

## Analisis instrumen ujian formatif mata pelajaran pendidikan jasmani olahraga dan kesehatan tingkat SMP

**Iswanto Iswanto**

Dinas Pendidikan Kabupaten Buton Sulawesi Tenggara. Jl. Gajahmada Poros Baubau - Batauga, Masiri, Batauga, Kabupaten Buton, Sulawesi Tenggara 93752  
E-mail: [iswantois.99@gmail.com](mailto:iswantois.99@gmail.com)

---

### Abstrak

Penelitian ini bertujuan untuk mengetahui cakupan dan kualitas instrumen ujian formatif mata pelajaran Penjasorkes dilihat dari aspek kognitif dan psikomotor. Penelitian ini merupakan penelitian deskriptif. Ada 2 jenis instrumen digunakan, yakni soal essay dan soal ujian praktik. Kualitas soal essay dan soal unjuk kerja dilihat dari validitas isi dengan menggunakan formula Aiken, hasil estimasi reliabilitasnya menggunakan koefisien *Alpha* dan *Cronbach ICC*. Hasil penelitian menunjukkan dari 12 sekolah, 9 sekolah memiliki instrumen lengkap, yakni kompetensi kognitif dan psikomotor. Selanjutnya, dapat dijelaskan sebagai berikut: (1) ada 2 sekolah tidak memiliki soal essay, karena mengutamakan praktik. (2) ada 1 sekolah tidak memiliki soal praktik, karena tidak memiliki lapangan. Karakteristik soal ujian formatif mata pelajaran Penjasorkes, yaitu: (1) soal essay 9 sekolah memiliki indeks Aiken tergolong baik, dan soal 1 sekolah memiliki indeks Aiken tidak baik; (2) soal essay memiliki tingkat kesukaran sedang dengan reliabilitas tidak memenuhi syarat; (3) soal praktik 11 sekolah memiliki indeks Aiken tergolong baik dengan reliabilitas 9 sekolah tidak memenuhi syarat, dan reliabilitas 2 sekolah memenuhi syarat.

**Kata Kunci:** instrumen, ujian formatif, deskriptif kualitatif dan kuantitatif.

## An analysis of formative test instruments for subjects of physical education and health among junior high school

### Abstract

*This study aims to determine the scope and quality test instruments for Subjects of Physical Education and Health seen formative subjects of cognitive and psychomotor aspects. This research is descriptive. There are 2 types of instruments are used, the essay and exam practice. Quality of essay questions and problems of performance seen from the content validity by the formula Aiken, reliability estimation results of coefficient Alpha and ICC. The result showed from 12 schools, 9 school has a complete instrument, namely the cognitive and psychomotor competency. Furthermore, it can be described as follows: (1) there are two schools which do not have the essay, because the two schools prioritize learning practice. (2) there is one school which has no practical matter, because the school does not have a field facility, because it does not have a field. Characteristics exam for Subjects of Physical Education and Health formative subjects, namely: (1) essay 9 school has an index of Aiken classified is good, and about 1 school has an index of Aiken is not good; (2) The essay has a moderate level of difficulty with reliability ineligible; (3) about the practices of 11 schools have relatively good index of Aiken with reliability 9 schools were not eligible, and the reliability of two schools qualify.*

**Keywords:** instrument, formative test, descriptive qualitative and quantitative

## PENDAHULUAN

Kualitas pembelajaran (*instructional quality*) merupakan salah satu faktor yang sangat berperan dalam meningkatkan kualitas pendidikan karena arah dari program pendidikan

adalah pada terlaksananya program pembelajaran yang baik. Oleh karena itu, usaha meningkatkan kualitas pendidikan tidak akan tercapai tanpa adanya peningkatan kualitas pembelajaran.

Salah satu faktor penting untuk efektifitas pembelajaran di nilai dari faktor evaluasi, baik evaluasi proses maupun evaluasi hasil pembelajaran, sehingga dapat diketahui hasil pembelajaran. Evaluasi tidak hanya bertumpu pada penilaian hasil belajar, tetapi juga perlu penilaian terhadap proses pembelajaran itu sendiri sehingga dapat diketahui kualitas proses pembelajaran. Sejauh ini evaluasi banyak dilakukan pada penilaian hasil belajar untuk mengetahui berhasil tidaknya suatu program pembelajaran. Oleh karenanya, tidak dapat dilihat pengaruh proses pembelajaran terhadap hasil pembelajaran secara signifikan. Dengan demikian, kegiatan evaluasi sebagai bagian dari proses pembelajaran perlu lebih dioptimalkan. Hal ini perlu mendapat perhatian dikarenakan evaluasi pembelajaran dapat mendorong guru untuk lebih meningkatkan kualitas pembelajaran dan siswa agar lebih giat belajar secara terus menerus dan bahkan dapat mendorong sekolah untuk lebih meningkatkan fasilitas dan kualitas manajemen sekolah.

Selanjutnya, untuk lebih mengoptimalkan hasil evaluasi program pembelajaran maka peran guru perlu ditingkatkan kompetensinya. Guru merupakan salah satu komponen dalam program pembelajaran yang dapat menentukan baik tidaknya *output* pendidikan. Oleh karena itu, dalam dunia pendidikan, peran dan fungsi guru merupakan salah satu faktor yang sangat signifikan. Demikian halnya, kompetensi guru dalam menjalankan tugasnya dituntut secara profesional. Oleh sebab itu, dalam setiap upaya peningkatan kualitas pendidikan, tidak dapat terlepas dari berbagai hal yang berkaitan dengan eksistensi guru itu sendiri.

Sehubungan dengan hal tersebut, maka alat untuk mengukur evaluasi pembelajaran yang dilakukan oleh guru mata pelajaran adalah dengan cara mengukur kompetensi siswa. Pengukuran mempunyai kedudukan yang sangat penting dalam proses evaluasi. Baik buruknya hasil evaluasi tergantung pada hasil pengukuran. Hasil pengukuran yang kurang cermat, akan memberikan hasil evaluasi yang kurang cermat. Sebaliknya, jika teknik-teknik pengukurannya tepat, maka diharapkan dapat memberikan landasan yang kokoh untuk mengadakan evaluasi yang tepat pula. Agar hasil pengukuran dapat menggambarkan hasil pengukuran yang sebenarnya, maka alat ukur seperti perangkat soal yang digunakan harus memenuhi semua persyaratan sebuah alat ukur yang baik dan teruji dalam berbagai aspek.

Mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan merupakan salah satu pelajaran yang diajarkan di sekolah. Pada tingkat SD, SMP, dan SMA/ sederajat. Mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan pada dasarnya merupakan bagian integral dari sistem pendidikan secara keseluruhan yang bertujuan untuk mengembangkan aspek kesehatan, kebugaran jasmani, keterampilan berfikir kritis, stabilitas emosional, keterampilan sosial, penalaran dan tindakan moral melalui aktivitas jasmani dan olahraga. Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan khususnya di tingkat SMP, diharapkan mampu mengenalkan siswa dengan konsep-konsep penjas yang mengarahkan siswa agar memahami konsep tentang olahraga, kesehatan, dan prestasinya.

