

PENGEMBANGAN INSTRUMEN ASESMEN DIRI DAN TEMAN SEJAWAT KOMPETENSI BIDANG STUDI PADA MAHASISWA

¹⁾Enny Wijayanti, ²⁾Mundilarto

¹⁾ Universitas Palangka Raya, ²⁾ Universitas Negeri Yogyakarta
¹⁾ennywijayanti28@yahoo.com, ²⁾mundilarto@uny.ac.id

Abstrak

Tujuan penelitian ini adalah untuk: (1) mengembangkan konstruk instrumen asesmen diri (ADM) dan asesmen teman sejawat (ATS) sebagai pelengkap asesmen formatif, (2) mengetahui karakteristik ADM dan ATS, (3) mengetahui efektivitas ADM & ATS, dan (4) mengetahui kebermanfaatan ADM & ATS yang digunakan sebagai pelengkap asesmen formatif. Prosedur pengembangan meliputi dua tahap, yaitu pengembangan dan validasi, mengacu pada model pengembangan Cennamo dan Kalk dengan lima fase pengembangan, yaitu *define, design, demonstrate, develop* dan *deliver*. *Content validity* dibuktikan melalui *expert judgment* dianalisis dengan *coefficient Aiken* dan *focus group discussion* dan teknik Delphi. Validitas konstruk dibuktikan dengan program SmartPLS. Subjek penelitian adalah mahasiswa semester enam Pendidikan Fisika Universitas Palangka Raya. Hasil penelitian adalah sebagai berikut: (1) instrumen ADM dan ATS mengukur konstruk yang sama pada kompetensi fisika, hasil uji kecocokan model nilai GFI sebesar 0,994, dan SRMR sebesar 0,062 menunjukkan model fit yang baik, (2) karakteristik instrumen mencakup validitas, reliabilitas, tingkat kesukaran dan daya pembeda termasuk kategori baik, (3) ADM dan ATS terbukti cukup efektif dalam meningkatkan hasil belajar, (4) ADM dan ATS sangat bermanfaat sebagai umpan balik terhadap proses, kemajuan dan perbaikan hasil belajar.

Kata kunci: *asesmen diri dan asesmen teman sejawat, kompetensi fisika*

DEVELOPING SELF-ASSESSMENT AND PEER-ASSESSMENT FOR STUDENTS' SUBJECT-MATTER COMPETENCY

¹⁾Enny Wijayanti, ²⁾Mundilarto

¹⁾ Universitas Palangka Raya, ²⁾ Universitas Negeri Yogyakarta
¹⁾ennywijayanti28@yahoo.com, ²⁾mundilarto@uny.ac.id

Abstract

This study aimed to: (1) develop the construct of the Students' Self-Assessment (SSA) and Students' Peer-Assessment (SPA) as supplementary formative assessment on physics teaching and learning, (2) find out the characteristic of the SSA and SPA, (3) find out the effectiveness and (4) find out the usefulness of the SSA and SPA. The development procedure consisted of two stages: development and validation, based of the spiral model from Cennamo and Kalk, with five phases, namely *define, design, demonstrate, develop, and deliver*. *Content validity* was determined by expert judgment through *coefficient validity analysis* from Aiken, *focus group discussion*, and *Delphi technique*. Construct validity was analyzed using software SmartPLS. The subjects of this study were sixth semester students majoring on Physics Education at Palangka Raya University. The result of the study can be concluded as follows: (1) the result of model fit, the GFI values 0.996 and the SRMR values of 0.062, may be taken to indicate good fit, (2) the characteristic of the SSA and SPA consisted of validity, reliability, difficulty index and discriminant index had been verified, (3) the effectiveness of group SSA and SPA is better than group of non SSA and SPA, besides students' response declares SSA and SPA are reasonably effective, (4) SSA and SPA are very useful as feedback for the learning process, progress, and the learning outcome improvement.

Keywords: *Peer-and Self-assessment, Physics competency*

Pendahuluan

Dalam bidang pendidikan dan pengajaran terdapat tiga komponen yang dianggap sama penting, yaitu: tujuan, kegiatan pembelajaran, dan evaluasi. Ketiga komponen itu disebut dengan tiga mata jangkar (*three anchor points*) yang merupakan perpaduan atau kesatuan. Di antara ketiga komponen tersebut terdapat konsistensi internal, artinya ada hubungan yang konsisten.

Asesmen kelas (*classroom assessment*) di perguruan tinggi digunakan untuk melihat seberapa baik mahasiswa belajar di awal, menengah, dan untuk memberikan informasi perbaikan ketika hasil belajar mahasiswa kurang memuaskan. Pelaksanaan asesmen hasil belajar mahasiswa dilakukan oleh dosen, sehingga peran dosen sangat dominan dalam mengases pencapaian kompetensi mahasiswa. Hal ini sesuai dengan UU RI Nomor 20 Tahun 2003 Pasal 58 ayat (1) tentang Sistem Pendidikan Nasional yang menyebutkan, bahwa evaluasi hasil belajar peserta didik dilakukan oleh pendidik untuk memantau proses, kemajuan dan perbaikan hasil belajar secara berkesinambungan.

Oleh karena itu, informasi hasil asesmen kelas harus digunakan sebagai pemantauan proses dan hasil belajar. Untuk itu, diperlukan informasi hasil asesmen dengan segera. Namun, informasi hasil asesmen kelas dari dosen tidak segera diperoleh, sehingga mahasiswa tidak dapat mengetahui kesulitannya dalam memahami materi fisika.

Mengingat bahwa tujuan *classroom assessment* adalah sebagai penelusuran (*keeping-track*), pengecekan (*checking-up*), pencarian (*finding-out*), serta penyimpulan (*summing-up*), yaitu untuk menyimpulkan apakah mahasiswa telah menguasai seluruh kompetensi yang telah ditetapkan dalam kurikulum (Hayat, 2010, p.5). Oleh karena itu, asesmen formatif dimaksudkan untuk mengetahui sejauh mana mahasiswa telah menguasai suatu kompetensi setelah mengikuti pembelajaran. Dalam kedudukan seperti ini menurut Birenbaun & Dochy (1996), dalam Arends (2007, p.231), tes formatif dapat dipandang sebagai tes diagnostik.

Informasi hasil tes formatif berfungsi sebagai umpan balik (*feedback*) bagi dosen ataupun mahasiswa. Oleh karena itu, sebaiknya mahasiswa segera mendapatkan informasi dari hasil kinerjanya. Hal ini sejalan dengan pendapat Amien (1987, p.99) yang menyatakan bahwa memberikan umpan balik dengan segera perlu dilakukan oleh dosen dalam pembelajaran, agar mahasiswa dapat segera mengetahui kesulitannya dalam memahami materi. Hasil asesmen yang dilakukan dosen digunakan sebagai pemantauan proses pembelajaran, kemajuan belajar dan prestasi belajar mahasiswa. Dengan demikian, ketiga hal tersebut menjadi tuntutan pokok pada pelaksanaan asesmen oleh dosen. Namun, sering dalam praktik asesmen di perguruan tinggi, hal tersebut jarang dilakukan oleh dosen.

Upaya peningkatan kualitas proses pembelajaran memerlukan informasi hasil asesmen pada pembelajaran sebelumnya. Berdasarkan hasil asesmen tersebut dapat diketahui kekuatan dan kelemahan atau kesulitan mahasiswa dalam mengikuti pembelajaran. Hal ini sesuai dengan pendapat Nitko (1989, p.447), bahwa evaluasi seharusnya terintegrasi dengan pembelajaran, artinya terdapat kaitan erat antara asesmen dan pembelajaran. Dengan demikian, dapat dikatakan bahwa, upaya perbaikan mutu pendidikan tidak dapat dilepaskan dari pemanfaatan hasil asesmen.

Ujung tombak perbaikan mutu pendidikan adalah perbaikan pembelajaran, sehingga sistem asesmen harus dipakai sebagai bagian dari upaya perbaikan mutu (Kumaidi, 2001a). Hal ini menunjukkan bahwa asesmen merupakan komponen penting yang tak dapat dipisahkan dari proses pendidikan dan pembelajaran. Asesmen merupakan satu kesatuan yang tak terpisahkan, seperti yang sudah diuraikan.

