ketidakwajaran skor calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 pada materi ujian dasar-dasar agama Buddha?

KAJIAN TEORI

Tes

Tes dapat dipandang sebagai perangkat pertanyaan yang memiliki jawaban benar atau salah. Surapranata (2005: 19) menjelaskan bahwa tes adalah sehimpunan pertanyaan yang harus dijawab, atau pertanyaan-pertanyaan yang harus dipilih, ditanggapi, atau tugas-tugas yang harus dilakukan oleh *testee* (orang yang diuji) dengan tujuan untuk mengukur suatu aspek (perilaku/atribut). Pada umumnya tes dikembangkan untuk mengukur aspek-aspek perilaku manusia seperti aspek pengetahuan, sikap, maupun aspek keterampilan.

Dalam dunia pendidikan selanjutnya, tes dapat digunakan untuk meningkatkan pembelajaran. Melalui tes, seorang guru dapat memperoleh informasi tentang berhasil tidaknya peserta didik dalam menguasai tujuan-tujuan pembelajaran yang telah ditetapkan. Melalui tes, seorang guru dapat dengan mudah mendeteksi peserta didik yang telah menguasai tujuan pembelajaran dan peserta didik yang belum menguasai tujuan pembelajaran. Melalui tes, seorang guru dapat mendeteksi berhasil tidaknya pembelajaran yang telah dilaksanakan. Hasil tes dapat digunakan untuk memberikan laporan tentang perkembangan kemajuan belajar peserta didik maupun tentang keberhasilan guru dalam mengajar (Surapranata, 2005: 19).

Sementara menurut Nurkancana (1983: 25), tes adalah suatu cara untuk mengadakan penilaian yang berbentuk suatu tugas atau serangkaian tugas yang harus dikerjakan oleh anak atau sekelompok anak sehingga menghasilkan suatu nilai tentang tingkah laku atau prestasi anak tersebut, yang dapat dibandingkan dengan nilai yang dicapai oleh anak-anak yang lain dengan nilai standar yang ditetapkan. Dari definisi tersebut, dapat diperoleh tiga hal yaitu tes dapat berupa tugas, serangkaian tugas atau pertanyaan yang harus dikerjakan, tes diberikan kepada anak atau sekelompok anak, dan jawaban atau respon yang diberikan melalui tes dinilai.

Beberapa ahli lain yang mengemukakan pendapat mengenai pengertian tes di antaranya Kerlinger (1973: 492) yang mengatakan bahwa tes adalah seperangkat rangsangan (stimulus) yang diberikan kepada seseorang dengan maksud untuk mendapatkan jawaban yang dapat dijadikan dasar bagi penetapan skor angka atau pun dijadikan dasar mengenai apa yang diukur. Demikian juga pendapat Cronbach yang dikutip oleh Sudijono (1996: 66) bahwa tes merupakan suatu prosedur yang sistematis untuk membandingkan tingkah laku dua orang atau lebih.

Dalam pengembangan sebuah tes, Karmel dan Karmel (Sudaryono, 2014: 90) menyatakan bahwa terdapat sepuluh kriteria sebuah tes yang baik yaitu: (1) tes harus relevan, (2) ada keseimbangan antara tujuan yang ingin dicapai dan jumlah butir tes yang mewakilinya, (3) efisiensi waktu yang digunakan untuk melakukan tes, penskoran, dan pengadministrasian tes, (4) objektivitas dalam memberikan skor dan interpretasinya, (5) kekhususan tes yang mengukur materi pelajaran yang diajarkan di kelas, (6) tingkat kesukaran setiap butir tes,

(7) kemampuan butir membedakan kelompok siswa yang memiliki kemampuan tinggi dan rendah, (8) reliabilitas, (9) kejujuran dan pemerataan kesempatan, dan (10) kecepatan menyelesaikan tes.

Dari beberapa pendapat yang dikemukakan di atas, terdapat kesamaan pengertian tentang tes yaitu alat yang digunakan untuk mengukur kemampuan atau kecakapan seorang individu pada aspek tertentu baik yang tampak maupun yang tidak tampak dan hasilnya berupa angka atau skor.

Tes Pilihan Ganda

Dalam penyusunan suatu tes, seorang pembuat soal harus memilih satu dari sekian banyak bentuk tes yang ada. Dari sekian bentuk tes yang ada, bentuk tes objektif lebih banyak digunakan. Hal ini dikarenakan penilaian dengan menggunakan bentuk tes ini lebih objektif. Dengan kata lain jika penilaian dilakukan oleh penilai yang berbeda, maka tetap akan menghasilkan nilai yang sama. Bentuk tes objektif terdiri atas pilihan ganda, betul-salah, dan menjodohkan (Gronlund, 1977: 34). Sementara itu, Mehrens dan Lehmann (1991: 107) menjelaskan bahwa seluruh tes objektif membuat siswa bekerja pada situasi terstruktur yang komplis dan meminta siswa untuk menjawab sejumlah pertanyaan. Meskipun demikian seluruh bentuk tes objektif yang ada akan menciptakan penilaian yang objektif, bahkan penilaian selain bentuk tes isian singkat atau melengkapi dapat dilakukan penilaian dengan bantuan mesin (komputer).

