

Seminar Nasional Hasil Pengabdian kepada Masyarakat
26 November 2022, Hal. 1415-1422
e-ISSN: 2686-2964

Pelatihan Analisis Karakteristik Tes bagi Guru SMA

Kusumarasyati¹, Kusumadyahdewi², Hakkun Elmunsyah³

Universitas Negeri Surabaya, Kampus Lidah Wetan, Surabaya¹
UIN Maulana Malik Ibrahim, Jl. Gajayana 50, Malang²
Universitas Negeri Malang, Jl Semarang 5, Malang³
kusumarasyati@unesa.ac.id

ABSTRAK

Tes merupakan salah satu alat yang paling sering digunakan dalam mengukur kemampuan pembelajar di pendidikan menengah. Untuk mendapatkan skor yang betul-betul mencerminkan kemampuan pembelajar, tes harus memenuhi beberapa karakteristik, yaitu sah, dapat diandalkan, tidak terlalu sulit maupun terlalu mudah, dan dapat membedakan pembelajar yang mampu dan kurang mampu. Kegiatan pengabdian kepada masyarakat ini bertujuan untuk membekali para guru SMA agar memiliki pengetahuan dan kemampuan dalam menyusun tes yang baik. Kegiatan ini dilaksanakan di SMA Vita Surabaya dalam tiga tahap. Pada tahap pertama, para guru mempelajari konsep-konsep yang berkaitan dengan karakteristik tes, yaitu kepraktisan, validitas (kesahihan), reliabilitas (keterandalan), tingkat kesulitan butir soal dan daya pembeda butir soal. Pada tahap berikutnya, mereka menganalisis karakteristik tes yang telah mereka gunakan di kelas pada semester sebelumnya. Tahap terakhir berupa evaluasi kegiatan yang dilakukan untuk menjaring pendapat para peserta mengenai pelatihan. Hasil kuesioner dan wawancara menunjukkan respon positif para guru terhadap kegiatan ini karena mereka memandang kegiatan ini sangat bermanfaat untuk evaluasi pembelajaran di kelas. Dokumen yang mereka susun selama pendampingan menunjukkan bahwa pengetahuan mereka mengenai karakteristik tes sudah baik, dan kemampuan mereka untuk menganalisis karakteristik tes juga meningkat.

Kata kunci: pengabdian kepada masyarakat, evaluasi pembelajaran, asesmen, karakteristik tes, pendidikan menengah

ABSTRACT

Tests are one of the tools frequently used to estimate learners' competences in the secondary education. To obtain scores that reflect these competences accurately, teachers have to ensure that the tests possess the following characteristics: high validity, high reliability, suitable difficulty level and appropriate discriminatory power. The current community service activity aims to enable the teachers of a senior high school to analyze the characteristics of test which could measure the learners' competences more accurately. It was held in SMA Vita Surabaya in three stages. In the first stage, the teachers learned the basic concepts of test characteristics, namely, validity, reliability, difficulty level and discriminatory power. In the next stage, they analyzed the characteristic of the tests they used in the previous semester, followed by evaluation in the final stage. The results of the questionnaire and interviews showed favorable response from the teachers, who thought this activity was very beneficial for evaluating the

learning process in their classes. They gained knowledge about the characteristics of a good test and improved their ability to analyze the characteristics of the tests they had developed.

Keywords : *community service, educational evaluation, assessment, test characteristics, secondary education*