Peningkatan kualitas pembelajaran melalui asesmen, didukung oleh Assessment Reform Group (ARG) (1999, p.4). Menurut ARG bahwa memberikan umpan balik dengan segera, melibatkan mahasiswa secara aktif dalam pembelajaran, serta melibatkan mahasiswa dalam mengases kompetensi diri

sendiri, penting dilakukan agar mahasiswa memahami bagaimana memperbaikinya. Hal senada juga dikatakan oleh Tola (2010, p.18) bahwa, keterlibatan mahasiswa dalam/proses asesmen merupakan bagian mendasar dari kesetimbangan asesmen. Dengan terlibat dalam kegiatan asesmen diri terhadap pencapaian kompetensi khususnya kognitif, mahasiswa akan mendapatkan informasi tentang kesulitan-kesulitan dalam menyelesaikan butir soal ataupun pada atribut-atribut, sehingga dapat segera dilakukan perbaikan terhadap belajarnya.

Keberhasilan proses belajar dapat dilihat dari hasil belajar yang dicapai. Keberhasilan ini selalu dikaitkan dengan tujuan pembelajaran atau kompetensi yang ingin dicapai. Berkaitan dengan hal itu, prestasi belajar fisika menunjukkan hasil yang kurang memuaskan. Hal ini terlihat dari hasil tes formatif pada mata kuliah keilmuan dan keterampilan I (MKK I) yang dilakukan oleh sebagian besar dosen pengampu mata kuliah tersebut. Hasilnya hampir tujuh puluh persen (70%) mahasiswa belum mencapai standar minimal lulus. Hasil ini tentu saja mengusik keprihatinan para dosen pengampu mata kuliah MKK I di Prodi Pendidikan Fisika FKIP Unpar.

Adanya kendala pada implementasi asesmen formatif dalam pembelajaran fisika, sesuai dengan hasil penelitian Kusairi (2010, p.7) yang menyebutkan bahwa ada kendala pada implementasi asesmen formatif dalam pembelajaran fisika. Penyebabnya adalah: (1) perencanaan dan pelaksanaan asesmen formatif memerlukan keterampilan, sementara masih banyak guru belum mendapatkan pelatihan profesional tentang teknik-teknik asesmen formatif; (2) pengembangan instrumen, implementasi dan analisis data-data asesmen formatif belum banyak dilakukan guru. Informasi ini menunjukkan bahwa pembelajaran fisika belum seperti yang diharapkan, artinya ada kendala yang terjadi dalam proses pembelajaran, terutama dalam pemanfaatan hasil asesmen formatif guna memperbaiki pembelajaran selanjutnya.

Hasil penelitian Kumaidi (2001a), menunjukkan bahwa, masih banyak guru kurang memanfaatkan hasil tes atau ujian untuk perbaikan proses pembelajaran. Hal ini menunjukkan bahwa ada kendala yang berkaitan dengan asesmen hasil belajar. Berdasarkan kendala-kendala yang terjadi pada praktik asesmen formatif perlu dicari penyelesaian yang sesuai.

Untuk mengatasi kelemahan dalam praktik asesmen formatif yang selama ini dilakukan diperlukan suatu asesmen pelengkap yang melibatkan mahasiswa dalam kegiatan asesmen terhadap pencapaian kompetensinya. Asesmen pelengkap tersebut salah satunya adalah asesmen diri (*self-assessment*). Hal ini sesuai dengan pendapat Black & William (1998, p.143) yang menyatakan bahwa *self-assessment* merupakan komponen utama dalam *assessment for learning*.

Umpan balik yang diperoleh mahasiswa dari hasil asesmen akan menimbulkan motivasi untuk segera memperbaiki cara belajarnya, menggunakan waktu belajar secara efektif, dan sebagainya. Bila hal ini terjadi maka dapat dikatakan bahwa hasil asesmen formatif dapat digunakan sebagai usaha perbaikan belajar. Hasil asesmen formatif berguna bagi dosen untuk mengetahui sejauh mana materi yang sudah diajarkan telah atau belum dikuasai oleh mahasiswa. Informasi ini digunakan untuk memperbaiki proses belajar berikutnya.

Atribut adalah kompetensi yang harus dimiliki peserta tes untuk menyelesaikan suatu butir soal. Perumusan atribut asesmen kognitif yang digunakan dalam penelitian ini diadopsi dari Tatsuoka, (2009, p.4). Menurut Tatsuoka asesmen kognitif (*cognitive assessment*) dipilah menjadi tiga atribut yaitu atribut isi (*content attributes*), atribut proses (*process attributes*), dan atribut skil atau keterampilan (*skill attributes*).

Pendapat para pakar tersebut, dapat disimpulkan bahwa melibatkan mahasiswa dan dosen dalam kegiatan asesmen terhadap proses dan hasil belajar adalah penting. Pada kegiatan asesmen diri, mahasiswa dilibatkan secara aktif dalam mengases penguasaan kompetensi diri sendiri secara jujur.

Hal ini dimaksudkan untuk mengetahui ketercapaian tujuan belajarnya, posisinya, dan bagaimana mengatasi kesenjangan antara tujuan dan ketercapaian kompetensi. Informasi yang diperoleh dari asesmen diri digunakan sebagai umpan balik, untuk segera dilakukan tindakan perbaikan dalam belajar.

Selanjutnya, Black & William (1998, p.14) menyatakan bahwa *peer-assessment* dalam praktik menjadi pelengkap dari asesmen sebelumnya dan sebagai syarat adalah *self-assessment*. Untuk melengkapi asesmen formatif, mahasiswa dilibatkan dalam kegiatan mengases kompetensi teman sejawat dalam kelompok, setelah melakukan kegiatan asesmen diri. Kegiatan saling mengases penguasaan kompetensi antarsejawat dalam kelompok disebut dengan istilah asesmen teman sejawat (ATS).

Asesmen teman sejawat adalah bentuk asesmen yang dilakukan untuk memperoleh informasi balikan yang diperoleh dari teman sejawat, selain yang diperoleh dari dosen. Pada kegiatan ATS mahasiswa terlibat saling mengakses penguasaan kompetensi teman sejawat dalam kelompok. Asesmen teman sejawat diterapkan sebagai pelengkap asesmen formatif, sehingga ATS dikembangkan untuk mengetahui kemajuan dan perkembangan belajar mahasiswa.

Sebagai asesmen pelengkap, ATS mempunyai banyak manfaat bagi mahasiswa antara lain: dapat melatih mahasiswa memiliki rasa tanggung jawab, kerja sama dalam belajar antarsejawat dalam kelompok. Mahasiswa menjadi berani merefleksikan kinerjanya berdasarkan standar ketuntasan belajar yang sudah ditentukan dan membandingkannya dengan yang lain, sehingga menimbulkan semangat berusaha dalam mencapai tujuan belajarnya. Untuk mengetahui kesulitan dan respon mahasiswa pada uji validasi ATS digunakan lembar refleksi ATS.

Oleh karena itu, pembaharuan asesmen dalam bidang pendidikan diperlukan pada saat ini adalah asesmen yang menunjang proses pembelajaran di kelas. Jadi, dengan mengembangkan instrumen ADM & ATS sebagai pelengkap asesmen formatif diharapkan akan memberikan kontribusi

yang berarti pada perbaikan proses pendidikan.

Penerapan kombinasi ADM & ATS menurut Sadler., White & Frederikson (Haris, 2007, p.28) adalah salah satu cara untuk memperbaiki praktik asesmen formatif yaitu dengan menggunakan *peer-and self-assessment*. Selanjutnya, dikatakan bahwa *peer-and self-assessment* merupakan esensi dari pembelajaran (*essential to learning*). Pakar lain yang mendukung ADM & ATS adalah Hagstrom (2006, p.28) menyatakan bahwa salah satu keuntungan menggunakan *peer-and self-assessment* dalam pembelajaran akan meningkatkan *self-esteem*.

Fungsi kegiatan ADM & ATS yang digunakan dalam pembelajaran adalah pencapaian tujuan atau penguasaan kompetensi dan koreksi diri terhadap proses dan pencapaian hasil belajar. Ketika mahasiswa melakukan kegiatan ADM dan ATS, artinya mahasiswa belajar mengembangkan kebiasaan untuk merefleksi diri sendiri.

Mengingat kegiatan ini baru pertama dilakukan, kemungkinan terjadi kesalahan yang dilakukan mahasiswa dalam melakukan asesmen diri ataupun dalam asesmen teman sejawat. Untuk mengatasi terjadinya perbedaan dalam kegiatan asesmen ini, maka asesmen dari dosen sebagai acuan untuk konfirmasi apakah asesmen diri yang dilakukan mahasiswa dan asesmen teman sejawat arahnya *off the track* atau *on the track*.