Pernyataan yang senada juga disampaikan oleh Lizza (Sudaryono, 2014: 90) bahwa yang menyebabkan tes pilihan ganda banyak digunakan di berbagai tingkat dan jenjang pendidikan adalah karena: (1) tipe tes disusun dan digunakan untuk mengukur semua standar kompetensi, mulai dari yang paling sederhana hingga yang kompleks, (2) jumlah alternatif jawaban lebih dari dua sehingga dapat mengurangi keinginan siswa untuk menebak, (3) tipe tes ini menuntut kemampuan siswa untuk membedakan berbagai tingkat kebenaran sekaligus, dan (4) tingkat kesukaran butir soal dapat dikendalikan dengan hanya mengubah tingkat homogenitas alternatif jawaban. Berdasarkan pemaparan tersebut tidak mengherankan jika bentuk tes yang paling lazim digunakan adalah bentuk tes objektif yaitu pilihan ganda.

Hal tersebut sejalan dengan penjelasan Kubiszyn dan Borich (2003: 107) bahwa tes pilihan ganda merupakan salah satu bentuk tes yang paling populer digunakan di dalam lingkup pendidikan. Tes jenis ini lebih banyak digunakan pada jenjang pendidikan yang tinggi termasuk universitas daripada di jenjang pendidikan dasar. Hal ini tidak berarti bahwa tes pilihan ganda tidak cocok untuk digunakan pada jenjang pendidikan dasar. Hanya saja memerlukan perhatian khusus mengenai penggunaannya pada siswa yang lebih muda.

Tes pilihan ganda merupakan suatu tes yang terdiri dari suatu pernyataan yang belum lengkap. Untuk melengkapi pernyataan tersebut disediakan beberapa pernyataan sambungan. Satu di antaranya adalah sambungan yang benar, sedang yang lain adalah sambungan yang salah. Tes pilihan ganda juga dapat berupa suatu pertanyaan yang telah disediakan

beberapa jawaban yang mana satu di antara jawaban-jawaban yang disediakan merupakan jawaban yang benar. Pertanyaan atau pernyataan yang belum lengkap dalam tes pilihan ganda disebut dengan *stem*. Alternatif pilihan yang disediakan disebut *option*. Jawaban-jawaban atau sambungan yang salah disebut pengecoh (distraktor) (Nurkencana, 1983: 31-32). ahsan

Hal senada juga dikemukakan oleh Haladyna (2004: 68) bahwa bentuk umum dari tes pilihan ganda adalah bersifat konvensional. Komponen utama dari soal pilihan ganda terdiri atas tiga bagian, yaitu *stem*, pilihan jawaban yang benar, dan pilihan jawaban yang salah biasa juga disebut distraktor. Setiap distraktor harus dapat memberikan peluang kepada peserta tes yang tidak memiliki pengetahuan atau keterampilan yang cukup baik untuk memilihnya. Sementara peserta tes dengan pengetahuan yang baik akan mengetahui bahwa distraktor merupakan pilihan yang salah. Pendapat yang lainnya dikemukakan oleh Osterlind (2002: 162) bahwa format tes pilihan ganda terdiri atas butir-butir dengan suatu *stem* berbentuk kalimat atau pernyataan yang menyatakan suatu permasalahan atau menanyakan sesuatu dengan dilengkapi dengan dua sampai lima alternatif pilihan yang satu alternatif pilihan menjawab *stem*.

Dari beberapa pemaparan mengenai tes pilihan ganda diperoleh kesamaan bahwa tes pilihan ganda merupakan tes objektif dengan butir-butir yang terdiri atas *stem* berupa pernyataan atau pertanyaan dengan dilengkapi beberapa alternatif pilihan jawaban yang salah satunya merupakan pilihan jawaban yang benar sementara lainnya merupakan pengecoh (distraktor).

Ketidakwajaran Skor

Dalam pengukuran pendidikan, ketimpangan skor yang mungkin terjadi perlu dicegah atau dihindari. Ketimpangan skor tersebut perlu dideteksi dan dari hasil deteksi ini diambil keputusan yang tepat tentang apa yang harus dilakukan terhadap hasil yang dicapai dalam pengukuran itu. Skor dicapai oleh peserta tes setelah mengerjakan sejumlah butir ujian, maka ada dua sumber tempat ketimpangan dapat terjadi masing-masing adalah peserta tes dan perangkat tes beserta butir-butirnya. Dengan demikian, ketimpangan skor pada tes dapat dipandang dari kombinasi antara letak ketimpangan dan bentangan ketimpangan. Ketimpangan itu dapat terjadi (a) pada peserta tes atau butir tes dan (b) secara individu atau secara kelompok (Santoso, 2011: 137).