PENDAHULUAN

Proses belajar mengajar tidak dapat terlepas dari asesmen, yaitu proses mengumpulkan, menggabungkan dan menginterpretasi informasi untuk mengambil keputusan (Russell & Airasian, 2012). Asesmen memiliki peran penting pembelajaran karena dari proses tersebut guru dapat memantau kemajuan belajar siswa dan mengetahui apakah tujuan pembelajaran telah tercapai (Miller et al., 2009). Ada beberapa instrumen yang digunakan dalam asesmen, seperti tes, observasi, angket, portfolio, dan sebagainya (Mehrens & Lehmann, 1991; Miller et al., 2009). Namun makalah ini hanya membahas instrumen berupa tes karena asesmen di sekolah pada umumnya lebih sering menggunakan tes daripada instrumen lain. Tes merupakan prosedur yang formal dan sistematis untuk mendapatkan informasi mengenai pencapaian siswa dan kemampuan kognitif lain yang dimiliki oleh siswa (Russell & Airasian, 2012). Untuk mengembangkan tes, guru harus mengacu pada tujuan pembelajaran seperti yang disebutkan di atas. Tujuan pembelajaran tersebut berisi kemampuan siswa yang hendak dicapai, dan menentukan materi pembelajaran serta metode pengajaran yang tepat. Pada akhir proses pembelajaran, guru menyusun butir-butir tes untuk mengukur pencapaian siswa dalam menguasai materi yang telah diajarkan. Tes seperti ini disebut tes buatan guru (*teacher-made test*). Tidak seperti tes terstandar (*standardized test*), tes buatan guru pada umumnya belum diujicoba sehingga ada kemungkinan tes tersebut memiliki beberapa kekurangan, seperti butir soal objektif yang ambigu (memiliki jawaban lebih dari satu), kalimat yang digunakan terlalu panjang, fokus tes yang kurang jelas dan penggunaan format tes yang kurang tepat (Mehrens and Lehmann, 1991). Apabila kekurangan tersebut tidak diatasi, tes akan memberikan hasil yang bias dan tidak dapat disebut sebagai alat yang baik untuk mengukur kemampuan siswa. Karena itu harus dipastikan tes tersebut memiliki beberapa karakteristik penting dan layak digunakan sebagai alat ukur kemampuan anak didik di kelas.

Tes yang baik memiliki karakteristik sebagai berikut: 1) praktis, 2) sah, 3) dapat diandalkan, 4) tingkat kesulitan yang baik, dan 5) daya pembeda yang baik. Tes yang praktis artinya tes tersebut tidak perlu biaya yang banyak untuk dikembangkan dan digunakan di kelas, mudah untuk dibuat dan dikerjakan, serta tidak rumit untuk dinilai. Tiap butir soal tes juga harus dibandingkan dengan silabus untuk memastikan apakah butir soal tersebut betul-betul menguji apa yang telah diajarkan guru. Apabila tiap butir soal sudah sesuai dengan materi yang terdapat di silabus dan pembelajaran di kelas, tes tersebut dianggap memiliki tingkat kesahihan yang tinggi. Selain sah, tes harus dapat diandalkan sehingga memberikan hasil yang konsisten meskipun digunakan berkali-kali. Agar tes tersebut sah dan dapat diandalkan, tingkat kesulitan dan daya pembeda tiap butir tes harus diukur. Butir tes dengan tingkat kesulitan yang baik tidak terlalu sulit ataupun terlalu mudah untuk dikerjakan oleh siswa. Guru juga harus memastikan bahwa daya pembeda tiap butir tes juga baik, sehingga dapat mengidentifikasi mana siswa yang mampu dan siswa yang kurang mampu. Setelah tes diujicobakan dan memenuhi karakteristik di atas, butir-butir soalnya disimpan dalam bank soal untuk digunakan lagi ke depannya.

Pada umumnya para guru mendapatkan pengetahuan mengenai cara mengembangkan tes yang baik termasuk lima karakteristik tersebut pada saat mereka menempuh pendidikan guru prajabatan di universitas, yaitu di fakultas keguruan atau program studi pendidikan yang terkait

dengan bidang ilmu mereka. Dalam pendidikan guru prajabatan juga ada praktek untuk menerapkan teori asesmen yang telah mereka pelajari di praktikum mengajar. Namun ada beberapa guru di SMA Vita Surabaya yang belum memiliki pengetahuan tersebut ataupun menempuh praktek karena mereka mendalami bidang ilmu di program studi non-kependidikan di universitas sebelum menjadi guru. Selama ini mereka melakukan asesmen dengan menggunakan tes, tetapi belum pernah menentukan kesahihan dan keterandalan tes yang telah mereka kembangkan. Mereka juga belum pernah mengukur tingkat kesulitan dan daya pembeda butir tes. Untuk meningkatkan pemahaman dan kemampuan mereka terkait karakteristik tes, SMA Vita mengadakan pelatihan bagi para guru. Pelatihan ini bertujuan untuk: (1) membekali para guru dengan pengetahuan mengenai konsep dasar tes dan karakteristiknya, dan (2) mengembangkan kemampuan mereka dalam menyusun tes yang praktis, sah, dapat diandalkan dan memiliki tingkat kesulitan dan daya pembeda yang baik. Setelah pelatihan ini, diharapkan para guru dapat membuat tes yang mengukur kemampuan siswa dengan akurat dan tidak bias.