Penggunaan ADM & ATS secara bersama-sama belum banyak dilakukan oleh dosen dalam pembelajaran pada umumnya serta pembelajaran fisika pada khususnya. Oleh karena itu, perlu dikaji konstruk instrumen ADM & ATS, karakteristik ADM & ATS, keefektifan ADM & ATS dan kebermanfaatannya ADM & ATS. Hal ini untuk mengetahui apakah ADM & ATS dapat meningkatkan pencapaian kompetensi kognitif, afektif serta psikomotor bila digunakan dalam pembelajaran fisika

Teknik ATS yang digunakan diadopsi dari Wiersema (2000, p.3), yaitu evaluasi kelompok/sejawat (*group/peer evaluation*), sedangkan metodenya diadopsi dari Kane & Lawler (Keaten, Richardson & Elisabeth,

1993, p. 3) yaitu *peer rating*. Apabila digunakan untuk mendeteksi kesulitan belajar mahasiswa, menggunakan *peer-rating* lebih tepat karena memberikan informasi yang rinci tentang tahapan kecakapan yang sudah berhasil dikuasai dan yang belum dikuasai (Latham & Wexley, 1982, p. 88).

Asesmen diri mahasiswa dan asesmen teman sejawat digunakan sebagai pelengkap asesmen formatif. Asesmen pelengkap menjadi sangat penting, karena dapat memberikan informasi yang lebih banyak tentang kemampuan mahasiswa baik unjuk kerja individu maupun dalam kelompok, asesmen terhadap dirinya sendiri dan bukan sekedar memperoleh informasi respon benar, salah dari dosen.

Asesmen formatif dari dosen juga dapat memberikan umpan balik kepada mahasiswa, namun informasi tersebut tidak segera diperoleh. Oleh karena itu, ADM dan ATS yang bersama-sama digunakan sebagai pelengkap asesmen formatif merupakan kebutuhan yang sangat penting. Oleh sebab itu, perlu dikembangkan dalam rangka mengungkap unjuk kerja kompetensi fisika pada domain kognitif, afektif dan psikomotor secara seimbang.

Penerapan kombinasi ADM & ATS didukung oleh beberapa ahli. Menurut Sadler (1989); White & Frederikson, (1998) dalam Haris (2007, p.28), bahwa salah satu cara untuk memperbaiki praktik asesmen formatif dalam pembelajaran fisika adalah menggunakan *peer and self-assessment*. Dalam hal ini ADM & ATS merupakan esensi dari pembelajaran (*essential to learning*). Selanjutnya, Hagstrom (2006, p.28) mengatakan bahwa salah satu keuntungan menggunakan *peer and self-assessment* dalam pembelajaran adalah akan meningkatkan *self-esteem*.

Pengukuran domain kognitif diturunkan dari standar kompetensi (SK), kompetensi dasar (KD) mata kuliah getaran gelombang. Selanjutnya, dilakukan identifikasi materi prasyarat hanya pada indikator-indikator yang bersesuaian dengan materi getaran gelombang. *Perumusan learning continuum* menggunakan indikator-indikator yang disepakati bersama pada kegiatan FGD. Materi pokok

yang digunakan, yaitu getaran (GT), gelombang (BL), dan bunyi (BN).

Domain afektif merupakan tindakan, sikap, dan perilaku Peringkat domain afektif yang diases diadopsi dari Trowbridge & Bybee (1986, p.134) ada lima kategori/*level* yaitu: *receiving, responding, valuing, organizing dan creating*. Pada level *receiving* mahasiswa memiliki keinginan memperhatikan suatu fenomena khusus, yaitu kegiatan ADM & ATS. *Responding* merupakan partisipasi aktif mahasiswa, yaitu sebagai bagian dari perilakunya. *Valuing* melibatkan penentuan nilai, keyakinan atau sikap dan komitmen.

Selanjutnya, domain psikomotor berkaitan dengan kemampuan gerak. Peringkat domain psikomotor yang diadopsi dari Trowbridge & Bybee (1986, p.135) ada empat kategori, yaitu: *moving, manipulating, communicating dan creating*. Acuan dalam asesmen kompetensi kognitif menggunakan pedoman penskoran sedangkan acuan dalam asesmen domain psikomotor menggunakan rubrik.

Pendekatan laboratorium deduktif dipilih karena pendekatan pembelajaran ini melibatkan mahasiswa dengan melakukan verifikasi atau konfirmasi terhadap teori yang telah diberikan sebelum dilakukan kegiatan laboratorium. Mengingat penekanan penelitian ini pada pengembangan instrumen asesmennya bukan pada pengembangan model pembelajarannya maka dipilih pendekatan laboratorium deduktif.

Pengetahuan fisika umumnya banyak menggunakan pengetahuan dan kemampuan pemecahan masalah yang proses belajarnya bersifat kompleks dan prosedural. Dengan demikian belajar secara kelompok relevan dan cocok digunakan untuk pembelajaran jenis pengetahuan prosedural (Johnson & Johnson, 1987, p.27). Kegiatan laboratorium dituntut untuk belajar bersama dalam kelompok dengan pembagian kerja/kolaboratif, balikan atau refleksi interaksi.

Asesmen pada pembelajaran kelompok melalui latihan soal dan kegiatan praktikum tidak berasumsi bahwa prestasi belajar adalah hasil dari kinerja individu sendiri, melainkan ada pengaruh kemampuan teman sejawat ketika individu terlibat dalam proses

belajar bersama. Mahasiswa yang kurang mampu dibantu belajar oleh teman-temannya yang lebih mampu dalam kelompok, sehingga dapat dikatakan bahwa asesmen teman sejawat mendukung belajar mahasiswa untuk mencapai tujuan pembelajaran. Oleh karena itu, ADM dan ATS menjadi kebutuhan yang perlu dikembangkan.

Terkait dengan hal tersebut, penelitian ini bertujuan untuk: (1) mengembangkan konstruk instrumen asesmen diri (ADM) dan asesmen teman sejawat (ATS) sebagai pelengkap asesmen formatif, (2) mengetahui karakteristik ADM dan ATS, (3) mengetahui efektivitas ADM & ATS, dan (4) mengetahui kebermanfaatan ADM & ATS yang digunakan sebagai pelengkap asesmen formatif.

Metode

Penelitian ini merupakan penelitian pengembangan disertai dengan eksperimen semu (*quasi experiment*). Asesmen teman sejawat dan asesmen diri mahasiswa mengintegrasikan kegiatan pengukuran hasil belajar dengan keseluruhan proses pembelajaran. Oleh karena itu, dipilih *model Spiral Cennamo dan Kalk* (2005, p.6) yang dikenal dengan lima fase pengembangan yaitu: (1) pendefinisian (*define*), sebelumnya dilakukan *preliminary study*, (2) desain (*design*), (3) peragaan (*demonstrate*), (4) pengembangan (*develop*) dan (5) penyajian (*deliver*). Tahap pengembangan meliputi, *define*, *desain*, *demonstrate*, sedangkan tahap validasi terdiri dari *develop* dan *deliver*.

Pengumpulan data dilakukan dalam dua tahap, yaitu tahap pengembangan dan tahap validasi. Pada tahap pengembangan, data dikumpulkan melalui (a) wawancara, (b) observasi, (c) kuesioner, (d) dokumentasi, (e) *focus group discussion* (FGD), dan (f) teknik Delphi, sedangkan tahap validasi dilakukan uji keterbacaan dan review ahli, uji kelompok kecil, uji kelompok diperluas dan uji validasi ADM & ATS.

Penelitian pengembangan produk dilakukan mulai 29 Januari 2013 sampai Agustus 2013. Subjek uji coba terbatas adalah mahasiswa di Program Studi Tadris Fisika Jurusan Tarbiyah STAIN Palangka Raya

dan uji coba diperluas adalah mahasiswa di Prodi Pendidikan Fisika Unpar yang sudah menempuh mata kuliah getaran gelombang. Subjek uji validasi ADM & ATS, adalah mahasiswa semester enam yang sedang menempuh mata kuliah getaran gelombang, terbagi dalam dua kelas. Kelas A sebagai kelompok pembanding dan kelas B sebagai kelompok uji validasi ADM & ATS.

Tahap Pengembangan

Tahap pengembangan meliputi kegiatan fase *define*, *design* dan *demonstrate*. Kegiatan awal melakukan *preliminary study*, melakukan review pustaka dan kajian penelitian yang relevan, observasi dan identifikasi terhadap proses pembelajaran fisika yang terjadi, dan asesmen hasil belajar yang dipakai. Untuk mendapatkan gambaran tentang pelaksanaan pembelajaran serta kesulitan mahasiswa dalam mengikuti perkuliahan fisika, perlu dilakukan wawancara terhadap dosen pengampu mata kuliah getaran gelombang serta mahasiswa. Informasi yang diperoleh digunakan untuk menyempurnakan *draft* rancangan pengembangan *draft* asesmen. Kegiatan selanjutnya merumuskan, merancang *learning continuum* kompetensi kognitif, pedoman penskoran, kisi-kisi instrumen afektif, instrumen psikomotor dan rubrik serta refleksi ADM dan ATS. Tahap pengembangan menghasilkan prototipe 1 preskripsi ADM & ATS. Sampel materi terdiri dari tiga materi pokok, yaitu getaran, gelombang dan bunyi, yang disusun dalam tujuh SAP. Oleh karena itu, diperlukan waktu tujuh kali tatap muka, atau tujuh minggu, ditambah satu minggu untuk sosialisasi melatih mahasiswa menggunakan pedoman penskoran.