Pada mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan dalam evaluasi pembelajarannya dilakukan proses penilaian kognitif, penilaian afektif, dan penilaian psikomotor yang dikenal dengan nama metode pembelajaran langsung. Penilaian ini dimaksudkan untuk melihat seberapa jauh penguasaan materi secara keilmuan siswa tentang materi-materi yang telah diajarkan baik dari segi teori maupun praktik. Untuk itu diperlukan sebuah instrumen penilaian yang baik untuk melakukan pengukuran alat yang digunakan harus memenuhi syarat, bahwa instrumen tersebut harus valid, reliabel, relevan, efisien, dan ekonomis.

Sedangkan pengevaluasian dari data pengukuran bisa menggunakan standar norma atau standar kriteria. Evaluasi bisa dilakukan untuk evaluasi formatif atau sumatif sesuai dengan tujuannya. Permasalahan utama yang dialami guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan tingkat SMP dalam penyusunan instrumen alat ukur dalam domain psikomotor dan kognitif adalah kurangnya pengetahuan mengenai penyusunan tes formatif yang memenuhi standar validitas dan reliabilitas, instrumen yang baik terdapat miskonsepsi terhadap analisis teknik-teknik atau psikomotor dan kognitif akan diukur dalam

satu cabang olahraga tertentu. Selain itu, masalah yang lain adalah kurang terdistribusinya secara merata guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di seluruh sekolah SMP yang berstatus pegawai negeri sipil sehingga memanfaatkan tenaga honorer yang kurang memiliki pengalaman dan pengetahuan mengajar, serta ketercapaian untuk menyusun instrumen alat ukur hasil belajar sangat rendah. Masalah selanjutnya selama ini belum ada guru yang melakukan analisis instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan, serta belum ada yang melakukan penelitian terhadap penyusunan instrumen alat ukur hasil belajar siswa utamanya pada mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan. Sehingga dengan penelitian ini dapat membantu guru dalam menyusun dan menganalisis instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan dengan baik.

Sesuai dengan permasalahan yang ada, peneliti mengharapkan bahwa dengan mengetahui permasalahan mengenai kompetensi guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan tingkat SMP diharapkan penelitian ini dapat membantu sebagai alternatif solusi kompetensi guru. Oleh karena itu, peneliti tertarik untuk melakukan analisis instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan tingkat SMP.

Tujuan penelitian ini adalah (1) mendeskripsikan cakupan kelengkapan instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan untuk SMP. (2) mendeskripsikan karakteristik instrumen soal ujian formatif praktik siswa SMP pada mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan berdasarkan pada kedua ranah penilaian, yaitu: ranah kognitif dan psikomotor.

Pendidikan jasmani pada dasarnya merupakan bagian integral dari sistem pendidikan secara keseluruhan, bertujuan untuk mengembangkan aspek kesehatan, kebugaran jasmani, keterampilan berfikir kritis, stabilitas emosional, keterampilan sosial, penalaran dan tindakan moral melalui aktivitas jasmani dan olahraga.

### **Pengertian Pendidikan Jasmani**

Dengan pendidikan jasmani siswa akan memperoleh berbagai ungkapan yang erat kaitannya dengan kesan pribadi yang menyenangkan serta berbagai ungkapan yang kreatif, inovatif, terampil, memiliki kebugaran jasmani, kebiasaan hidup sehat dan memiliki pengetahuan serta pemahaman terhadap gerak manusia.

Dalam proses pembelajaran pendidikan jasmani guru diharapkan mengajarkan berbagai keterampilan gerak dasar, teknik dan strategi permainan dan olahraga, internalisasi nilai-nilai (sportivitas, jujur, kerjasama dan lain-lain) serta pembiasaan pola hidup sehat. Pelaksananya bukan melalui pembelajaran konvensional di dalam kelas yang bersifat kajian teoritis, namun melibatkan unsur fisik, mental, intelektual, emosi dan sosial.

Istilah pendidikan jasmani berawal dari Amerika Serikat berawal dari istilah *gymnastics*, *hygiene*, dan *physical culture* Siedentop (1972). Pengertian pendidikan jasmani menurut Rahayu (2013: 5), menyebutkan bahwa "tubuh adalah tempat bersemayamnya pikiran." Ada unsur kesatuan pemahaman antara tubuh dengan pikiran. Kesatuan unsur tubuh dan pikiran salah satu masalah besar, untuk selama bertahun-tahun lamanya seolah tidak akan pernah tuntas, adalah perdebatan antara intelektual dan jasmani. Kepercayaan banyak orang adalah bahwa tubuh terpisah dari pikiran, yang kemudian memunculkan pemahaman "dualisme" dan cenderung mengarah pada pikiran adalah sesuatu yang diutamakan, sementara tubuh adalah sesuatu yang *inferior*.

Keyakinan ini dapat dengan mudah dikenali, seperti yang sering di dengar sebuah semboyan *Orandum est ute sit men sana in corpora sano* atau seperti; *a sound mind in a sound body* (Krethmar, 2005: 51). Moto seperti ini, sering dijadikan rujukan dalam setiap pelaksanaan pendidikan jasmani. Pendidikan jasmani memanfaatkan aktivitas jasmani untuk mengembangkan aspek tubuh dan pikiran, dan bahkan aspek spiritual. Hal ini pun menjadi fokus orientasi utama dalam pengembangan aktivitas jasmani sebagai upaya pengembangan utuh-manusia. Pada kenyataannya di masyarakat sering ditemukan keyakinan bahwa tubuh dan pikiran berada pada sifat *dualism*. Sesungguhnya, pendidikan jasmani mencoba

membuktikan dan meyakinkan setiap orang bahwa tubuh dan pikiran berpadu menjadi satu kesatuan dalam konsep *holism*, meskipun pikiran berada di atas kedudukan tubuh.

Konsep penjas merupakan bagian penting dari proses pendidikan. Artinya, Penjas bukan hanya dekorasi atau ornamen yang ditempel pada program sekolah sebagai alat untuk membuat anak sibuk. Menurut Husdarta (2011: 17), bahwa "pendidikan jasmani adalah bagian penting dari pendidikan. Melalui penjas yang diarahkan dengan baik, anak-anak akan mengembangkan keterampilan yang berguna bagi pengisian waktu senggang, terlibat dalam aktivitas yang kondusif untuk mengembangkan hidup sehat, berkembang secara sosial, dan menyumbang pada kesehatan fisik dan mentalnya." Meskipun penjas menawarkan kepada anak untuk bergembira, tidaklah tepat untuk mengatakan penjas diselenggarakan semata-mata agar anak-anak bergembira dan bersenang-senang.