METODE

Pelatihan ini diselenggarakan oleh SMA Vita, Surabaya selama dua hari dengan tim PKM sebagai narasumber. Peserta terdiri dari sepuluh guru SMA Vita yang mengampu berbagai mata pelajaran. Meskipun beberapa di antara mereka dulunya menempuh pendidikan prajabatan sebagai guru, mereka tetap berpartisipasi dalam pelatihan ini untuk pendalaman dan pengayaan sebagai bagian dari pengembangan profesi dalam jabatan.

Kegiatan ini dilaksanakan dalam tiga tahap, yaitu pelatihan, pendampingan dan evaluasi. Pada tahap *pelatihan* di hari pertama, tim PKM menjelaskan konsep-konsep dasar asesmen dan tes, termasuk berbagai karakteristik tes, yaitu kepraktisan, validitas, reliabilitas, tingkat kesulitan dan daya pembeda butir tes. Lalu tim PKM menunjukkan cara mengembangkan tes yang praktis sehingga penggunaannya untuk mengases anak didik lebih efisien. Untuk menentukan validitas tes, tim PKM mendemonstrasikan cara membuat tabel spesifikasi yang menunjukkan kesesuaian antara materi di silabus dan butir tes. Reliabilitas tes dapat diuji dengan menggunakan berbagai metode, tetapi tim PKM memilih metode yang paling sederhana, yaitu konsistensi internal (*internal consistency*). Rumus yang digunakan untuk mengukur konsistensi internal adalah K-R21, dan penghitungannya menggunakan kalkulator K-R21 online. Yang terakhir, analisis butir tes diperagakan dengan menggunakan Excel®. Dari satu halaman data di sana, tingkat kesulitan dan daya pembeda butir soal dapat diukur.

Pada tahap *pendampingan* di hari berikutnya, peserta pelatihan menganalisis tes yang pernah mereka gunakan untuk mengases anak didik mereka di kelas. Mereka bekerja dalam tiga kelompok, dan tiap kelompok memilih satu tes objektif untuk dianalisis dari aspek kepraktisan, validitas, reliabilitas, tingkat kesulitan dan daya pembeda butir tes. Setelah itu tiap kelompok menyajikan hasil analisisnya di hadapan peserta lain sehingga mereka dapat saling menimba ilmu dari rekannya.

Kegiatan ini diakhiri dengan tahap *evaluasi*, di mana para guru memberikan pendapatnya mengenai kegiatan ini. Instrumen yang digunakan untuk menjaring data tersebut ada dua, yaitu angket terstruktur dan wawancara semi-terstruktur. Peserta mengisi angket secara online, lalu berdasarkan jawaban mereka tim PKM mewawancarai mereka untuk mendapatkan data yang lebih mendalam mengenai beberapa aspek pelatihan dan pendampingan dari perspektif mereka.

HASIL, PEMBAHASAN, DAN DAMPAK

Setelah mengikuti pelatihan dan pendampingan, hasil yang didapatkan adalah meningkatnya pengetahuan para guru SMA Vita, Surabaya mengenai ciri-ciri tes yang baik. Bahkan bagi guru yang belum menempuh pendidikan prajabatan di fakultas keguruan atau

prodi pendidikan, materi pelatihan ini merupakan ilmu yang baru bagi mereka. Selain itu, kemampuan mereka dalam menguji kepraktisan, validitas, reliabilitas, tingkat kesulitan dan daya pembeda butir tes juga semakin meningkat. Hasil ini akan dijabarkan lebih lanjut sebagai berikut.

Kepraktisan Tes

Suatu tes bisa dianggap praktis apabila tes tersebut mudah dan murah untuk dikembangkan oleh guru dan dibagikan kepada siswa (Brown, 2004). Dalam pelatihan ini, tiap kelompok memilih satu tes yang pernah mereka gunakan di kelas sebelumnya, sehingga ada tiga tes yang dianalisis. Untuk mengetahui apakah tes yang telah mereka pilih memiliki karakteristik praktis atau tidak, tiap kelompok membuat *checklist* atau daftar cek yang berisi sejumlah kriteria (Tabel 1).

Tabel 1. Daftar cek kepraktisan tes

No	Kriteria	Y/T	Catatan
1	Prosedur pembuatan tes tidak terlalu rumit.		
2	Kunci jawaban sudah disiapkan sebelum tes digunakan.		
3	Prosedur penilaian sudah disiapkan sebelum tes digunakan.		
4	Pembuatan tes tidak makan waktu terlalu banyak.		
5	Pembuatan tes tidak perlu biaya terlalu banyak.		
6	Siswa dapat menyelesaikan tes dalam alokasi waktu yang telah ditentukan.		
7	Siswa memandang petunjuk pengerjaan tes sudah jelas.		
8	Prosedur melaporkan hasil tes sudah disiapkan sebelum tes digunakan.		