Tahap Validasi

Rancangan yang sudah dihasilkan pada tahap pengembangan, selanjutnya dilakukan uji keterbacaan dan review ahli, analisis dan revisi 1 yang menghasilkan prototipe 2 preskripsi ADM & ATS, serta instrumen pendukung. Hasil analisis selanjutnya dilakukan uji coba terbatas, untuk melihat ketercukupan waktu, reliabilitas instrumen. Subjek uji coba terbatas adalah mahasiswa

Tadris Fisika Jurusan Tarbiyah STAIN Palangka Raya, yang berjumlah 24 mahasiswa. Namun, sebelum instrumen digunakan, sebelumnya sudah dilakukan analisis *content validity ratings* menggunakan formula Aiken melalui *judgment* rater. Hasilnya semua instrumen valid secara konten. Selanjutnya, hasil uji coba dilakukan analisis dan revisi 2, menghasilkan model tentatif. Uji coba diperluas dilakukan pada mahasiswa Prodi Pendidikan Fisika FKIP Unpar yang sudah menempuh mata kuliah getaran gelombang. Uji coba diperluas untuk melihat reliabilitas komposit, tingkat kesukaran dan daya pembeda butir. Hasil uji coba diperluas, kemudian dilakukan analisis dan revisi 3. Selanjutnya, instrumen ADM & ATS beserta instrumen pendukung digunakan dalam pembelajaran fisika yang dilakukan selama tujuh kali tatap muka. Berdasarkan hasil uji coba ini dilakukan analisis menggunakan program SmartPLS 2.0M3, untuk melihat validitas konstruk, reliabilitas komposit, *R-Square*. Untuk melihat apakah model yang dikembangkan telah sesuai antara teori dan penerapannya di lapangan, dilakukan uji kecocokan model digunakan *software* program *generalized structured component analysis* (GSCA) (Heungsun Hwang, 2011).

Teknik Analisis Data

Validitas isi digunakan untuk melihat apakah semua instrumen yang akan digunakan sudah valid secara konten, menggunakan formula Aiken's (1985, p.132-133) sebagai berikut.

$$V = \frac{\sum s}{n(c-1)}$$

Keterangan:

- c : banyaknya kategori
- r : skor penilaian rater
- n : banyaknya rater
- lo : skor rating minimal
- s : r - lo (skor dari rater-skor terendah) untuk setiap butir; ($\sum s = S$).

Kriteria yang digunakan untuk menentukan butir-butir tersebut valid secara konten adalah dengan membandingkan harga V hitung dengan V tabel V pada Tabel

Right-tail Probabilities (p) for selected values of the *validity coefficient (V)* dengan $p = 0,05$ pada kategori rating 4 (empat) dan banyaknya rater ($n = 9$ (sembilan)), pada tabel menunjukkan harga 0,74. Jika V hitung > harga tabel 0,74, maka instrumen tersebut valid secara konten. Untuk menghitung besarnya koefisien reliabilitas interrater diadopsi dari Aiken (1980, p.959). Rumus koefisien reliabilitas interrater Aiken's berikut.

$$R = 1 - \frac{\sum_{j=0}^2 \sum_{i=0}^2 n_{ij} |i-j|}{N(c-1)}$$

Keterangan:

- n_{ij} : banyaknya sel pada tabel matriks 3x3
- N : banyaknya rater dan
- c : banyaknya kategori

Reliabilitas uji coba terbatas diestimasi dengan Cronbach Alpha. Teknik analisis data uji coba diperluas, meliputi tingkat kesukaran (TK) dan daya pembeda (DP) bentuk tes uraian objektif dihitung menggunakan program excel. Menurut Crocker & Algina (1986, p.311), Ebel & Frisbie (1986, p.231), Gulliksen (1950, p.366), Linn, & Gronlund (2009, p.356), Reynolds, Livingston, & Willson (2010, pp. 148-149). Kesulitan butir didefinisikan sebagai proporsi dari jawaban-jawaban yang benar. Dengan demikian, menurut Ebel & Frisbie (1986, p.231), semakin tinggi indeks kesulitan semakin mudah butir/tes yang bersangkutan. Berdasarkan definisi kesukaran butir di atas, maka persamaan kesulitan butir untuk tes uraian objektif adalah sebagai berikut.

$$P_j = \frac{\sum_{i=1}^n X_{ij}}{nX_{jmax}}$$

Keterangan:

- P_j : proporsi jawaban benar atau indeks kesulitan butir ke- j .
- n : ukuran sampel
- nX_{jmax} : skor maksimal untuk butir ke- j
- $\sum_{i=1}^n X_{ij}$: jumlah skor butir ke- j dari peserta tes ke- i , dengan i mulai dari 1 sampai ke n .

Menurut McDonald (1999, p.78), Miller, Linn, & Gronlund (2009, p.357), Reynolds, Livingston, & Willson (2010,

p.150) diskriminasi butir atau daya pembeda butir adalah suatu indeks yang merujuk pada derajat bagaimana suatu butir membedakan antara peserta tes yang mendapatkan skor tinggi dan skor rendah pada butir tes tertentu. Menurut Ebel & Frisbie (1986, p.230) jika tujuan utama seleksi butir adalah untuk memaksimalkan reliabilitas tes, maka butir yang memiliki diskriminasi tinggi adalah butir yang harus dipilih. Rumus untuk menghitung indeks daya pembeda (DP) butir sebagai berikut.

$$DP = \frac{\text{Mean kelompok atas} - \text{Mean kelompok bawah}}{\text{Skor minimum soal}}$$

Kriteria yang digunakan untuk menentukan kategorisasi indeks DP diadaptasi dari (Cracker & Algina, 1986, p. 315)

Pada analisis butir acuan kriteria, jika seluruh mahasiswa sudah berhasil menguasai indikator dari suatu kompetensi, maka daya bedanya akan sebesar 0. Namun, butir tersebut tetap dinyatakan baik dan tetap dapat dipakai untuk menunjukkan efektivitas proses (Mardapi, 2012, p.188).

Untuk mengukur reliabilitas pengukuran observasi kegiatan laboratorium dilakukan oleh tiga observer, dan dianalisis dengan reliabilitas ratings atau reliabilitas pengukuran observasi diadopsi dari Ebel (Azwar, 1986, p.44).

Kategorisasi penguasaan kompetensi kognitif untuk masing-masing materi pokok, digunakan kriteria yang diadopsi dari Mardapi (2012, p.107), disajikan pada Tabel 1.

Tabel 1. Kategorisasi Penguasaan Kognitif dan Psikomotor

No.	Tingkat Penguasaan (TP)	Kategori
1.	90% - 100%	Baik Sekali
2.	80% - 89%	Baik
3.	70% - 79%	Cukup
4.	< 70%	Kurang

$$TP = \frac{\text{Jumlah jawaban benar}}{\text{Skor maksimal}} \times 100\%$$

Kategorisasi kompetensi afektif yang diadopsi dari Mardapi (2012, p.162), disajikan pada Tabel 2.

Data hasil validasi empiris digunakan untuk mendeskripsikan keefektifan, dan kebermanfaatan ADM & ATS dianalisis dengan program excel. Teknik analisis data kualitatif digunakan untuk menjelaskan prosedur pengembangan model ADM & ATS.

Tabel 2. Kategorisasi Penguasaan Kognitif dan Psikomotor

No.	Tingkat Penguasaan	Kategori
1	$X \geq \bar{X} + 1.SB_x$	Baik Sekali
2	$\bar{X} + 1.SB_x > X \geq \bar{X}$	Baik
3	$\bar{X} > X \geq \bar{X} - 1.SB_x$	Cukup
4	$X < \bar{X} - 1.SB_x$	Kurang

Hasil Penelitian dan Pembahasan

Pengembangan ADM & ATS ini difokuskan untuk mengukur penguasaan kompetensi kognitif mahasiswa, namun kompetensi afektif dan psikomotor juga diukur untuk melihat pengaruh ADM & ATS pada kompetensi tersebut. Hal ini merupakan keterbatasan dari penelitian ini, selanjutnya perlu dilakukan penelitian lanjutan yang mengases kompetensi afektif dan psikomotor dengan ADM & ATS dengan menggunakan sampel yang lebih besar.