Penjas merupakan wahana pendidikan, yang memberikan kesempatan bagi anak untuk mempelajari hal-hal yang penting. Oleh karena itu, pelajaran penjas tidak kalah penting dibandingkan dengan pelajaran lain, seperti: Matematika, Bahasa, IPS, dan IPA, dan lain-lain. Dengan demikian tidak semua guru penjas menyadari hal tersebut, sehingga banyak anggapan bahwa penjas boleh dilaksanakan secara serampangan. Hal ini tercermin dari berbagai gambaran negatif tentang pembelajaran penjas, mulai dari kelemahan proses yang menetap. Misalnya membiarkan anak bermain sendiri hingga rendahnya mutu hasil pembelajaran, seperti kebugaran jasmani yang rendah. Di kalangan guru penjas sering ada anggapan bahwa pelajaran penjas dapat dilaksanakan seadanya, sehingga pelaksanaannya cukup dengan cara menyuruh anak pergi ke lapangan, menyediakan bola sepak untuk laki-laki dan bola voli untuk perempuan. Guru tinggal mengawasi di pinggir lapangan. Menurut Husdarta (2011: 18), bahwa "kelemahan ini berpangkal pada ketidakpahaman guru tentang arti dan tujuan penjas di sekolah, disamping guru mungkin kurang mencintai tugas itu dengan sepenuh hati."

Berdasarkan pada definisi tersebut di atas, maka secara umum penjas dapat didefinisikan penjas adalah proses pendidikan melalui aktivitas jasmani, permainan atau olahraga yang terpilih untuk tujuan pendidikan.

### **Pengertian Sistem dan Penilaian**

Sistem penilaian merupakan gabungan dari dua kata, yaitu sistem dan penilaian. Istilah sistem menunjukkan pada serangkaian subsistem-subsistem yang saling berinteraksi, berkolerasi, berintelerasi, dan saling berhubungan membentuk satu kesatuan yang utuh dan dinamis (Usman, 2008: 39-40). Hal yang senada disampaikan oleh Scott (1992) dalam bukunya Hoy & Miskel (2001: 10), "*system is the extent to which a set of actions is organized and implemented to achieve predetermined goals with maximum efficiency*". Sistem merupakan serangkaian kegiatan yang terorganisasi dan dilaksanakan untuk mencapai tujuan yang telah ditentukan secara efisien. Menurut Hartati Sukirman, dkk (1998: 6), sistem adalah unsur-unsur yang secara fungsional saling berkaitan sesuai dengan fungsinya masing-masing. Selain itu, Luxenburg & Ornstein (1998: 16), "*System are inputs, transformation process, outputs...*" Sistem terdiri dari input, proses, dan output.

Kegiatan untuk mengetahui keberhasilan belajar peserta didik dilakukan pengujian dan dilanjutkan penilaian. Data hasil pengukuran dan informasi hasil penilaian digunakan dalam melakukan *judgement* terhadap evaluasi program.

*Measurement, assessment and evaluation are hierarchical. The comparison of an observation with the criteria is a measurement, the interpretation and description of the evidence is an assessment and the judgement of the value or implications of the behavior is an evaluation* (Griffin & Nix, 1991: 3).

Pengukuran, penilaian dan evaluasi merupakan kegiatan yang saling menunjang dimana ketiganya merupakan hirarki. Pengukuran membandingkan hasil pengamatan dengan kriteria, penilaian menjelaskan dan menafsirkan hasil pengukuran, dan evaluasi adalah penetapan nilai dan implikasi perilaku.

Pengukuran menurut Allen & Yen (1979: 2), "*measurement is the assigning of numbers to individuals in a systematic way as a means of representing properties of the individuals*". Pengukuran adalah penetapan angka dengan cara yang sistematis untuk menyatakan

keadaan individu. Menurut Griffin & Nix (1991: 3), "*the process of measurement as assigning numbers to, or quantifying, things according to a set of rules*". Pengukuran adalah proses penetapan angka terhadap suatu gejala menurut aturan tertentu.

Berdasarkan pendapat-pendapat tersebut, dapat disimpulkan pengukuran merupakan suatu proses yang sistematis untuk mendapatkan informasi dalam bentuk angka-angka dari karakteristik individu dengan menggunakan alat tertentu berdasarkan prosedur dan aturan yang jelas.

Nana Sudjana (2014: 3), menyatakan penilaian sebagai proses memberikan atau menentukan nilai kepada objek tertentu berdasarkan suatu kriteria tertentu. Proses pemberian nilai tersebut berlangsung dalam bentuk interpretasi yang diakhiri dengan *judgement*. Interpretasi dan *judgement* merupakan tema penilaian yang mengimplikasikan adanya suatu perbandingan antara kriteria dan kenyataan dalam konteks situasi tertentu. Pendapat yang senada menurut Griffin & Nix (1991: 4):

*An assessment is a statement based on a series of facts to describe some characteristic of somebody or something. This definition of 'assessment' can relate to any part of the education process, it allows anything related to the teaching and learning process. It is not limited to student characteristic, as are some definitions. Characteristics of teaching methods, curricula, facilities, and administrations.*

Penilaian merupakan suatu pernyataan berdasarkan sejumlah fakta untuk menjelaskan karakteristik seseorang atau sesuatu. Definisi penilaian berhubungan dengan setiap bagian dari proses pendidikan, mencakup semua proses belajar mengajar. Kegiatan penilaian tidak terbatas pada karakteristik peserta didik saja, tetapi juga mencakup karakteristik metode mengajar, kurikulum, fasilitas, dan administrasi.

Senada dengan penjelasan penilaian tersebut, Carroll (1994: 6) menyatakan:

*Assessment is not only an integral part of teaching and learning, it should not be seen as a separate process even in the form of examinations and other summative assessments. It should be seen as part of a teaching assessment evaluation cycle, which at all points in the cycle feeds back information to both teacher and pupil for teaching purpose (formative assessment), to affect further action in the teaching-learning context and for curriculum and teaching evaluation.*

Penilaian tidak hanya merupakan bagian integral dari pengajaran dan pembelajaran, tidak harus dilihat sebagai proses yang terpisah bahkan dalam bentuk pemeriksaan dan penilaian sumatif lainnya. Ini harus dilihat sebagai bagian dari siklus pengajaran-penilaian-evaluasi, yang di semua titik dalam siklus feed kembali informasi kepada guru dan murid untuk tujuan pengajaran (penilaian formatif), untuk mempengaruhi tindakan lebih lanjut dalam konteks belajar-mengajar dan untuk kurikulum dan evaluasi pengajaran.

Penjelasan-penjelasan tersebut di atas, menunjukkan luasnya makna penilaian, maka penilaian dapat disimpulkan sebagai proses pengumpulan informasi untuk mengetahui pencapaian hasil belajar peserta didik yang dilakukan secara teratur dan berkesinambungan. Penilaian dalam penelitian ini difokuskan pada penilaian hasil belajar.