Ketimpangan skor dapat terjadi pada responden secara perorangan. Responden yang biasanya memperoleh skor tinggi, secara timpang memperoleh skor rendah. Ketimpangan ini dikenal sebagai ketidakwajaran skor. Ketidakwajaran skor (*inappropriateness*) terjadi pada responden secara perorangan sehingga sangat kaitannya dengan kondisi/keadaan responden tersebut. Ada beberapa sebab yang menghasilkan ketidakwajaran skor pada responden. Penyebab tersebut antara lain kecemasan, ketidakhati-hatian, kurang sehat, atau tidak terbiasa dengan sistem baru yang digunakan responden ketika proses tes berlangsung (Naga, 2013: 391). Hal ini juga

dijelaskan oleh Nitko (1996: 91-94) bahwa tekanan mental peserta tes, seperti perasaan cemas, khawatir, takut akan gagal, ketidakmampuan dalam menulis, dan perasaan tidak disukai, dapat menyebabkan peserta tes tidak berhasil menjawab secara benar butir-butir tes. Sebagai akibatnya, peserta tes seperti ini akan memperoleh skor yang tidak tepat, yakni tidak sesuai dengan kemampuan mereka yang sebenarnya.

Drasgow (1987: 59-79) mengemukakan bahwa ketidakwajaran skor terjadi ketika peserta tes gagal menghasilkan pengukuran yang tepat tentang hal yang diukur oleh tes yang dimaksud. Sementara Hulin, *et. al.* (1983: 110) mengemukakan bahwa ketidakwajaran skor (*inappropriateness*) adalah bila seorang peserta tes dengan kemampuan tinggi salah dalam menjawab soal mudah dan peserta tes dengan kemampuan rendah banyak menjawab benar soal-soal yang sulit, seorang peserta tes tidak menjawab terlalu banyak soal yang mudah, atau seorang peserta tes menjawab secara acak keseluruhan tes. Dengan demikian, ketidakwajaran skor pada tes dapat dipandang sebagai kombinasi mereka yakni dari kombinasi di antara letak yang sukar sehingga berhasil menjawabnya dengan benar, tetapi menjadi sangat ceroboh ketika berhadapan dengan butir yang mudah sehingga menjawabnya dengan salah. Sebagai akibatnya, ditemukan skor yang tidak wajar pada peserta tes itu.

Tidak hanya itu, Rosana dan Sukardiyono (2015: 132) juga mengemukakan bahwa penyebab lain dari ketidakwajaran pengukuran adalah kecurangan saat pelaksanaan tes, misalnya ada peserta didik yang menyontek. Pola jawaban dari peserta didik yang curang dalam tes mungkin akan tampak ganjil. Kelompok jawaban benar akan bercampur dengan kelompok jawaban yang hampir semuanya tidak benar. Jenis pola jawaban seperti ini berbeda pada setiap peserta ujian. Keadaan ini dapat dideteksi melalui analisis ketidakwajaran pengukuran. Berdasarkan hal tersebut, Widyastuti (2015: 74) mengemukakan bahwa kecurangan tersebut lebih mudah lagi dilakukan apabila tes matematika berbentuk tes pilihan ganda, dan juga memberi peluang untuk menebak jawaban benar jika mereka tidak menguasai materi pada butir tes tersebut. Jawaban siswa yang diidentifikasi terdapat ketidakwajaran skor akan terlihat sangat ganjil dan bisa terlihat setelah skor selesai dianalisis, soal diurutkan dari yang tingkat kesukarannya rendah ke tingkat kesukarannya tinggi, kemudian terlihatlah ketidakwajaran skor ketika terdapat peserta tes yang dapat menjawab benar butir soal sulit sedangkan butir soal mudah tidak mampu dijawab benar. Apabila sudah ada indikasi bahwa skor peserta tes itu tidak wajar, maka penting dilakukan pendeteksian ketidakwajaran skor, hal ini dilakukan apabila kita sudah tidak dapat mencegah atau menghindari ketidakwajaran skor tersebut, karena itulah fungsi dari pendeteksian ketidakwajaran skor, yaitu suatu metode untuk mendeteksi skor peserta tes yang dicurigai tidak wajar.

Berdasarkan hal tersebut di atas, maka dapat ditarik kesimpulan bahwa ketidakwajaran skor adalah adanya ketidaksesuaian skor yang diperoleh dengan kemampuan yang dimiliki yang diakibatkan oleh adanya kesalahan

dalam menjawab butir soal dengan taraf sukar mudah namun betul dalam menjawab soal yang sukar.