Pengembangan ADM & ATS diawali dengan proses kajian teoritik, hasil penelitian yang relevan, dilanjutkan dengan observasi terhadap praktik pembelajaran fisika di lapangan. Berdasarkan informasi yang diperoleh digunakan untuk membuat *draft* pengembangan asesmen. Merumuskan dan merancang *draft* instrumen ADM & ATS kompetensi kognitif yang dirumuskan melalui *learning continuum* materi getaran gelombang. Dasar perumusan *learning continuum* adalah indikator-indikator yang telah disepakati bersama pada kegiatan FGD. Kegiatan pada tahap pengembangan menghasilkan prototipe 1, sedangkan kegiatan tahap validasi secara umum merupakan kegiatan uji coba. Uji coba terbatas, uji coba diperluas, analisis dan revisi, menghasilkan model tentatif. Untuk melihat konstruk, keefektifan dan ke-

terpakaian ADM & ATS, dilakukan secara empiris di dalam pembelajaran fisika.

Hasil estimasi *content validity* dari Aiken's, bila dirangkum secara keseluruhan hasilnya disajikan pada Tabel 3.

Tabel 3. Hasil Validasi Isi Instrumen

Nama Instrumen	V		Kategori	Hasil
	rata-rata	tabel		
Pedoman Penskoran Kognitif	0,95	0,74	valid	DDRK
Refleksi ADM	0,94	0,74	valid	DDRK
Refleksi ATS	0,90	0,74	valid	DDRK
Afektif	0,89	0,88	valid	DDRK
Psikomotor	0,93	0,74	valid	DDRK
LKM	0,96	0,74	valid	DDRK
Rubrik Observasi Kegiatan Laboratorium	0,89	0,74	valid	DDRK
Refleksi Kegiatan Laboratorium deduktif	0,91	0,74	valid	DDRK

Hasil analisis *Content Validity* Aiken menunjukkan bahwa semua *draft* instrumen yang digunakan pada penelitian ini telah divalidasi rater memenuhi validitas konten, artinya instrumen valid secara konten. Hasil semua instrumen dapat digunakan dengan revisi kecil (DD-RK). Bila hasil *raters* tersebut dianalisis secara deskriptif kualitatif, hasilnya seperti yang nampak pada Tabel 4.

Tabel 4. Kategorisasi Penilaian instrumen oleh Raters pada Prauji Coba

Nama Instrumen	Rata-rata skor	Rentang Skor	Kategori
Pedoman Penskoran Kognitif	30,44	$X > 26$	Sangat Baik
Refleksi ADM	49,56	$X > 42,25$	Sangat Baik
Refleksi ATS	34,11	$X > 29,25$	Sangat Baik
Afektif (AFE)	116,40	$X > 100$	Sangat Baik
Psikomotor (PSI)	21,78	$X > 19,5$	Sangat Baik
LKM	26,44	$X > 22,75$	Sangat Baik
Rubrik Observasi Kegiatan Laboratorium	64,44	$X > 55,25$	Sangat Baik
Refleksi Kegiatan Laboratorium deduktif	20,44	$X > 18,95$	Sangat Baik

Tabel 4 menunjukkan bahwa skor rata-rata untuk semua instrumen nilainya di atas rentang skor pada kategori sangat baik. Dengan demikian, dapat disimpulkan bahwa berdasarkan penilaian raters, semua instrumen berada pada kategori sangat baik. Hasil perhitungan koefisien reliabilitas antarrater disajikan pada Tabel 5.

Tabel 5. Reliabilitas Antar-Raters Aiken

No	Instrumen	Nilai Koefisien Reliabilitas Antarrater (R)	Ket.
1	Kegiatan Laboratorium Deduktif (Psikomotor)	0,94	Reliabel
2	Pedoman Penskoran	0,94	Reliabel
3	Afektif (AFE)	0,875	Reliabel
4	Refleksi ATS	0,75	Reliabel
5	Refleksi ADM	0,96	Reliabel
6	Lembar Kegiatan Mahasiswa (LKM)	0,96	Reliabel
7	Rubrik Observasi Kegiatan Laboratorium	0,90	Reliabel

Tabel 5 menunjukkan bahwa semua nilai reliabilitas *draft* instrumen pada tahap pra uji coba mempunyai nilai lebih besar dari 0,70, yang merupakan nilai reliabilitas minimal yang disyaratkan, sehingga dapat disimpulkan bahwa semua instrumen reliabel.

Uji terbatas dimaksudkan untuk melihat kejelasan tiap butir soal dan ketercukupan waktu, serta reliabilitas instrumen. Reliabilitas instrumen dihitung dengan rumus alpha, hasil perhitungan reliabilitas sebesar 0,73. Hasil analisis uji coba diperluas, meliputi tingkat kesulitan, indeks daya pembeda, reliabilitas komposit, berturut-turut disajikan pada Tabel 6 & Tabel 7.

Tabel 6 menunjukkan bahwa nilai TK untuk seluruh butir berada pada rentang 0,30-0,70. Dengan demikian, TK butir tes kognitif termasuk dalam kategori sedang. Oleh karena itu, seluruh butir dapat diterima. Nilai DP untuk butir 9, 20, dan 21 berada pada rentang 0,30-0,40, sehingga butir-butir tersebut memiliki DP yang tinggi.

Tabel 6. Analisis Tingkat Kesulitan (TK) dan Daya Pembeda (DP) Butir Kognitif

Materi Pokok	Butir	Karakteristik Butir		Kategori		Ket		
		TK	DP	TK	DP			
Getaran (GT)	SAP 1	1	0,300	0,423	S	ST	D	
	Osilasi Harmonik Sederhana pada Pegas	2	0,430	0,731	S	ST	D	
		3	0,308	0,776	S	ST	D	
	SAP 2	4	0,430	0,769	S	ST	D	
		5	0,498	0,657	S	ST	D	
	Osilasi Harmonik Sederhana pada Ayunan Matematis	6	0,576	0,444	S	ST	D	
Gelombang (GL)	SAP 3	7	0,320	0,923	S	ST	D	
	Persamaan Getaran	8	0,459	0,538	S	ST	D	
		9	0,504	0,399	S	T	D	
		10	0,530	0,548	S	ST	D	
	SAP 4	Gelombang	11	0,515	0,462	S	ST	D
			12	0,420	0,757	S	ST	D
	SAP 5	Energi Gelombang	13	0,384	0,657	S	ST	D
			14	0,638	0,581	S	ST	D
			15	0,471	0,487	S	ST	D
16			0,633	0,650	S	ST	D	
17			0,583	0,615	S	ST	D	
Bunyi (BN)	SAP 6	Taraf Intensitas Bunyi	18	0,693	0,731	S	ST	D
			19	0,697	0,846	S	ST	D
	SAP 7	Efek Doppler	20	0,627	0,256	S	T	D
			21	0,698	0,333	S	T	D
	SAP 7	Efek Doppler	22	0,694	0,469	S	ST	D
			23	0,689	0,633	S	ST	D
			24	0,697	0,628	S	ST	D
			25	0,700	0,550	S	ST	D
Rerata		0,540	0,595	S	ST			

Tabel 7. Reliabilitas Komposit Uji Coba II

No.	Instrumen	Reliabilitas komposit
1.	KOG	0,751
2.	AFE	0,601
3.	ADM	0,785
4.	ATS	0,790

Uji coba ketiga dilakukan dalam pembelajaran fisika, sebagai subjek coba adalah kelas A sebagai kelompok kontrol dan kelas B sebagai kelompok uji validasi ADM & ATS. Pada uji empiris ini untuk melihat validitas konstruk ADM dan ATS, reliabilitas komposit, ketercapaian pada kompe-

tensi kognitif, afektif dan psikomotor antara kelompok kontrol dan kelompok uji ADM & ATS. Reliabilitas kompetensi psikomotor di-ukur melalui kegiatan praktik yang diases oleh tiga observer. Hasil pengukuran reliabilitas antarrater sebesar 0,84. Hasil analisis uji coba ketiga atau validasi ADM & ATS.