Menurut Kaufman & Thomas (1980: 4), "*evaluation is a process used to assess the quality of what going on. Evaluation may determine what is working, what is not working, what to change, and what to keep*". Evaluasi merupakan proses yang digunakan untuk menilai kualitas dari apa yang terjadi. Evaluasi dapat menentukan apa yang dikerjakan, apa yang tidak dikerjakan, apa yang harus dirubah, dan apa yang harus dipertahankan. Hal yang senada disampaikan oleh Tyler (Fernandes, 1984: 1), menyatakan "*evaluation as the process of determining to what extent the educational objectives are being realized*". Evaluasi adalah suatu proses menentukan sejauh mana tujuan pendidikan telah tercapai. Hal tersebut senada dengan pendapat Guskey (1999), bahwa "*evaluation is a judgement regarding the quality, value or worth a response, product, or performance, based on established criteria and curriculum standards*". Evaluasi adalah keputusan mengenai suatu kualitas, atau respon yang berharga, atau hasil suatu kinerja berdasarkan standar kurikulum dan kriteria yang telah ditentukan. Menurut Gay (Iman Sodikun, 2004: 3):

*(a) Evaluation is a systematic process of collecting and analyzing data in order to determine whether, and to what degree, objectives have been or are being achieved; (b)*

*evaluation is systematic process of collecting and analyzing data in order to make decision.*

Pernyataan pertama menunjukkan bahwa menentukan tingkat ketercapaian tujuan melalui proses yang sistematis mulai dari mengumpulkan data, menganalisis dan memberikan penilaian. Pernyataan kedua menunjukkan bahwa memberikan keputusan (*decision*) melalui proses sistematis dari mengumpulkan data, menganalisisnya, dan mengambil keputusan. (4) Kiefer (2010), memberikan pengertian:

*Evaluation is the process of examining a subject and rating it based on its important features. We determine how much or how little we value something, arriving at our judgement on the basis of criteria that we can define.*

Evaluasi adalah proses menguji suatu pokok bahasan dan menilainya berdasarkan unsur-unsur pentingnya. Kita dapat mengetahui seberapa besar atau seberapa kecil kita menilai sesuatu, membuat pertimbangan berdasarkan kriteria yang bisa kita tetapkan.

Berdasarkan pendapat-pendapat di atas, dapat disimpulkan bahwa evaluasi merupakan suatu kegiatan pengumpulan data dalam menilai suatu program sebagai umpan balik (*feedback*) untuk perbaikan pelaksanaan program serta menjadi bahan pertimbangan dalam pengambilan keputusan. Evaluasi merupakan kegiatan yang dilakukan secara sistematis untuk membantu pengambilan keputusan bagi *decision makers* dalam menentukan kebijakannya yang menyangkut program yang dievaluasi.

Menurut Hamalik (2003: 148), sistem penilaian merupakan subsistem dalam proses belajar mengajar. Sistem ini memberikan informasi, baik tentang proses pengajaran maupun tentang keberhasilan studi para siswa. Sistem penilaian terdiri dari komponen-komponen masukan, proses, dan produk. Hal senada pendapat Coll dkk (2007: 787):

*Assessment activities are to be intersed into teaching and learning activities, organized and sequenced around broad thematic areas, and teachers encouraged to provide follow-up, support and tutoring to students during the development of assessment activities.*

Kegiatan penilaian termasuk bagian dari kegiatan belajar mengajar, terorganisasi dan terangkai pada bidang utama yang luas, dan guru terdorong untuk memberikan tindak lanjut, dorongan dan bimbingan untuk peserta didik selama kegiatan pengembangan penilaian.

Sistem penilaian dalam penelitian ini dimaksudkan sebagai kesatuan rangkaian kegiatan penilaian yang terdiri atas komponen-komponen yang saling berhubungan guna mencapai tujuan pembelajaran (kompetensi). Sehingga diketahui hasil belajar peserta didik pada mata pelajaran tertentu. komponen-komponen tersebut terdiri atas input (perencanaan penilaian), proses (pelaksanaan penilaian), dan output (tindak lanjut hasil penilaian).

Berdasarkan pedoman penilaian terhadap kemampuan akademik siswa, meliputi: ranah kognitif, afektif, dan psikomotor. Bentuk tes untuk penilaian berupa tes tertulis, tes lisan, dan tes perbuatan. Cara penilaian yang digunakan adalah penilaian acuan angka yang memiliki bobot berbeda-beda.

**Tabel Nilai dan Predikat**

Nilai	Predikat
> 85 – 100	Baik Sekali
> 80 – 85	Baik
> 70 – 80	Cukup
> 65 – 70	Kurang
≤ 60	Sangat Kurang

Ketentuan dalam penilaian evaluasi hasil belajar, sistem penilaian hasil belajar, yaitu: (1) Evaluasi hasil belajar siswa dalam suatu mata pelajaran sekurang-kurangnya merupakan gabungan dari tiga penilaian, yaitu: tugas/ujian harian, ujian tengah semester, dan ujian akhir semester, (2) Bobot masing-masing penilaian yang digunakan ditetapkan berbeda. Evaluasi dengan bobot berbeda, yaitu: ujian harian atau tugas lain memiliki bobot 1, UTS memiliki bobot 2, dan UAS memiliki bobot 3. Penetapan bobot ini harus diberitahukan guru pengampuh mata pelajaran kepada siswa pada awal pembelajaran, dan (3) Perbaikan nilai hanya dapat dilakukan dengan menempuh remedial atau pengayaan.

## Penilaian Formatif

Cowie & Bell (2002: 6), mendefinisikan penilaian formatif sebagai proses yang digunakan guru dan siswa dalam mengenali dan merespon belajar siswa dalam rangka meningkatkan proses belajar dan pembelajaran. Penilaian formatif membantu guru dalam menggambarkan kemajuan belajar siswa dan menginformasikan keputusan tentang langkah selanjutnya dalam pembelajaran. Jadi informasi penilaian formatif dapat digunakan oleh guru dan siswa untuk memodifikasi cara belajar atau cara mengajarnya dengan harapan mendapatkan hasil yang lebih efektif. Popham (2008: 5), mendefinisikan penilaian formatif sebagai berikut: "*formatif assessment is a planned process in which assessment-elicited evidence of students' status is used by teachers to adjust their on going instructional procedures or by students to adjust their current learning tactics*".

Penilaian formatif merupakan proses terencana dimana bukti yang dihasilkan penilaian tentang status siswa digunakan oleh guru untuk menyesuaikan prosedur pembelajaran yang sedang berlangsung atau untuk menyesuaikan taktik belajar oleh siswa. Senada dengan itu, *State Collaborative on Assessment and Student Standards (SCASS) Formative Assessment for Students and Teachers (FAST)*, Popham (2008: 5), menyatakan bahwa penilaian formatif merupakan proses yang digunakan oleh guru dan siswa selama pembelajaran yang memberikan balikan untuk mengatur belajar dan pembelajaran yang berkelanjutan dalam rangka meningkatkan pencapaian luaran pembelajaran. Jadi, penilaian formatif merupakan sebuah proses yang terencana tentang segala sesuatu yang berkaitan dengan status siswa dan dimanfaatkan oleh guru untuk mengatur prosedur pembelajaran yang berkelanjutan dan dimanfaatkan oleh siswa untuk mengatur strategi belajar mereka.

Penilaian formatif merupakan sebuah proses yang terencana dan melibatkan berbagai kegiatan yang berbeda. Salah satu kegiatan yang dilakukan dalam proses penilaian formatif adalah mengumpulkan berbagai kejadian tentang status siswa, baik secara formal maupun informal, untuk menentukan derajat penguasaan siswa terhadap indikator kurikulum dan keterampilan-keterampilan. Berdasarkan kejadian-kejadian dan informasi inilah seorang guru akan melakukan penyesuaian terhadap pembelajaran yang dilakukan dan siswa akan melakukan pengaturan cara-cara mempelajari materi pelajaran.