Kriteria yang mereka masukkan dalam daftar meliputi pengembangan dan penggunaan tes. Untuk mengembangkan tes, menurut mereka perlu dipertimbangkan mengenai prosedur dan kelengkapannya. Pada saat tes dibagikan ke siswa, kriteria yang perlu diperhatikan adalah alokasi waktu dan petunjuk pengerjaan tes. Apabila tes memenuhi suatu kriteria, mereka menulis 'Y' (ya) pada kolom di sebelah kanannya. Sedangkan kode 'T' (tidak) dicantumkan di sana jika kriteria tidak terpenuhi. Baik memenuhi kriteria ataupun tidak, para guru dapat menuliskan catatan untuk tiap kriteria. Misalnya, guru menganggap tes yang telah dikembangkannya memenuhi kriteria karena siswa dapat menyelesaikan tes tersebut dalam waktu yang telah ditentukan, maka dibuat catatan mengenai jumlah soal dan alokasi waktu di kolom paling kanan. Hal ini bermanfaat pada saat mereka mengembangkan tes yang serupa di masa yang akan datang. Mereka jadi lebih mudah menentukan alokasi waktu yang tepat untuk tes tertentu. Sebaliknya, kriteria yang tidak terpenuhi untuk suatu tes juga perlu disebutkan penyebabnya dalam catatan, sehingga mereka dapat menghindari masalah yang sama nantinya.

Validitas Tes

Karakteristik tes paling penting adalah validitas atau kesahihan. Meskipun suatu tes memiliki karakteristik lain seperti praktis dan dapat diandalkan, tes ini tidak akan dapat digunakan apabila validitasnya sangat rendah. Suatu tes dianggap memiliki validitas yang tinggi apabila tes tersebut mengukur kemampuan siswa yang memang seharusnya diukur (Miller et al., 2009). Ada empat metode untuk mengetahui validitas tes, namun dalam pelatihan ini metode validitas isi (*content validity*) yang digunakan karena paling sesuai untuk tes buatan guru. Kesahihan isi tes dapat diketahui dengan membuat tabel spesifikasi (Tabel 2).

Tabel 2. Tabel spesifikasi untuk mapel Bahasa Inggris

No	Materi	Butir Soal																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Greetings			√																	
2	Congratulations						√	√													
3	Getting information				√	√															
4	Descriptive text											√	√	√							
5	Announcement										√	√									
6	Past experience								√	√											
7	Recount text														√	√	√				
8	Narrative text																	√	√	√	
9	Song lyrics	√	√																		

Tabel spesifikasi ini berisi semua materi pembelajaran yang tercantum di silabus mapel Bahasa Inggris kelas X dan telah diajarkan di kelas pada semester sebelumnya. Untuk mengetahui validitas isi tes akhir semester yang dibuat oleh guru, tiap butir soal dicocokkan dengan materi. Misalnya soal nomor 1 dan 2 mengenai lirik lagu, soal nomor 3 mengenai ungkapan sapaan, dan seterusnya. Karena semua topik sudah direpresentasikan dalam soal-soal tes, maka tes tersebut bisa dianggap memiliki validitas yang tinggi.

Reliabilitas Tes

Tes yang baik menghasilkan skor yang konsisten meskipun diberikan kepada siswa berulang kali. Karakteristik ini disebut reliabilitas (Miller et al., 2009), yaitu seberapa konsisten tes tersebut mengukur kemampuan siswa. Reliabilitas dapat diukur dengan menggunakan berbagai metode, dan metode yang paling tepat untuk tes buatan guru adalah konsistensi internal karena proses penghitungannya sederhana dan hasilnya cukup akurat. Dalam pelatihan ini para guru dianjurkan untuk menghitung koefisien reliabilitas dengan rumus K-R21. Untuk memudahkan mereka lebih lanjut, pemakaian kalkulator K-R21 secara online (Walker, 2007) disarankan dalam penghitungan koefisien reliabilitas (Gambar 1).