Tabel 8. Hasil Estimasi Reliabilitas Komposit

No	Konstruk	Komponen	Composite Reliability
1.	KOG	GT	0,735
		GL	0,745
		BN	0,773
		KOG	0,852
		AFE	
2.	AFE	A1	0,717
		A2	0,581
		A3	0,469
		A4	0,604
		A5	0,636
		AFE	0,780
3.	ADM	GT	0,788
		GL	0,822
		BN	0,744
		ADM	0,876
		ATS	
4.	ATS	GT	0,866
		GL	0,795
		BN	0,710
		ADM	0,873
		ATS	

Pada Tabel 9 nampak bahwa pada instrumen ADM, komponen GT, GL dan BN, hasil nilai T-statistiknya, ada beberapa yang < 1,96, yaitu: GL02, GL07, GL10, BN18, BN20, BN21, BN25. Namun demikian, hal itu memiliki nilai positif, sehingga membuangnya dari instrumen dapat menurunkan tingkat reliabilitas, sehingga tetap dipertahankan. Selain itu, instrumen sudah dinyatakan valid melalui *expert judgment*. Dengan demikian dapat disimpulkan bahwa semua indikator ADM kecuali tujuh indikator yang digunakan adalah valid.

Tabel 9. Results for Outer Weights

Komponen	NO Butir	Original	Sample	Standart	Standart	T-Statistic
		Sample (O)	Mean (M)	Deviation (STDEV)	Error (STERR)	
Getaran (GT)	GT01	0,276	0,273	0,045	0,045	6,117
	GT02	0,207	0,179	0,107	0,107	1,928
	GT03	0,288	0,282	0,057	0,057	5,038
	GT04	0,230	0,224	0,058	0,058	3,930
	GT05	0,355	0,358	0,067	0,067	5,263
	GT06	0,225	0,211	0,066	0,066	3,408
ADM Gelombang (GL)	GL07	0,080	0,073	0,050	0,050	1,605
	GL08	0,210	0,197	0,027	0,027	7,847
	GL09	0,154	0,165	0,030	0,030	5,115
	GL10	0,083	0,089	0,044	0,044	1,883
	GL11	0,127	0,148	0,054	0,054	2,365
	GL12	0,164	0,138	0,038	0,038	4,364
	GL13	0,192	0,197	0,031	0,031	6,210
	GL14	0,102	0,100	0,031	0,031	3,241
	GL15	0,206	0,199	0,032	0,032	6,365
	GL16	0,155	0,137	0,037	0,037	4,170
	GL17	0,230	0,238	0,034	0,034	6,667
Bunyi (BN)	BN18	0,106	0,072	0,094	0,094	1,134
	BN19	0,214	0,166	0,093	0,093	2,318
	BN20	0,111	0,100	0,086	0,086	1,287
	BN21	0,163	0,152	0,089	0,089	1,825
	BN22	0,338	0,336	0,042	0,042	7,977
	BN23	0,339	0,327	0,051	0,051	6,634
	BN24	0,263	0,249	0,077	0,077	3,430
	BN25	0,157	0,186	0,084	0,084	1,878
	Getaran (GT)	GT01	0,239	0,240	0,030	0,030
GT02		0,252	0,249	0,035	0,035	7,140
GT03		0,217	0,213	0,035	0,035	6,183
GT04		0,179	0,172	0,030	0,030	5,974
GT05		0,264	0,266	0,023	0,023	11,679
GT06		0,233	0,232	0,024	0,024	9,835
ADM Gelombang (GL)	GL07	0,115	0,124	0,043	0,043	2,656
	GL08	0,224	0,211	0,035	0,035	6,401
	GL09	0,162	0,153	0,038	0,038	4,207
	GL10	0,086	0,073	0,049	0,049	1,754
	GL11	0,112	0,106	0,043	0,043	2,607
	GL12	0,195	0,180	0,041	0,041	4,762
	GL13	0,182	0,169	0,030	0,030	6,008
	GL14	0,143	0,131	0,040	0,040	3,584
	GL15	0,215	0,209	0,043	0,043	4,952
	GL16	0,180	0,172	0,034	0,034	5,273
	GL17	0,224	0,219	0,042	0,042	5,332
Bunyi (BN)	BN18	0,226	0,223	0,068	0,068	3,314
	BN19	0,160	0,133	0,121	0,121	1,330
	BN20	0,256	0,231	0,088	0,088	2,919
	BN21	0,221	0,220	0,090	0,090	2,448
	BN22	0,305	0,261	0,067	0,067	4,579
	BN23	0,356	0,320	0,064	0,064	5,568
	BN24	0,231	0,190	0,140	0,140	1,650
	BN25	0,209	0,205	0,098	0,098	2,120

Hasil estimasi validitas konstruk instrumen ATS nampak bahwa komponen GL dan BN ada yang nilai T-statistiknya < 1,96, yaitu GL10, BN19 dan BN24. Namun demikian, hal itu memiliki nilai positif, bila dibuang dari instrumen dapat menurunkan tingkat reliabilitas, sehingga tetap dipertahankan. Selain itu, instrumen sudah dinyatakan valid melalui *expert judgment*. Dengan demikian, dapat disimpulkan bahwa semua indikator ATS kecuali tiga indikator tersebut adalah valid. Bila dilihat dari uji *result for inner weights* instrumen ADM, ATS, KOG, AFE dan PSI, hasilnya seperti Tabel 10.

Tabel 10. Result for inner weights

Komponen	Original Sample (O)	Sample Mean (M)	Std. Deviation (STDEV)	Std. Error (STERR)	T-Statistic	
ADM	ADM -> GT	0,820	0,845	0,036	0,036	22,686
	ADM -> GL	0,915	0,909	0,028	0,028	33,236
	ADM -> BN	0,796	0,809	0,046	0,046	17,134
ATS	ATS -> GT	0,871	0,876	0,032	0,032	26,814
	ATS -> GL	0,895	0,905	0,024	0,024	36,998
	ATS -> BN	0,714	0,747	0,054	0,054	13,219
KOG	KOG -> GT	0,836	0,849	0,036	0,036	22,983
	KOG -> GL	0,850	0,871	0,028	0,028	30,430
	KOG -> BN	0,868	0,892	0,032	0,032	27,533
AFE	AFE -> A1	0,869	0,866	0,060	0,060	14,448
	AFE -> A2	0,657	0,698	0,079	0,079	8,319
	AFE -> A3	0,555	0,577	0,102	0,102	5,417
	AFE -> A4	0,739	0,751	0,081	0,081	9,141
	AFE -> A5	0,730	0,775	0,098	0,098	7,470
PSI	PSI -> P1	0,905	0,804	0,435	0,435	2,078
	PSI -> P2	0,918	0,623	0,674	0,674	1,362
	PSI -> P3	0,886	0,267	0,853	0,853	1,039

Tabel 10 menunjukkan bahwa seluruh komponen ADM, ATS, KOG, AFE, dan PSI memiliki nilai T-statistik > 1,96, kecuali komponen PSI, yaitu P2 dan P3 nilai T-statistiknya < 1,96, tetapi nilainya positif. Dengan demikian, secara umum dapat disimpulkan instrumen ADM, ATS, KOG, AFE dan PSI tersebut valid.

Nilai reliabilitas komposit instrumen KOG, AFE, ADM dan ATS uji coba ketiga/ uji validasi, secara keseluruhan berturut-turut 0,852, 0,780, 0,876 dan 0,873, sehingga dapat disimpulkan semua instrumen di atas reliabel.

Untuk membuktikan validitas konstruk digunakan hasil T-statistik, kriteria digunakan nilai kritis untuk uji t pada taraf signifikansi $\alpha = 5\%$ yaitu 1,96. *Result for Outer Weights* menunjukkan bahwa semua indikator konstruk ADM dan ATS valid, jika nilai T-statistik yang dihasilkan $>1,96$ (Ghazali & Latan, 2012, p.133).

Hasil uji kecocokan model ADM & ATS, menunjukkan bahwa nilai FIT sebesar 0,671 dan AFIT sebesar 0,641, artinya bahwa model yang dispesifikasi mampu menjelaskan 67,1% varians pada data. Nilai GFI sebesar 0,994, dan SRMR sebesar 0,062 menunjukkan fit yang baik karena nilai GFI mendekati 1 dan $SRMR \leq 0,08$. Dengan demikian, dapat disimpulkan bahwa model teoretis yang dikembangkan didukung oleh data empirik.

Tabel 11. Ketercapaian Kompetensi Kognitif

Materi Kegiatan Belajar	Skor Max	Skor Rata-rata Kelompok		Ketercapaian Rata-rata (%)	
		A	B	A	B
1 Osilasi Pegas	30	3,4	10	11	33,3
2 Ayunan Matematis	26	4,6	16	17,7	61,5
3 Persamaan Getaran	60	3,7	16	16,2	26,7
4 Gelombang Stasioner	40	4,1	15,6	10,3	39
5 Energi Gelombang	40	4,5	28,3	11,3	70,7
6 Taraf Intensitas Bunyi	30	5,5	22,5	18,3	75
7 Efek Doppler	42	6,5	31,3	15,5	74,5

Dari Tabel 11 nampak kelompok ADM & ATS (B) lebih baik dari kelompok non ADM & ATS (A). Ada tren kenaikan ketercapaian, walaupun dalam kategori cukup namun sangat berarti bila dibandingkan dengan kelompok B. Hasil ketercapaian rata-rata pada kompetensi afektif dan psikomotor disajikan pada Tabel 12 dan Tabel 13.