Beberapa pengertian tentang penilaian formatif yang telah dikemukakan memiliki beberapa kesamaan. Kesamaan ini antara lain: (1) Penilaian formatif merupakan proses yang dilakukan dalam pembelajaran, (2) Hasil penilaian formatif tidak saja digunakan oleh guru tetapi juga digunakan oleh siswa, (3) Penilaian formatif memberikan umpan balik terhadap proses belajar siswa dan proses pembelajaran yang dilakukan oleh guru, (4) Umpan balik yang diberikan oleh penilaian formatif akan berguna bagi siswa dan guru untuk melakukan pengaturan-pengaturan sehingga belajar dan pembelajaran dapat mencapai tujuan kurikulum.

### *Validitas dan Reliabilitas Instrumen*

Validasi instrumen merupakan tahap yang mendasar dalam pengembangan dan evaluasi instrumen. Validitas merupakan ketepatan tes terhadap komponennya dan kecocokan skor dengan penafsirannya (Wright, 2008: 146-147). Proses validasi meliputi pengumpulan bukti-bukti untuk menunjukkan dasar saintifik penafsiran skor seperti yang tercantum pada tujuan penggunaan instrumen penilaian. Dengan kata lain skor hasil penilaian dapat ditafsirkan sesuai dengan tujuan penggunaan instrumen. Namun, langkah agar mendapatkan penafsiran yang tepat adalah dengan validasi instrumen terlebih dahulu.

Ada beberapa cara untuk membuktikan validitas suatu instrumen, kebanyakan dikelompokkan menjadi 3 kategori, yaitu: isi, berdasarkan kriteria, dan konstruk (Ebel & Firsbie, 1986: 90). Suatu instrumen disebut valid apabila isi instrumen penilaian tersebut komprehensif, relevan, dan tidak keluar dari batasan tujuan ukur. Penetapan isi instrumen dipandang valid dapat dibuktikan berdasarkan analisis rasional terhadap isi instrumen, yang penilaiannya didasarkan atas pertimbangan subjektif individual. Individu yang tepat untuk menentukan validitas isi adalah mereka yang dipandang ahli dalam komponen yang diukur, sehingga hasil yang diperoleh dapat dipertanggungjawabkan.

Prosedur pendekatan validitas berdasar kriteria menghendaki tersedianya kriteria eksternal yang dapat dijadikan pengujian skor penilaian (Azwar. S., 2010: 50). Adapun kriteria

variabel yang dijadikan pengujian skor penilaian tergantung pada ketentuan pengguna instrumen. Ada dua versi model kriteria, yaitu: konkuren dan prediktif. Maksud dari kedua versi tersebut menurut Brennan (2006: 18) adalah :

*Concurrent validity studies employed criterion scores obtained at about the same time as the test score and could be used to validate a proxy measure that would be cheaper, easier, or safer than the criterion. Predictive validity studies employed a criterion of future performance (e.g., on the job, in college), which was not available at the time of testing.*

Antara validitas konkuren dan prediktif dibedakan pada fungsi penggunaan dan waktu perolehan antara skor tes dengan skor kriteria. Validitas konkuren digunakan untuk mendiagnosa hubungan variabel satu dengan variabel lainnya dan waktu perolehan skor tes dengan skor kriteria dapat diperoleh bersama-sama. Adapun validitas prediktif digunakan untuk memprediksi performansi di waktu yang akan datang, sehingga waktu perolehan skor tes dengan skor kriteria tidak bisa dilakukan bersama-sama.

Prosedur validasi juga dapat dilakukan pendekatan konstrak. Dalam *Standards* yang dikeluarkan *American Educational Research Association* (1999: 12), menyebutkan konstrak merupakan jawaban tes yang diperoleh sebagai hasil respon terhadap item. Pengujian validitas konstruk ini dilakukan terus menerus sejalan dengan perkembangan sifat (*trait*) yang diukur. Dalam pelaksanaannya validitas konstruk tidak hanya teridentifikasi berdasarkan hasil penelitian namun juga berdasarkan orientasi peneliti (Moss, 2007: 472).

Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel. Walaupun reliabilitas mempunyai berbagai nama lain seperti: kepercayaan, keterandalan, keajegan, kestabilan, dan konsistensi, namun ide pokok yang terkandung dalam konsep reliabilitas adalah seberapa jauh hasil suatu pengukuran dapat dipercaya. Berkenaan dengan konsep reliabilitas ini, Miller (2002: 59), mengatakan bahwa reliabilitas adalah tingkat ketetapan atau konsistensi pengukuran oleh sebuah tes dalam mengukur kualitas yang sama dalam setiap pelaksanaan tes tersebut. Reliabilitas diartikan sebagai ketetapan pengukuran berarti, bahwa semua orang yang menggunakan prosedur ini mendapatkan hasil yang sama secara ajeg. Safrit dan Wood (1989: 45), menjelaskan bahwa "*if a test or measuring instrument is reliable, it consistently will provide the same measurement of a person. If a measurement instrument is totally reliable, it can be administered several times in one day or on several different days, and each times a person would obtain the same score.*" Wahjoedi (2000: 32), mengatakan bahwa suatu alat ukur atau tes dikatakan reliabel apabila hasil-hasil penggunaan tes tersebut menunjukkan ketetapan atau diperoleh hasil pengukuran yang ajeg atau tetap terhadap suatu yang seharusnya diukur. Menurut Gronlund & Linn (1990: 77), *reliability refers to the consistency of the test scores that is, how consistency they are from one measurement to another.* Pernyataan ini berarti bahwa reliabilitas mengacu pada konsistensi skor-skor tes, apabila tes tersebut mengukur hal yang sama, atau seberapa konsisten tes-tes tersebut dari satu pengukuran ke pengukuran lainnya. Linn dan Gronlund (1990: 78-79), menjelaskan, "(1) *reliability refers to the results obtained with an assessment instrument and not to instrument itself. (2) an estimate of reliability always refer to particular type of consistency, (3) reliability is a necessary but not sufficient condition for validity, and (4) reliability is primarily statistical.*"

Pendapat di atas, dapat ditarik suatu makna bahwa reliabilitas sebuah tes menunjuk pada tingkat keajegan atau konsistensi skor-skor yang relatif bebas dari kesalahan-kesalahan. Kecenderungan ini mengarah pada ketetapan yang ditunjukkan dengan memberikan ulangan prestasi dari sebuah perilaku pada setiap siswa. Gejala atau unsur-unsur dalam gejala yang diungkapkan dalam pengukuran pertama, ternyata tidak berubah atau sama pada pengukuran kedua dan seterusnya apabila pengukuran dilakukan dengan menggunakan instrumen yang sama.