METODE PENELITIAN

Penelitian ini merupakan penelitian yang menggunakan metode deskriptif dengan pendekatan kuantitatif. Penelitian ini dilakukan tidak untuk menguji hipotesis penelitian, melainkan untuk memberikan deskripsi atau penggambaran terkait data yang dikumpulkan. Adapun objek di dalam penelitian ini adalah ketidakwajaran skor mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 pada bidang tes dasar-dasar agama Buddha.

Populasi di dalam penelitian ini terdiri atas populasi subjek dan populasi atribut. Populasi subjek di dalam penelitian ini adalah seluruh mahasiswa angkatan 2018 STABN Sriwijaya Tangerang Banten. Sementara populasi atribut di dalam penelitian ini adalah seluruh ketidakwajaran skor mahasiswa angkatan 2018 saat proses seleksi penerimaan mahasiswa baru yang telah dilaksanakan pada tahun 2018 pada bidang tes dasar-dasar agama Buddha. Dikarenakan hasil pekerjaan (respon) seluruh mahasiswa STABN Sriwijaya Tangerang Banten angkatan 2018 saat seleksi penerimaan mahasiswa baru pada materi ujian dasar-dasar agama Buddha masih dapat dijangkau oleh peneliti, maka keseluruhan populasi di dalam penelitian ini dijadikan sebagai sampel penelitian. Sehingga teknik *sampling* yang digunakan di dalam penelitian ini adalah teknik *non-probability sampling*, yaitu *sampling* jenuh atau dikenal juga dengan sensus.

Jenis data yang digunakan di dalam penelitian ini adalah data sekunder. Pengumpulan data yang dimaksudkan dalam tahapan ini adalah menghimpun seluruh data yang telah diperoleh sebelumnya oleh panitia seleksi penerimaan mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 dalam bentuk respon (jawaban A, B, C, D, atau E) calon mahasiswa terhadap masing-masing butir pada materi ujian dasar-dasar agama Buddha.

Analisis data penelitian diawali dengan melakukan analisis butir soal pada tes dasar-dasar agama Buddha menggunakan software ITEMAN 4.3. agar diperoleh gambaran taraf sukar untuk setiap butir yang dilanjutkan dengan perhitungan indeks ketidakwajaran skor menggunakan metode Donlon-Fisher. Naga (2013: 394-396) menjelaskan bahwa pada tahun 1968, T.F. Donlon dan F.E. Fisher mengemukakan suatu metode yang dikenal sebagai metode Donlon Fisher. Mereka menganggap bahwa distribusi taraf sukar seluruh butir mengikuti probabilitas normal sehingga taraf sukar masing-masing butir dapat dinyatakan dalam skala delta (Δ).

Untuk semua butir yang masing-masing memiliki taraf sukar butir π_b , secara total diperoleh rerata dan simpangan baku taraf sukar butir dalam skala delta yaitu μ_Δ dan σ_Δ . Untuk setiap responden ke-g berdasarkan butir yang dijawab betul diperoleh rerata dan simpangan baku dalam skala delta yaitu $\mu_{\Delta g}$ dan $\sigma_{\Delta g}$. Sehingga indeks ketidakwajaran Donlon-Fisher dinyatakan sebagai

korelasi biserial di antara setiap skor responden dengan total sebagai mana disajikan sebagai berikut

$$\rho_{g-bis} = \frac{\mu_{\Delta} - \mu_{\Delta g} \pi_g}{\sigma_{\Delta} \gamma_{zg}}$$

dengan:

π_g : proporsi jawaban betul responden ke-g

z : nilai probabilitas normal baku berdasarkan nilai $(1 - \pi_g)$. Nilai ini diperoleh dengan menggunakan tabel fungsi distribusi bawah pada distribusi probabilitas normal baku

Δ : taraf sukar butir dalam skala delta yang diperoleh dengan rumus $\Delta = 4z + 13$

μ_{Δ} : rerata taraf sukar seluruh butir dalam skala delta

$\mu_{\Delta g}$: rerata taraf sukar seluruh butir dalam skala delta berdasarkan butir yang dijawab betul oleh responden ke-g

σ_{Δ} : simpangan baku taraf sukar butir dalam skala delta

γ_{zg} : densitas pada distribusi probabilitas normal baku di titik z yang dicapai pada π_g . Nilai ini diperoleh dari tabel fungsi densitas distribusi probabilitas normal baku

Pendeteksian ketidakwajaran skor menggunakan metode Donlon-Fisher, perhitungannya menggunakan korelasi biserial, acuan dalam menyatakan tingginya suatu indeks kewajaran skor peserta tes sebagai pembanding dapat menggunakan indeks daya beda butir soal yang dinyatakan oleh Ebel dan Frisbie (1991: 232) yaitu: indeks daya beda antara 0,70 – 1,00 dinyatakan butir memiliki daya beda baik sekali; 0,40 – 0,69 dinyatakan butir memiliki daya beda cukup baik; 0,30 – 0,39 dinyatakan butir memerlukan revisi sedikit atau tidak; 0,20 – 0,29 dinyatakan butir memerlukan revisi atau disisihkan; dan 0,00 – 0,19 dinyatakan butir direvisi total atau disisihkan.