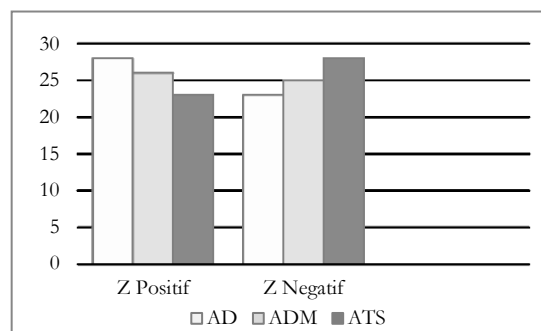
Bila nilai *z-score* antara AD, ADM dan ATS dibandingkan, ternyata hasil asesmen ADM lebih dekat ke asesmen dosen (AD) dari pada ATS. Hal ini bisa dipahami karena sifat kognitif adalah laten, yang hanya bisa dipahami oleh dirinya sendiri.

Tabel 12. Ketercapaian Kompetensi Afektif

Aspek yang Diases	Skor Max	Skor Rata-rata Kelompok		Ketercapaian Rata-rata (%)	
		A	B	A	B
Menerima (A1)	4	3,8	3,8	95	95
Menanggapi (A2)	4	3,4	3,6	85	90
Menilai (A3)	4	3,8	3,8	95	95
Mengorganisasi (A4)	4	3,8	3,8	95	95
Mengkarakterisasi (A5)	4	3,8	3,8	95	95

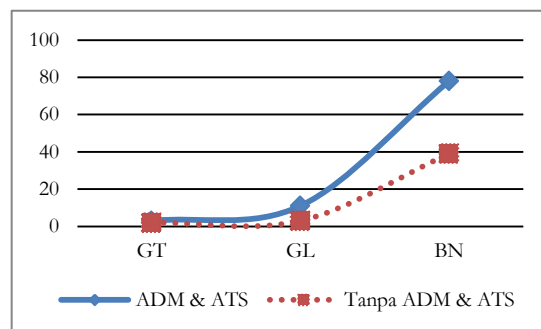
Tabel 13. Ketercapaian pada Kompetensi Psikomotor

Aspek yang Diases	Skor Max	Skor Rata-rata Kelompok		Ketercapaian Rata-rata (%)	
		A	B	A	B
Mendemonstrasikan (P1)	5	4,2	3,9	84	78
Menipulasi (P2)	5	4,3	4,1	86	82
Mengkomunikasi (P3)	5	3,9	4,2	78	84



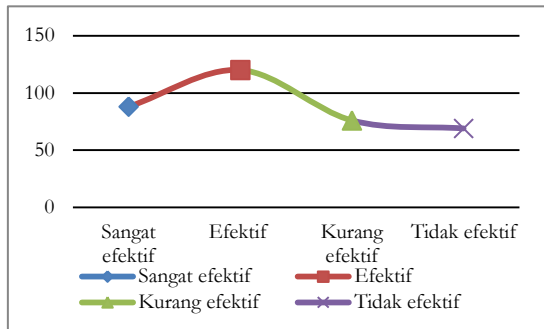
Gambar 1. Perbandingan Nilai *z-score* antara AD, ADM dan ATS

Hasil ketuntasan tiap materi pokok kelompok ADM & ATS lebih banyak tuntas dari pada kelompok non ADM & ATS, bila digambarkan dengan diagram garis disajikan pada Gambar 2.



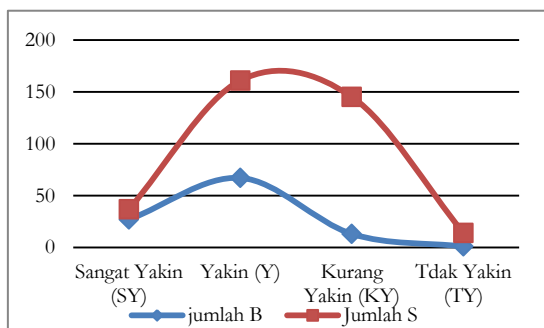
Gambar 2. Diagram Garis Perbedaan Ketuntasan tiap Materi Pokok

Bila hasil keefektifan ditinjau dari respon mahasiswa, hasilnya menunjukkan bahwa ADM & ATS efektif digunakan dalam pembelajaran fisika. Hal ini nampak dari tren kenaikan hasil belajar mahasiswa. Diagram garis keefektifan ADM & ATS ditinjau dari respon mahasiswa disajikan pada gambar 3.



Gambar 3. Keefektifan ADM & ATS ditinjau dari respon Mahasiswa

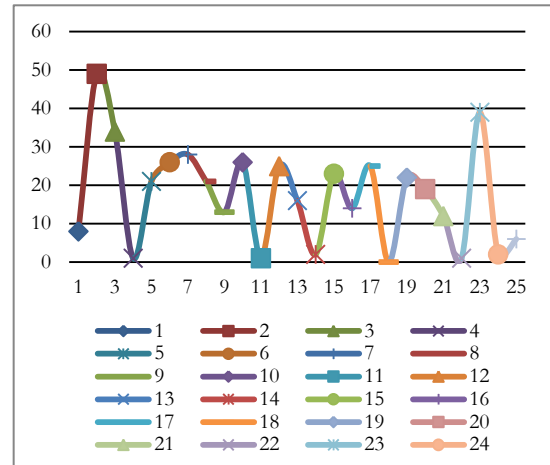
Hasil keterpakaian ADM dan ATS, berdasarkan estimasi validasi secara empiris, diperoleh informasi umpan balik terhadap mahasiswa maupun dosen sebagai pengampu mata kuliah. Informasi yang diperoleh hasil keterpakaian ADM & ATS antara lain: (1) analisis kesulitan mahasiswa pada asesmen kognitif terkait pada atribut, (2) butir-butir soal yang dianggap sulit oleh mahasiswa, dan (3) ketercapaian dan ketuntasan, berdasarkan hasil tersebut mahasiswa dapat mengetahui posisinya dan tujuan belajar yang belum tercapai. Hasil refleksi ADM tentang keyakinan atas jawaban benar seperti yang diperlihatkan pada Gambar 4.



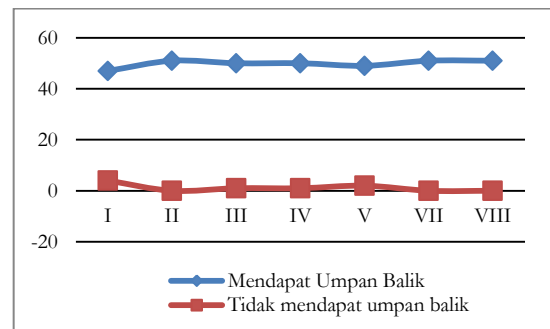
Gambar 4. Tingkat Keyakinan Atas Jawaban Benar

Pada Gambar 4 nampak bahwa banyak mahasiswa yang merasa yakin jawab-

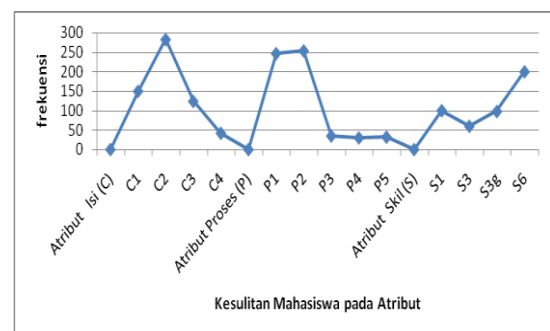
annya benar, namun ternyata salah berdasarkan asesmen dosen. Informasi ini bermanfaat sebagai refleksi diri, terhadap pekerjaannya. Dengan demikian, mahasiswa belajar mengembangkan kebiasaan untuk merefleksikan diri.



Gambar 5. Butir-butir Tes yang Dianggap Sulit



Gambar 6. Respon Mahasiswa terhadap ADM



Gambar 7. Diagram Garis Kesulitan Mahasiswa pada Atribut

Gambar 5 merupakan hasil releksasi ADM, butir sulit yang paling banyak dipilih oleh mahasiswa, yaitu butir nomor 2 dan 23.