Budiwanto, S., (2003: 141), mengatakan bahwa, ada tiga cara menaksir reliabilitas tes keterampilan olahraga, yaitu cara tes dan tes ulang (*test-retest*), cara belah dua (*split half*), dan menggunakan tes setara (*equivalent*). Cara tes dan tes ulang (*test-retest*) maksudnya adalah tes pertama dilakukan kemudian selang beberapa waktu disusul dilakukan tes ulang dengan menggunakan tes yang sama. Untuk memperoleh koefisien reliabilitas tes eksperimen, hasil tes pertama dan hasil tes ulang dikorelasikan menggunakan teknik statistik korelasi *product moment* dari Pearson. Koefisien korelasi antara hasil tes pertama dan hasil

tes kedua merupakan koefisien reliabilitas tes eksperimen. Mengetes orang-orang yang sama menggunakan tes yang sama akan menghasilkan indeks stabilitas.

Memperoleh reliabilitas tes keterampilan olahraga dengan cara belah dua (*split half*) hanya digunakan jika jumlah percobaan tes keterampilan tersebut terdiri dari beberapa kali percobaan. Contoh, tes servis bawah dalam permainan bola voli dilakukan sepuluh kali servis. Akan ada sepuluh skor dari usaha servis pertama sampai dengan kesepuluh. Skor-skor tersebut dapat dibelah atau dibagi menjadi dua kelompok.

Cara lain memperoleh reliabilitas tes adalah menggunakan tes yang setara (*equivalency* atau *alternate form*). Cara ini sering disebut juga menggunakan cara tes paralel (*parallel form*). Dalam hal ini harus dibuat dua perangkat tes yang pada dasarnya sama atau mempunyai tingkat kesetaraan (koefisien ekuivalen). Kedua tes yang dipakai memiliki kualitas dan spesifikasi yang sama. Reliabilitas tes eksperimen diperoleh dengan cara mengkorelasikan antara kedua tes tersebut.

## METODE

Penelitian ini merupakan penelitian deskriptif, yaitu mendeskripsikan cakupan dan kualitas instrumen yang digunakan dalam melakukan penilaian formatif. Penelitian ini bertujuan untuk mengetahui cakupan ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan kelas VII SMP Tahun Pelajaran 2014/2015 dilihat dari aspek kognitif dan psikomotor. Peneliti dibantu oleh ahli menganalisis instrumen ujian formatif secara kualitatif untuk mendeskripsikan bentuk instrumen yang baik berdasarkan aspek materi, konstruksi, dan bahasa.

Penelitian dilaksanakan bulan Maret – April 2015 pada tingkat SMP Tahun Pelajaran 2014-2015. Sasaran dari penelitian ini adalah instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan tingkat SMP di kelas VII dengan materi yang telah dibatasi oleh peneliti mengenai permainan bola besar.

Subjek penelitian ini adalah siswa kelas VII SMP Tahun Pelajaran 2014/2015. Objek dalam penelitian ini adalah seluruh instrumen ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan yang dibuat oleh guru di kelas VII, dan lembar jawaban siswa yang diperoleh dari guru beserta hasil pengamatan guru saat melakukan ujian formatif praktik mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan dengan materi permainan bola voli di semester genap Tahun Pelajaran 2014/2015. Objek penelitian tersebut yang digunakan dalam penelitian ini sebanyak 12 sekolah diperoleh dari 44 SMP yang diambil menggunakan teknik *purposive sampling* yaitu hanya berfokus pada sekolah yang memiliki guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan berstatus PNS serta memiliki perangkat tes berbeda pada setiap sekolah. Objek penelitian tersebut akan dianalisis untuk melihat cakupan ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan SMP pada aspek kognitif dan psikomotor.

Pengumpulan data dalam penelitian ini adalah menggunakan instrumen *checklist* (lembar telaah butir soal) dan dokumentasi. *Checklist* digunakan untuk melihat kriteria butir soal dilihat dari aspek materi, konstruk, dan bahasa. Sedangkan teknik dokumentasi yang digunakan adalah untuk mengumpulkan data-data yang berkaitan dengan penelitian yang akan dilakukan, seperti kisi-kisi soal, perangkat soal, perangkat unjuk kerja, kunci jawaban, pedoman penskoran, respon peserta didik yang diambil di sekolah-sekolah yang menjadi objek penelitian. Pelaksanaannya dengan mendatangi sekolah yang dijadikan objek penelitian dan menemui kepala sekolahnya untuk meminta izin, kemudian peneliti menemui guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan untuk meminta data berkaitan dengan instrumen yang digunakan pada ujian formatif serta hasil penilaian siswa.

Berdasarkan dari dua ranah yang akan dianalisis dalam penelitian ini, yaitu: ranah kognitif dan psikomotor, maka peneliti membagi dua model analisis didasarkan pada jenis obyek data yang diteliti. Pada penilaian instrumen kognitif dan psikomotor. Adapun dua jenis teknik analisis data tersebut, yaitu.

### Analisis Kualitatif

Penelaahan dari aspek materi, konstruksi, dan bahasa dilakukan secara terpisah yang ditelaah oleh masing-masing ahli, yaitu: ahli dalam bidang studi yang diujikan, ahli dalam bidang pengukuran, dan ahli dalam bidang gagasan. Aspek materi berkaitan dengan substansi keilmuan yang ditanyakan serta tingkat berpikir yang terlibat. Aspek konstruksi berkaitan dengan teknik penulisan soal, baik bentuk objektif, maupun yang non-objektif. Aspek bahasa berkaitan dengan kekomunikatifan atau kejelasan hal yang ditanyakan.

Hasil penelaahan instrumen peneliti membuat rangkuman dan kesimpulan terhadap butir-butir yang sudah memenuhi kriteria penilaian pada pedoman telaah. Kriteria keputusan yang diambil adalah memperoleh dua dukungan dari ketiga penelaah dengan katagori diterima, direvisi, dan ditolak.

Pedoman penelaahan butir soal yang digunakan untuk menelaah butir soal mengacu pada pedoman yang telah ditetapkan oleh Pusat Penilaian Pendidikan. Pedoman penelaahan soal bentuk uraian adalah sebagai berikut.

#### Aspek Materi

- a). Soal sesuai dengan indikator;
- b). Batasan pertanyaan dan jawaban yang diharapkan sesuai;
- c). Materi yang diukur sesuai dengan kompetensi; dan
- d). Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah, dan tingkat kelas.

#### Aspek konstruksi

- a). Ada petunjuk yang jelas mengenai cara mengerjakan soal;
- b). Rumusan kalimat soal/pertanyaan menggunakan kata tanya atau perintah yang menuntut jawaban terurai;
- c). Gambar/grafik/tabel/diagram dan sebagainya, jelas dan berfungsi; dan
- d). Ada pedoman penskoran.

#### Aspek bahasa

- a). Rumusan kalimat soal/pertanyaan komunikatif;
- b). Butir soal menggunakan Bahasa Indonesia yang baku;
- c). Tidak mengandung kata-kata/kalimat yang mengandung penafsiran ganda atau salah pengertian;
- d). Tidak mengandung kata yang menyinggung perasaan; dan
- e). Tidak menggunakan bahasa yang berlaku setempat/tabu.