Gambar 1. Kalkulator K-R21 online

Untuk mengoperasikan kalkulator ini, diperlukan tiga jenis data saja, yaitu rerata (M), simpangan baku (SD) dan jumlah soal dalam tes (k). Karena itu para guru terlebih dahulu menghitung rerata dan simpangan baku dari skor tes akhir semester di Excel[®]. Setelah memasukkan data dan menekan tombol 'Calculate', koefisien reliabilitas muncul di laman tersebut dan dapat diinterpretasi oleh para guru.

Tingkat Kesulitan dan Daya Pembeda Butir Soal Tes

Tiga karakteristik yang telah dibahas di atas berlaku untuk tes secara keseluruhan, sedangkan dua karakteristik terakhir mengacu pada butir soal dalam suatu tes. Soal tes yang baik hendaknya tidak terlalu sulit ataupun terlalu mudah. Apabila soal terlalu mudah, semua siswa dapat menjawabnya dengan benar sehingga soal tersebut tidak dapat membedakan siswa yang pandai dan kurang pandai. Begitu pula jika soal terlalu sulit, tidak ada atau sangat sedikit siswa yang memberi jawaban benar sehingga sulit pula membedakan kedua kelompok siswa tersebut (Heaton, 1990; Hughes, 2002). Dengan demikian, tingkat kesulitan dan daya pembeda soal tes sangat berhubungan erat meski konsepnya berbeda. Keduanya dikategorikan sebagai analisis butir soal (*item analysis*).

Untuk mengukur keduanya, data yang diperlukan sama, yaitu jumlah siswa pandai (*high students*) yang menjawab suatu soal dengan benar dan jumlah siswa kurang pandai (*low students*) yang menjawab soal tersebut dengan benar. Perbedaan hanya terletak pada rumusnya, karena itu para guru menganalisis keduanya dalam satu tabel di Excel[®] (Gambar 2).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Siswa																Tingkat Kesulitan	Daya Pembeda	
2	Soal	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			16
3	1	1	1	1	1	0	1	0	1	1	0	0	1	1	1	0	0	0.63	0.13
4	2	1	0	1	1	1	1	1	1	1	0	1	0	0	1	0	0	0.63	0.25
5	3	1	1	1	1	0	0	0	1	0	1	0	0	1	1	1	0	0.56	0.06
6	4	1	0	0	1	1	1	1	1	1	1	0	1	1	0	0	1	0.69	0.06
7	5	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0.88	0.13
8	6	1	1	1	1	1	1	0	0	0	0	1	0	0	1	0	0	0.50	0.25
9	7	0	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	0.63	0.25
10	8	1	1	0	1	1	1	0	1	0	1	0	0	1	1	0	0	0.56	0.19
11	9	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	1	0.56	0.31
12	10	1	1	1	1	0	1	1	1	0	1	1	0	1	0	1	0	0.69	0.19
13	11	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0.75	0.25
14	12	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	0	0.69	0.31
15	13	1	0	1	1	1	0	1	1	1	0	1	0	1	0	0	0	0.56	0.19
16	14	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	1	0.63	0.13
17	15	1	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0.50	0.38
18	16	1	1	1	1	0	1	1	0	0	0	0	1	0	0	0	1	0.50	0.25
19	17	1	1	1	1	1	1	1	0	0	0	1	0	0	1	0	0	0.56	0.31

Gambar 2. Penghitungan Tingkat Kesulitan dan Daya Pembeda Soal Tes

Tabel tersebut berisi nomor siswa, skor siswa untuk tiap soal, serta angka yang menunjukkan tingkat kesulitan dan daya pembeda soal. Karena tes yang dianalisis adalah tes objektif, skornya 1 untuk tiap jawaban yang benar dan 0 untuk tiap jawaban yang salah. Tingkat kesulitan dan daya pembeda soal dihitung dengan rumus di kolom paling kanan. Setelah didapatkan hasilnya, para guru memilah soal yang sudah memenuhi kriteria tingkat kesulitan (antara 0,2 hingga 0,8) dan daya pembeda (lebih dari 0,2) dengan soal yang masih perlu diperbaiki. Soal yang belum memenuhi kriteria ditindaklanjuti dengan revisi setelah kegiatan pendampingan selesai.

Evaluasi

Berdasarkan hasil angket yang diisi oleh para guru, pendapat mereka mengenai kegiatan ini sangat positif. Tingkat kepuasan mereka terhadap berbagai aspek pelatihan dan pendampingan berkisar antara 76% hingga 82% (Tabel 3).