Sebagian besar mahasiswa merasakan bahwa ADM & ATS memberikan umpan balik yang berdampak positif terhadap kenaikan hasil belajar. Pada Gambar 7 nampak bahwa mahasiswa masih banyak mengalami kesulitan pada atribut-atribut asesmen kognitif. Namun, yang paling menonjol, kesulitan pada atribut C2, P1, P2 dan S6.

Simpulan dan Saran

Simpulan

Hasil penelitian dan pembahasan terhadap ADM & ATS, yang dikembangkan dapat disimpulkan sebagai berikut.

1. Konstruksi ADM & ATS dibentuk oleh tiga dimensi yaitu kognitif afektif dan psikomotor. Hasil uji kecocokan model ADM & ATS menunjukkan bahwa model fit, karena nilai GFI sebesar 0,994, dan nilai SRMR sebesar 0,062. Dengan demikian, dapat disimpulkan bahwa konstruksi teoritik yang dikembangkan didukung oleh data empirik
2. Karakteristik ADM & ATS, proses asesmennya terintegrasi dalam proses pembelajaran dan bersifat *on going*, sebagai pelengkap asesmen formatif. Instrumen ADM & ATS hasilnya telah memenuhi syarat sebagai instrumen yang baik, yaitu valid secara konten, tingkat kesulitan kategori sedang, daya pembeda kategorinya sangat tinggi, reliabel serta valid secara konstruk.
3. Hasil analisis deskriptif ketercapaian dan ketuntasan menunjukkan bahwa kelompok ADM & ATS ketercapaian dan ketuntasan cukup baik. Profil ketuntasan individu pada kelompok ADM & ATS lebih baik dari pada kelompok non ADM & ATS. Oleh karena itu dapat disimpulkan bahwa ADM & ATS cukup efektif bila digunakan sebagai pelengkap asesmen formatif, khususnya dalam pembelajaran fisika.
4. Hasil analisis deskriptif dan refleksi ADM dan ATS menunjukkan bahwa ketercapaian ADM dan ATS sangat bermanfaat. Informasi hasil belajar yang diperoleh secara langsung setelah melakukan tes, sa-

ngat mendukung karena dapat digunakan untuk memantau proses pembelajaran, kemajuan belajar serta perbaikan hasil belajar mahasiswa secara kontinyu.

Saran Pemanfaatan dan Diseminasi

1. Model ADM & ATS beserta instrumen pendukungnya telah dikembangkan, sehingga dapat digunakan sebagai pelengkap asesmen formatif dalam pembelajaran fisika pada materi gelombang dan bunyi. Penggunaan ADM & ATS akan membantu mahasiswa belajar mengenali kekuatan dan kelemahannya di dalam memahami materi fisika. Selain itu, melalui ADM & ATS dapat mendorong mahasiswa menjadi lebih reflektif dan memiliki kritik diri secara berkelanjutan. Meningkatkan rasa tanggung jawab mahasiswa dan secara aktif terlibat dalam proses pembelajaran. Model asesmen ini sangat diperlukan dalam pembelajaran yang berpusat pada mahasiswa (*student-centered*).
2. Model ADM & ATS yang digunakan sebagai pelengkap asesmen formatif sudah diujicobakan dalam pembelajaran fisika di Prodi Pendidikan Fisika FKIP Unpar. Namun, belum semua mahasiswa dan dosen terlibat dalam penelitian ini, sehingga proses diseminasi dapat dilakukan pada semua civitas akademika di Program Studi Pendidikan Fisika, maupun pada guru-guru fisika se-Provinsi Kalimantan Tengah.

Keterbatasan Penelitian

Uji kelayakan model ADM & ATS melibatkan jumlah subjek yang terbatas, tidak semua mahasiswa mempunyai kemampuan dan kemampuan yang sama, sehingga akan berpengaruh terhadap perolehan skor mahasiswa. Produk pengembangan instrumen ADM & ATS kiranya masih memerlukan penyempurnaan melalui analisis dan revisi lebih lanjut. Mengingat bahwa produk ADM & ATS baru sampai pada tahap validasi belum dilakukan analisis dan revisi terhadap hasil empiris.

Ukuran sampel yang masih belum memadai, hal ini nampak dari beberapa hasil analisis yang memperlihatkan nilai, *Observed power* yang masih rendah. Selain itu instrumen ADM & ATS baru difokuskan pada domain kognitif, belum dikembangkan untuk domain afektif dan psikomotor.

Daftar Pustaka

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142.
- Amien, M. (1987). *Mengajarkan ilmu pengetahuan alam (IPA) dengan menggunakan metode "discovery dan inquiry"*. Jakarta: Dirjen Dikti.
- Anderson, L.W. & Krathwohl, R. (Ed). (2001). *A taxonomy for learning teaching and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). New York: Longman.
- Arends, R.I. (2007). *Learning to teach II* (7th ed.). New York: McGraw Hill Companies.
- Assessment Reform Group. (1999). *Assessment for Learning: Beyond the black box*. University of Cambridge School of Education. Diakses tanggal 30 Mei 2010 dari <http://www.assessment-reformgroup.org.uk>.
- Azwar, S. (2012). *Penyusunan skala psikologi*. Yogyakarta: Pustaka Pelajar.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*. 80, 2, 139-148. Diambil 30 Juni 2011 dari http://www.academicleadership.org/articles/9/1_full.html.
- Black, P. et al. (2004). Working in side the black box: Assessment for learning in the classroom. *Phi Delta Kappan*. 86, 1, 8-14.
- Cennamo, K. & Kalk, D. (2005). *Real world instructional design*. Australia: Thomson Learning.
- Croker, L. & Algina, J. (1986). *Intoduction to classical and modern test theory*. Newyork, NY: CBS College Publishing.
- Depdiknas. (2003). *Undang-undang Republic Indonesia Nomor 20 pasal 58 ayat (1) Tahun 2003 tentang Sistem Pendidikan Nasional*.
- Ebel, R.L. & Frisbie, D.A. (1986). *Essential of educational measurement*. Englewood Cliffs, NJ: Prentice Hall, Inc.
- Ghazali, I., & Latan, H. (2012). *Partial least squares konsep, teknik dan aplikasi smartPLS 2.0M3*. Semarang: Badan Penerbit Undip.
- Gronlund, N. E. & Linn, R. L. (1990). *Measurement and evaluation in teaching*. (6th ed.). New York: Publishing Company.
- Hair, J.F. Hult, G.T. Ringle, C.H. & Sarstedt, H. (2014). *A primer on partial least square structural equation modelling (PLS – SEM)*. Los Angeles: Sage Publication, Inc.
- Haris, L. (2007). *Employing formative assessment in the classroom*. Diakses tanggal 28 september 2010 dari <http://imp.sagepub.com/content/10/3/249.refs.html>.
- Hayat, B. (2010). *Prinsip-prinsip dan strategi penilaian kelas*. Jakarta: Puspendik
- Johnson, D.W., & Johnson, R. (2002). *Meaningful assessment: A manageable and cooperate process*. Boston: allyn Bacon.
- Keaten, J.A. et al. (1993). *A field investigation of peer-assessment as part of the student group grading process*. EDRS ED 361 753.
- Kumaidi, (2001a). *Pengujian sebagai bagian peningkatan kualitas pembelajaran*. Pidato pengukuhan guru besar madya dalam ilmu evaluasi pada Fakultas Teknik Universitas Negeri Padang.
- Kusairi, S. (2011). *Analisis asesmen formatif fisika SMA berbantuan computer*. Disertasi tidak diterbitkan. Universitas Negeri Yogyakarta

- Linn, Robert L. (1989). *Educational measurement (3rd ed.)*. New York: Macmillan Publishing Company.
- Mardapi, D. (2012). *Pengukuran, penilaian & evaluasi pendidikan*. Yogyakarta: Nuha Medika.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Miller, M. D. Linn, R. L. & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. (10th ed). Upper Saddle River, NJ: Pearson Education, Inc.
- Nitko, A.J. & Brookhart, S.M. (1989). Designing tests that are integrated with instruction. Dalam Linn, R. *Educational assessment of students (5th ed.)*, p. 447. Columbus, Ohio: Pearson Merrill Prentice Hall.
- Noonan, B. & Duncan, C.R. (2005). Peer and self-assessment in high school. *Practical Assessment, Research & Evaluation*, 4, 17, 1-10.
- Tatsuoka, K. K. (2009). *Cognitive assessment*. New York: Routledge.
- Trowbridge, L. W. & Bybee, R. W. (1986). *Becoming a secondary school science teacher (4th ed)*. Columbus: Charles E Merril Publishing Company.