Klasifikasi indeks daya beda pada besaran 0,70 sampai 1,00 digunakan sebagai batasan indikator yang menyatakan skor peserta tes wajar, karena mempertimbangkan korelasi antara pola kesukaran butir soal responden secara individu dengan pola.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Pada bagian ini disajikan hasil analisis data yang meliputi hasil analisis taraf sukar terhadap butir yang terdapat pada instrumen seleksi penerimaan mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha dan hasil analisis ketidakwajaran skor seleksi penerimaan mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha.

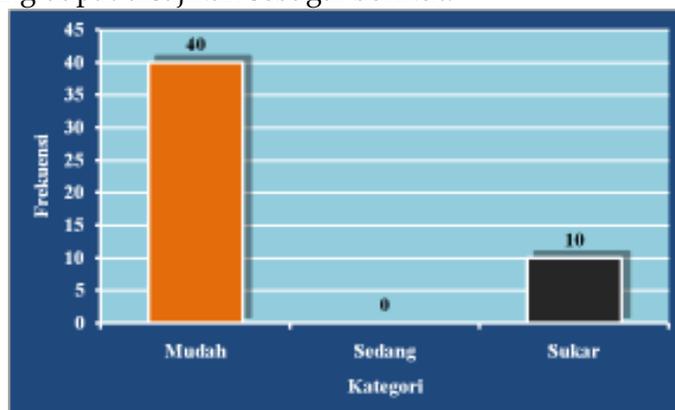
Berdasarkan hasil analisis terhadap taraf sukar 50 butir soal dalam tes dasar-dasar agama Buddha tahun 2018 dengan menggunakan *software*

ITEMAN 4.3. diperoleh hasil distribusi kategori taraf sukar butir tes sebagaimana disajikan pada Tabel 1 sebagai berikut.

Tabel 1. Distribusi Taraf Sukar Butir Tes Dasar-Dasar Agama Buddha Tahun 2018

No	Kategori	Frekuensi	Persentasi
1.	Mudah	40	80,00
2.	Sedang	0	0,00
3.	Sukar	10	20,00
Jumlah		50	100

Berdasarkan Tabel 1 di atas, maka dapat diperoleh informasi bahwa dari 50 butir soal yang ada, terdapat sebanyak 40 butir soal atau 80,00% yang memiliki taraf sukar butir dalam kategori mudah, sebanyak 10 butir soal atau 20,00% yang memiliki taraf sukar butir sukar, sementara tidak terdapat satupun butir dengan taraf sukar sedang. Secara visualisasi dalam bentuk diagram batang dapat disajikan sebagai berikut.



Gambar 1. Diagram Batang Distribusi Taraf Sukar Butir

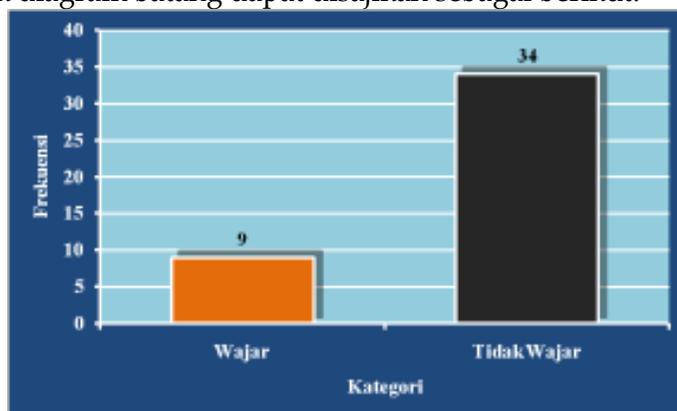
Sementara berdasarkan hasil pengujian ketidakwaajaran skor penerimaan mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha dengan menggunakan metode Donlon-Fisher, diperoleh hasil sebagai berikut.

Tabel 2. Distribusi Ketidakwaajaran Skor Mahasiswa Baru STABN Sriwijaya Tangerang Banten Tahun 2018

No	Kategori	Frekuensi	Persentasi
1.	Wajar	9	20,93
2.	Tidak Wajar	34	79,07
Jumlah		43	100

Berdasarkan Tabel 2 di atas, maka dapat diperoleh informasi bahwa dari 43 calon mahasiswa baru STABN Sriwijaya Tangerang Banten tahun 2018, terdapat sebanyak 9 orang atau 20,93% calon mahasiswa yang memiliki skor

wajar dalam tes dasar-dasar agama Buddha sementara sebanyak 34 orang atau 79,07% calon mahasiswa yang memiliki skor tidak wajar dalam tes dasar-dasar agama Buddha. Hal ini menunjukkan bahwa pada umumnya calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 pada bidang tes tes dasar-dasar agama Buddha berada dalam kategori tidak wajar. Secara visualisasi dalam bentuk diagram batang dapat disajikan sebagai berikut.