Tabel 3. Tingkat Kepuasan Peserta Pelatihan

Aspek Penilaian	Tingkat kepuasan (%)
Isi Materi	83
Penyampaian Materi	81
Praktek	78
Fasilitas Praktek	77

Dari hasil wawancara dengan beberapa guru diketahui bahwa materi pelatihan ini sangat bermanfaat, utamanya bagi para guru yang dulunya menempuh pendidikan tinggi di prodi non-kependidikan. Bagi mereka, analisis karakteristik tes merupakan hal yang baru dan penting untuk dipelajari serta diterapkan dalam asesmen pembelajaran di kelas. Selama ini mereka menyusun tes tanpa memiliki dasar ilmu yang kuat mengenai asesmen sebagai bagian dari evaluasi pembelajaran. Setelah mengikuti pelatihan ini mereka paham bagaimana mengembangkan tes yang baik karena sudah mengetahui berbagai karakteristik tes yang harus diperhatikan. Hal ini sejalan dengan kegiatan PKM berupa pelatihan untuk menganalisis validitas, reliabilitas, tingkat kesulitan dan daya pembeda butir soal yang dilaksanakan oleh Fauzan dkk. (2021). Selain itu, penggunaan teknologi berupa *software* dalam proses penghitungan juga sangat membantu tugas guru dalam analisis karakteristik tes karena penghitungan jadi lebih cepat dan akurat (Nuraeni dkk., 2021). Para guru SMA Vita berharap pelatihan ini dapat berkelanjutan dengan membahas jenis-jenis asesmen yang lain seperti tes essay dan portfolio.

SIMPULAN

Pelatihan ini dilaksanakan untuk membekali para guru SMA Vita dengan pengetahuan mengenai karakteristik tes dan meningkatkan kemampuan mereka dalam menganalisis tes untuk mengetahui apakah tes yang mereka susun sudah memenuhi kriteria tes yang baik. Tim PKM mengajarkan konsep dasar karakteristik tes, lalu mendampingi para guru untuk praktek menganalisis tes dalam hal kepraktisan, validitas, reliabilitas, tingkat kesulitan dan daya pembeda. Untuk mengetahui apakah tes mereka bisa dikategorikan praktis dan valid, mereka membuat tabel yang digunakan untuk menganalisis tes. Sedangkan reliabilitas, tingkat kesulitan dan daya pembeda dihitung dengan menggunakan aplikasi online maupun offline sehingga akurat dan menghemat waktu. Setelah mengikuti pelatihan, mereka mendapatkan pengetahuan yang mendalam mengenai karakteristik tes yang baik. Selain itu mereka juga memiliki kemampuan yang baik untuk menganalisis karakteristik tes sehingga tes yang mereka kembangkan dapat mengukur kemampuan siswa mereka secara akurat. Hasil evaluasi kegiatan ini menunjukkan bahwa para guru antusias untuk mempelajari lebih lanjut berbagai aspek asesmen dan berharap pelatihan ini masih berkelanjutan untuk memperkaya pengetahuan dan kemampuan mereka dalam mengases anak didik mereka.

UCAPAN TERIMAKASIH

Tim PKM mengucapkan banyak terima kasih kepada SMA Vita, Surabaya selaku penyelenggara pelatihan ini, dan juga kepada para guru SMA Vita yang telah berpartisipasi sebagai peserta.

DAFTAR PUSTAKA

- Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. Longman.
- Fauzan, A., Hariyanto, Rispawati dan Tripayana, I.N.A. (2021). Pendidikan Dan Latihan Menghitung Validitas, Reliabilitas, Tingkat Kesukaran, Daya Beda Soal Bagi Guru PPKn SMAN Se-Kota Mataram. *Jurnal Pengabdian Magister Pendidikan IPA*, 4(4):35-41.
- Heaton, J. B. (1990). *Writing English language test*. Longman.
- Hughes, A. (2002). *Testing for language teachers*. Cambridge University Press.
- Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. Wadsworth/Thomson Learning.

- Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. Pearson Education, Inc.
- Nuraeni, Z., Simarmata, R.H., Sukmaningthias, N. dan Sari, N. (2021). Pelatihan Software SPSS untuk Menghitung Validitas, Reliabilitas, dan Analisis Butir Soal bagi Mahasiswa Calon Guru di Palembang. *Jurnal Anugerah*, 3(1):15-23.
- Russell, M., & Airasian, P. (2012). *Classroom Assessment: Concepts and Applications*. McGraw-Hill.
- Walker, D. A. (2007). *Walker's Calculators: Kuder-Richardson 21*. <https://www.cedu.niu.edu/~walker/calculators/kr.asp>.