Butir soal diterima, apabila butir soal tersebut memenuhi semua kriteria yang ada pada aspek materi, konstruk, dan bahasa. Butir soal direvisi, apabila butir soal memenuhi paling tidak satu kriteria dari aspek materi, tiga kriteria dari aspek konstruk, dan satu kriteria dari aspek bahasa. Untuk soal ditolak, apabila butir soal tidak memenuhi semua kriteria penilaian yang telah ditetapkan.

### Analisis Kuantitatif

Untuk analisis secara kuantitatif di analisis menggunakan formula teori tes klasik untuk menganalisis respon peserta didik terhadap perangkat soal. Analisis ini dilakukan untuk melihat karakteristik instrumen yang dirancang oleh masing-masing guru pada SMP se-Kabupaten Buton Tahun Pelajaran 2014/2015. Pada soal uraian hasil analisis memberikan informasi tentang butir soal yang layak dan tidak layak untuk diujikan, yaitu tingkat kesukaran yang di analisis dengan menggunakan program *EXCEL*. Butir yang baik dan handal akan memberikan informasi yang akurat tentang keberhasilan pembelajaran peserta didik.

Tingkat kesukaran, yaitu proporsi peserta tes yang menjawab benar. Rentang tingkat kesukaran berada antara interval 0 sampai dengan 1. Jika proporsi menjawab benar ( $p$ ) < 0,30, maka disimpulkan bahwa tingkat kesukarannya tinggi, apabila proporsi menjawab benar ( $p$ ) diantara 0,30 sampai 0,70, maka soal memiliki tingkat kesukaran sedang, dan jika proporsi menjawab benar ( $p$ ) > 0,70, maka disimpulkan bahwa tingkat kesukarannya rendah.

## HASIL DAN PEMBAHASAN

Proses pelaksanaan penelitian ini sesuai dengan petunjuk pada metode penelitian yang mengambil subjek penelitian siswa peserta ujian formatif dengan materi permainan bola voli semester genap pada mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di kelas VII SMP Tahun Pelajaran 2014/2015 dengan sejumlah 12 sekolah. Subjek penelitian ini adalah siswa yang mengikuti ujian formatif dengan materi permainan bola voli semester genap. Objek penelitian ini adalah kisi-kisi soal, soal, kunci jawaban, lembar penskoran, kisi-kisi ujian praktik, praktik, lembar penskoran praktik, kisi-kisi sikap, penilaian sikap, dan penskoran sikap yang dikumpulkan hingga menjadi sebuah dokumentasi. Selain itu, respon siswa berupa lembar jawaban dan nilai hasil ujian psikomotor. Dokumentasi yang diperoleh melalui 12 sekolah melalui guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan yang dijadikan sampel penelitian ini memiliki keragaman. Tahapan-tahapan yang harus diperhatikan pada instrumen kognitif dan psikomotor berupa karakteristik butir soal, unjuk kerja, dan kuesioner yang dianalisis secara kualitatif dan ditelaah oleh ahli, selanjutnya dianalisis secara kuantitatif dengan menggunakan program komputer.

Untuk analisis kognitif hasil lembar kerja siswa selanjutnya dilakukan telaah Aiken, perhitungan penskoran yang mana setiap tahapan pengerjaan jawabannya diperhitungkan sesuai lembar penskorannya. Ketentuan penskoran yang digunakan dalam penilaian lembar jawaban siswa dapat dicontohkan yaitu jika skor 0 apabila lembar jawaban siswa tidak sesuai dengan kunci jawaban, jika skor 1 apabila lembar jawaban siswa benar dari 1 atau 2 bagian dari kunci jawaban, jika skor 2 apabila lembar jawaban siswa benar 3 bagian dari kunci jawaban, jika skor 3 apabila seluruh jawaban siswa sesuai dengan kunci jawaban.

Berdasarkan dari perolehan deskriptif kualitatif dan kuantitatif, maka dapat dikemukakan hasil penelitian sebagai berikut.

1. Hasil penelitian menunjukkan bahwa dari 12 sekolah, hanya 9 sekolah yang memiliki instrumen lengkap, yakni memiliki soal ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan kelas VII pada kompetensi kognitif dan psikomotor. Selanjutnya, secara rinci dapat dijelaskan sebagai berikut: (1) ada 2 sekolah tidak memiliki soal essay, dikarenakan 2 sekolah tersebut lebih mengutamakan pembelajaran bersifat praktik. (2) ada 1 sekolah yang tidak memiliki soal praktik, dikarenakan sekolah tersebut tidak memiliki fasilitas lapangan.
2. Karakteristik instrumen soal ujian formatif mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan SMP, yaitu: (1) untuk soal uraian yang mengungkap aspek kognitif dari 9 sekolah memiliki indeks Aiken tergolong baik (Aiken validitas 0,54, sampai dengan 0,69), dan hanya soal dari 1 sekolah yang tidak baik, (2) soal uraian tersebut rata-rata memiliki tingkat kesukaran sedang (TK 0,42 sampai dengan 0,58) dengan reliabilitas memenuhi syarat (*Alpha* 0,41 sampai dengan 0,59). (3) untuk soal praktik yang mengungkap aspek psikomotor dari 11 sekolah memiliki indeks Aiken tergolong baik (Aiken validitas 0,47 sampai dengan 0,59) dengan reliabilitas soal praktik 9 sekolah yang tidak memenuhi syarat (*ICC* 0,37 sampai dengan 0,49) dan reliabilitas soal praktik 2 sekolah yang memenuhi syarat (*ICC* 0,84 dan 0,94).

### Temuan Penelitian

Berdasarkan hasil penelitian dan pembahasan, ada beberapa hal yang ditemukan sebagai berikut : (1). Kelengkapan instrumen penelitian yang dikumpulkan dari 12 sekolah pada 3 ranah, yaitu ranah kognitif dan psikomotor terdapat beberapa sekolah kurang lengkap. Untuk sekolah tidak memiliki soal kognitif terdapat 2 sekolah, dikarenakan guru mata pelajarannya lebih memprioritaskan kegiatan psikomotor. Untuk sekolah tidak memiliki soal psikomotor hanya 1 sekolah, ini dikarenakan tidak memiliki lapangan. (2). Karakteristik butir soal yang dianalisis secara kualitatif menunjukkan bahwa ada beberapa indikator dari setiap aspek yang tidak terpenuhi, sehingga dari segi kualitatif perangkat soal yang dibuat oleh guru dari masing-masing sekolah kurang baik.

## SIMPULAN

Berdasarkan pada hasil dan pembahasan penelitian ini, maka ada beberapa hal yang dapat disimpulkan, yakni: (1) Penelitian ini hanya terbatas pada satu unsur materi pokok bahasan, yaitu materi permainan bola voli. (2) Penelitian ini dalam memperoleh dokumen dari sekolah, utamanya pada ranah kognitif dan psikomotor yang selanjutnya dilakukan analisis dengan menggunakan analysis faktor (CFA) secara umum terdapat tingkat koefisien reliabilitas tidak memenuhi syarat, hal ini dikarenakan jumlah item butir sangat minim digunakan kepada siswa cukup besar sehingga menghasilkan tingkat reliabilitas yang rendah. (3) Penelitian ini hanya melihat menganalisis instrumen ujian formatif berbentuk uraian, sehingga bentuk soal lain berupa pilihan ganda, maupun soal campuran lainnya belum dapat terlaksana. (4) Penelitian ini pula, hanya melihat pada instrumen yang dibuat oleh guru mata pelajaran pada masing-masing sekolah, sehingga instrumen yang dinilai dalam penelitian ini tidak digeneralisasikan.