Gambar 2. Diagram Batang Distribusi Ketidakwaian Skor

Pembahasan

Pada umumnya tes dikembangkan untuk mengukur aspek-aspek perilaku manusia seperti aspek pengetahuan, sikap, maupun aspek keterampilan. Dalam dunia pendidikan selanjutnya, tes dapat digunakan untuk meningkatkan pembelajaran. Melalui tes, seorang guru dapat memperoleh informasi tentang berhasil tidaknya peserta didik dalam menguasai tujuan-tujuan pembelajaran yang telah ditetapkan. Melalui tes, seorang guru dapat dengan mudah mendeteksi peserta didik yang telah menguasai tujuan pembelajaran dan peserta didik yang belum menguasai tujuan pembelajaran. Melalui tes, seorang guru dapat mendeteksi berhasil tidaknya pembelajaran yang telah dilaksanakan. Hasil tes dapat digunakan untuk memberikan laporan tentang perkembangan kemajuan belajar peserta didik maupun tentang keberhasilan guru dalam mengajar (Surapranata, 2005: 19).

Hasil yang diperoleh dari suatu tes yang telah dilaksanakan masih dapat dipertanyakan, apakah angka-angka tersebut murni hasil usaha peserta tes? Apakah angka tersebut terbebas dari ketidakwaian pengukuran? Apakah hasil tersebut murni merupakan pengetahuan peserta tes? Kenyataannya, capaian peserta tes tersebut tidak dapat digunakan sepenuhnya sebagai indikator keberhasilan peserta tes. Hal ini dikarenakan hampir semua tes tidak pernah terlepas pada pengaruh ketidakwaian pengukuran atau bias.

Tes memiliki peranan penting dalam pengukuran hasil belajar. Tes dapat digunakan untuk menentukan prestasi siswa, menentukan persyaratan ujian, merencanakan dan mengevaluasi hasil belajar, menyeleksi, menempatkan siswa dalam kelas tertentu sesuai kompetensi, serta untuk berbagai tujuan lainnya. Namun demikian, terdapat sejumlah kritikan terhadap penggunaan tes baku atau tes prestasi, terutama jika hasil tes tersebut digunakan untuk

memilih mahasiswa di perguruan tinggi dan untuk menempatkan siswa dalam kelas tertentu.

Ada kalanya skor pada hasil tes itu tidak memberikan informasi yang benar tentang peserta tes. Mungkin saja informasi itu tidak menjangkau sampai ke besaran atau dimensi yang hendak diukur oleh tes itu. Mungkin pula, hasil tes itu tercampur dengan besaran atau dimensi lain yang tidak dimaksudkan untuk diukur oleh tes itu sehingga hasil tes menjadi rancu. Mungkin pula, pelaksanaan tes itu sendiri kurang layak sehingga menghasilkan informasi yang tidak benar. Apa pun yang terjadi pada pengujian semacam ini, di dalam hal ini, kita berhadapan dengan skor yang tidak benar atau skor yang timpang. Dengan demikian, ketidakwajaran skor pada tes merupakan skor yang tidak memberikan informasi yang benar tentang hal-hal yang dimaksud untuk diukur. Ketidakwajaran skor tes merupakan ketidakwajaran skor yang bersumber dari responden yang tidak memberikan informasi yang sebenarnya, mungkin lebih rendah atau mungkin pula lebih tinggi dari yang seharusnya dicapai tentang hal-hal yang diukur dalam suatu tes.

Sebagaimana hasil yang telah disajikan pada bagian sebelumnya, terlihat bahwa terdapat sebanyak 9 orang atau 20,93% calon mahasiswa yang memiliki skor wajar dalam tes dasar-dasar agama Buddha sementara sebanyak 34 orang atau 79,07% calon mahasiswa yang memiliki skor tidak wajar dalam tes dasar-dasar agama Buddha. Sehingga dapat dikatakan bahwa sebanyak 34 orang atau 79,07% calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 berhasil menjawab dengan betul soal dengan taraf sukar butir dalam kategori sukar sementara mereka salah dalam menjawab soal dengan taraf sukar butir dalam kategori mudah. Dengan hasil tersebut dapat dikatakan bahwa ketidakwajaran skor calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha berada dalam kategori tinggi.