### Saran

1. Instrumen yang digunakan oleh guru mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan tingkat SMP sebaiknya perlu dikondisikan dengan keadaan masing-masing sekolahnya.
2. Forum Musyawarah Guru Mata Pelajaran (MGMP) perlu diaktifkan kembali, guna instrumen yang telah disusun oleh guru dapat terpenuhi baik validitas dan reliabilitas instrumen yang akan digunakan.
3. Sebelum soal digunakan untuk evaluasi belajar peserta didik perlu dilakukan telaah oleh ahli (*expert judgment*) dengan meminta ahli yang memiliki pengalaman dalam pembelajaran Pendidikan Jasmani Olahraga dan Kesehatan serta penilaian. Analisis kualitatif bertujuan untuk melihat karakteristik soal ditinjau dari aspek materi, konstruk, dan bahasa.
4. Setelah dilakukan analisis secara kualitatif, selanjutnya perlu dianalisis secara kuantitatif dengan teori tes klasik dapat menggunakan program EXCEL. Analisis kuantitatif ini bertujuan untuk melihat tingkat kesukaran untuk soal essay.
5. Penyusunan bank soal agar guru lebih mudah dalam melakukan evaluasi hasil belajar peserta didik.

## DAFTAR PUSTAKA

- Allen, M.J. & Yen, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company.
- Azwar.S., (2015). *Dasar-dasar Psikometri. Ed. II*. Yogyakarta: Pustaka Pelajar.
- Azwar.S., (2014). *Validitas dan Reliabilitas. Ed. IV*. Yogyakarta: Pustaka Pelajar.
- Budiwanto.S., (2003). *Prosedur menyusun tes keterampilan olahraga*. Jurnal IPTEK Olahraga, Volume 5 Nomor 2 Mei 2003, 133-150.
- Bell, B., & Cowie B. (2002). *Formative assessment and science education*. New York: Kluwer Academic.
- Bompa, T.O. (1994). *Theory and methodology of training* (Terjemahan). Program Pascasarjana Universitas Pajajaran Bandung.
- Brennan. (2006). *Educational Measurement*. Westpost: Greenwood Publishing Group.
- Carroll.B., (1994). *Assessment in Physical Education: A Teacher's Guide to the Issues*. Washington, D.C: The Falmer Press.
- Coll, C., Rochera, M.J., Mayordomo, R.M., et al. (2007). Continuous assessment and support for learning: an experience in educational innovation with ICT support in higher education [Versi elektronik]. *Journal of Research in Educational Psychology*, N. 13 Vol 5(3), 783-804.
- Ebel, R.K & Frisbie, D.A. (1986). *Essensial of educational measurement (4<sup>th</sup> ed.)*. Englewood Cliffs New Jersey: Prentice-Hall. Inc.

- Guskey, T.R. (1999). *Using assessment to improve student learning. National conference on standards and assessment*. Alberta Assessment Consortium. Diambil dari <http://www.aac.ab.ca/aboutqa.html>.
- Griffin, P. & Nix, P. (1991). *Educational assessment and reporting: A new approach*. Harcourt Brace Jovanovich (Australia).
- Gronlund, N.E & Linn, R.L. (1990). *Measurement and evaluation in teaching 6<sup>th</sup> edition*. New York: MacMillan Publishing Company.
- Hamalik.O., (2003). *Manajemen belajar di perguruan tinggi: pendekatan kredit semester (sks)*. Bandung: Sinar Baru Algensindo.
- Husdarta, (2011). *Manajemen Pendidikan Jasmani*. Alfabeta. Bandung.
- Hoy, W.K & Miskel, C.G. (2001). *Educational administration: theory, research, and practice (6<sup>th</sup> ed.)*. United States: School Management and Organization.
- Kiefer, K. (2010). *A definition of evaluation*. Colorado State University. Diambil dari <http://writing.colostate.edu/guides/processes/evaluate/pop2a.cfm>.
- Linn, R.L. & Miller, M.D. (2005). *Measurement and assessment in teaching (9<sup>th</sup> ed.)*. New Delhi: Pearson Education, Inc.
- Luxenburg, F.C. & Ornstein, A.C. (1998). *Educational administration: concepts and practices (3<sup>rd</sup> ed)*. United States: School Management and Organization.
- Mardapi.D., (2012). *Pengukuran penilaian & evaluasi pendidikan*. Yogyakarta: Nuha Media.
- Mardapi.D., (2008). *Teknik Penyusunan Tes & Non Tes*. Yogyakarta: Mitra Cendekia Press.
- Mardapi.D., (2004). *Pengembangan sistem penilaian berbasis kompetensi*. Disampaikan pada Seminar HEPI "Rekayasa Sistem Penilaian untuk Meningkatkan Kualitas Pendidikan" Yogyakarta 26-27 Maret 2004.
- Miller, D.K. (2002). *Measurement by the physical educator 4<sup>th</sup> edition*. San Fransisco: McGraww Hill.
- Morrow, J.R; Jackson, A.W; Disch, J.G; dan Mood, D.P., (2005). *Measurement and Evaluation in Human Performance*. 3<sup>rd</sup> Edition USA: Human Kinetics.
- Popham, W.J. (2008). *Transformative Assessment*. Alexandria: Association of supervision and curriculum development (ASCD).
- Rahayu.E.T., (2013). *Strategi Pembelajaran Pendidikan Jasmani: Implementasi pada Pembelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan*. Bandung: Alfabeta.
- Safrit, M.J & Wood, T.M. (1989). *Measurement concept in physical education and exercise science*. Illionois: Human Kinetics Books.
- Sudjana.N., (2014). *Penilaian hasil proses belajar mengajar*. Bandung: PT Remaja Rosdakarya.
- Sukirman, H. dkk (1998). *Administrasi dan supervise pendidikan*. Yogyakarta: IKIP Yogyakarta.
- Sodikun, I. (2004). *Pengembangan sistem penilaian hasil belajar keterampilan gerak*. Disampaikan pada seminar HEPI "Rekayasa Sistem Penilaian dalam rangka Meningkatkan Kualitas Pendidikan" Yogyakarta 26-27 Maret 2004.
- Thomas, J.R., & Nelson, J. Krecthmar. (1990). *Methods research in physical activity (2 nd ed)*. New Jersey: Human Kinetics Publisher.
- Usman, H. (2008). *Manajemen: teori, praktik, dan riset pendidikan (edisi kedua)*. Jakarta: Bumi Aksara.
- Wahjoedi. (2000). *Landasan Evaluasi Pendidikan Jasmani*. Jakarta: PT Raja Grafindo Persada.
- Wright. R.J., (2008). *Educational Assessment: Test and Measurements in the Age of Accountability*. California: Widener University, Sage Publication.