Hal ini sejalan dengan pendapat yang dikemukakan oleh Hulin, *et. al.* (1983: 110) bahwa ketidakwajaran skor (*inappropriateness*) adalah bila seorang peserta tes dengan kemampuan tinggi salah dalam menjawab soal mudah dan peserta tes dengan kemampuan rendah banyak menjawab benar soal-soal yang sulit, seorang peserta tes tidak menjawab terlalu banyak soal yang mudah, atau seorang peserta tes menjawab secara acak keseluruhan tes. Demikian halnya Naga (2011: 6) mengemukakan bahwa manifestasi dari skor yang tidak wajar dapat terjadi ketika peserta ujian menjawab salah soal yang mudah dan menjawab benar soal yang sulit. Dengan demikian, ketidakwajaran skor pada tes dapat dipandang dari kombinasi mereka yakni dari kombinasi di antara letak yang sukar sehingga berhasil menjawabnya dengan benar, tetapi menjadi sangat ceroboh ketika berhadapan dengan butir yang mudah sehingga menjawabnya dengan salah. Sebagai akibatnya, ditemukan skor yang tidak wajar pada peserta tes itu.

Ketidakwajaran skor dapat diantisipasi jika semua kegiatan dan langkah pelaksanaan tes berlangsung dengan baik, selain itu kita juga akan mendapatkan gambaran kemampuan yang benar tentang peserta tes. Dari

hasil pendeteksian ketidakwajaran skor tersebut dapat diambil keputusan tentang apa yang harus dilakukan berikutnya terhadap peserta tes yang dideteksi skornya tidak wajar atau juga dapat menjadi bahan evaluasi pada proses pelaksanaan tes.

Sehingga dengan memperhatikan hasil penelitian tersebut, maka sepatutnya kriteria penetapan kelulusan calon mahasiswa baru tidak hanya didasarkan pada skor mentah (*raw score*) dari pengerjaan tes tertulis, melainkan perlu dilakukan pengecekan terhadap ketidakwajaran skor mereka. Hal ini dimaksudkan agar pelaksana tes dapat dengan tepat memprediksi peluang keberhasilan mereka dalam menempuh pendidikan di STABN Sriwijaya Tangerang Banten. Selain itu pelaksana tes juga akan dapat memperoleh gambaran kemampuan akademik calon mahasiswa baru.

PENUTUP

Berdasarkan hasil analisis yang telah dilakukan diperoleh informasi bahwa terdapat sebanyak 9 orang atau 20,93% calon mahasiswa yang memiliki skor wajar dalam tes dasar-dasar agama Buddha, sementara sebanyak 34 orang atau 79,07% calon mahasiswa yang memiliki skor tidak wajar dalam tes dasar-dasar agama Buddha. Dengan hasil tersebut, dapat dikatakan bahwa sebanyak 34 orang atau 79,07% calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 berhasil menjawab dengan betul soal dengan taraf sukar butir dalam kategori sukar sementara mereka salah dalam menjawab soal dengan taraf sukar butir dalam kategori mudah. Sehingga dapat dikatakan bahwa ketidakwajaran skor calon mahasiswa STABN Sriwijaya Tangerang Banten tahun 2018 dalam bidang tes dasar-dasar agama Buddha berada dalam kategori tinggi.

Melalui penelitian ini diharapkan kepada *stakeholder* dalam lingkup STABN Sriwijaya Tangerang Banten menjadikan ketidakwajaran skor sebagai salah satu komponen yang perlu diperhatikan dalam penerimaan mahasiswa baru. Sementara bagi pihak jurusan sekiranya dapat menyiapkan program khusus bagi mahasiswa yang teridentifikasi memiliki skor yang tidak wajar dalam bidang dasar-dasar agama Buddha. Hal ini dimaksudkan untuk membantu mereka dalam mengejar ketertinggalan mereka terkait dasar-dasar agama Buddha.

Melalui penelitian ini diharapkan kepada *stakeholder* dalam lingkup STABN Sriwijaya Tangerang Banten menjadikan ketidakwajaran skor sebagai salah satu komponen yang perlu diperhatikan dalam penerimaan mahasiswa baru. Sementara bagi pihak jurusan sekiranya dapat menyiapkan program khusus bagi mahasiswa yang teridentifikasi memiliki skor yang tidak wajar dalam bidang dasar-dasar agama Buddha. Hal ini dimaksudkan untuk membantu mereka dalam mengejar ketertinggalan mereka terkait dasar-dasar agama Buddha.

DAFTAR PUSTAKA

- Ansori, Ari Hasan. (2017). Pengaruh Metode Penyekoran dan Kecemasan Tes terhadap Indeks Ketidakwaian Skor Tes Matematika setelah Mengontrol Pengetahuan Awal. *Journal of Madrasah Ibtidaiyah Education*, Vol. 1, No. 1. Diakses dari <http://e-journal.adpgmiindonesia.com/index.php/jmie> pada 10 Agustus 2018.
- Dragow, F. Levine. (1987). Detecting Inappropriate Test Scores with Optimal and Practical Appropriateness Indices. *Applied Psychological Measurement*, Vol. 11, No. 1, March 1987. Diakses dari <http://journals.sagepub.com/doi/abs/10.1177/014662168701100105?journalCode=apmapada> 10 Agustus 2018.
- Ebel, Robert L. dan David A. Frisbie. (1991). *Essentials of Educational Measurement Fifth Edition*. New Delhi: Prentice Hall of India.
- Gronlund. (1977). *Constructing Achievement Tests Second Edition*. London: Prentice-Hall, Inc.
- Haladyna, Thomas M. (2004). *Developing and Validating Multiple-Choice Test Items Third Edition*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Hulin, Charles L., Fritz Dragow, dan Charles K. Parsons. (1983). *Item Response Theory: Applications to Psychological Measurement*. Illinois: Dow Jones-Irwin.
- Kerlinger, Fred N. (1973). *Foundations of Behavioral Research Second Edition*. New York: Holt, Rinehart and Winston, Inc.
- Kubiszyn, Tom dan Gary Borich. (2003). *Educational Testing and Measurement: Classroom Application and Practice*. New York: John Wiley & Sons, Inc.
- Mehrens, William A. dan Irvin J. Lehmann. (1991). *Measurement and Evaluation in Education and Psychology*. California: Wadsworth/Thomson Learning.
- Naga, Dali Santun. (2013). *Teori Skor pada Pengukuran Mental Edisi Kedua*. Jakarta: PT. Nagarani Citrayasa.
- Nitko, Anthony J. (1996). *Educational Assessment of Student*. New Jersey: Prentice Hall, Inc.
- Nurkencana, Wayan. (1983). *Evaluasi Pendidikan*. Surabaya: Usaha Nasional.
- Osterlind. (2002). *Constructing Test Items: Multiple-Choice, Constructed-Response, Performance, and Other Formats Second Edition*. New York: Kluwer Academic Publisher.
- Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan.
- Robb, Thomas N., dan Jay Ercanbrack. (1999). A Study of the Effect of Direct Test Preparation on the Toeic Scores of Japanese University Student. *Teaching English as a Second or Foreign Language Journal*, Vol. 3, No. 4, Januari 1999. <http://www.cc.kyoto-su.ac.jp/information/tesl-ej/ej12/a2.html>, pada 10 Agustus 2018.
- Rosana, Dadan dan Sukardiyono. (2015). Analisis Butir dan Identifikasi Ketidakwaian Skor Ujian Akhir Sekolah untuk Standarisasi Penilaian. *Jurnal Kependidikan*, Vol. 45, No. 2, November 2015. Diakses dari <https://journal.uny.ac.id/index.php/jk/article/view/7490> pada tanggal 10 Agustus 2018.

- Santoso, Budi. (2011). Ketidakwajaran Skor Berdasarkan Teori Responsi Butir. *Jurnal Evaluasi Pendidikan, Vol. 2, No. 2, Oktober 2011*. Diakses dari <http://journal.unj.ac.id/unj/index.php/jep/article/view/2257> pada 10 Agustus 2018.
- Setyaningsih, Nur Hayati. (2017). Komparasi Metode Pendeteksian Ketidakwajaran Skor Berdasarkan Gender dan Ruang Lingkup Materi pada Instrumen yang Mengukur Kemampuan Pemahaman Matematika. *Jurnal Evaluasi Pendidikan, Vol. 8, No. 2, Oktober 2017*. DOI: doi.org/10.21009/JEP.082.05. Diakses dari <http://journal.unj.ac.id/unj/index.php/jep/article/download/7111/5127/> pada 10 Agustus 2018.
- STABN Sriwijaya. (2018). Pedoman Penerimaan Mahasiswa Baru dan Pindahan Tahun Akademik 2018/2019 Sekolah Tinggi Agama Buddha Negeri Sriwijaya Tangerang Banten. STABNSW/P-PMB/12.
- Sudaryono. (2014). *Pengantar Evaluasi Pendidikan Berdasarkan Teori Tes Klasik dan Modern*. Jakarta: Lentera Ilmu Cendekia.
- Sudijono, Anas. (1996). *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali.
- Surapranata, Sumarna. (2005). *Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2014*. Bandung: PT. Remaja Rosdakarya.
- Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional*.
- Widyastuti, Suciati Rahayu. (2015). Perbandingan Indeks Ketidakwajaran Skor Metode SHL dan Metode Donlon-Fisher pada Tes Hasil Belajar Matematika. *Edusentris, Jurnal Ilmu Pendidikan dan Pengajaran, Vol. 2, No. 1, Maret 2015*. Diakses dari <http://ejournal.sps.upi.edu/index.php/edusentris/article/view/162> pada 10 Agustus 2018.