

# Pendampingan Pembuatan Soal Ujian Kredibel bagi Guru Madrasah di Kabupaten Soppeng

Rukli<sup>1</sup>, Munirah<sup>2</sup>

<sup>1</sup>Prodi Pendidikan Matematika, Universitas Muhammadiyah Makassar

<sup>2</sup>Prodi Pendidikan Bahasa Indonesia, Universitas Muhammadiyah Makassar

**Abstract.** The purpose of this activity is how to empower teachers to make credible questions through mentoring method. The results of the identification of the problem in making credible questions at Islamic boarding school in Soppeng shows that this mentoring is needed for the teachers to increase their knowledge in understanding credible questions through, to improve their skills in making questions based on the indicators, and to improve their skills in analysing questions qualitatively. The result of this mentoring showed that the implementation of this mentoring was at Madrasah Tsanawiah Yasrib Islamic Boarding School in Watansoppeng, South Sulawesi. It is one of favourite schools in Soppeng for elementary graduates. The questions used by teachers have not been arranged well where most of the questions are categorized as very bad and do not reflect the activities of productive students. This mentoring helps them to make good questions, considering the questions theory, good indicators, especially the main problem (stem), and the competency in analysing the problems qualitatively. Long-term planning of mentoring activities where the teachers have their own standard questions for school exams and other exams. Generally, teachers make 20 sample questions for each subject to be developed further as an incubator question and question bank with good criteria.

**Keywords:** assistance in making questions, credible, madrasah

## I. PENDAHULUAN

### A. Analisis Situasi

Undang-undang Nomor 14 Tahun 2005 Pasal 1 menegaskan bahwa guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan anak usia dini jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah. Oleh karena itu, mengevaluasi kemampuan peserta didik baik sebelum, saat, dan setelah proses pembelajaran merupakan tugas guru. Agar proses pengambilan keputusan akurat dan terpercaya, maka alat evaluasi tersebut menggunakan alat ukur valid, reliabel, objektif, dan adil. Disamping persyaratan tersebut, sesuai kebutuhan perkembangan teknologi dan informasi maka peta domain kognitif berpikir tingkat tinggi menjadi prioritas utama di sekolah sebagai ciri guru abad 21.

Hasil observasi pada Madrasah Tsanawiyah Yasrib Lapajung sejauh kurang lebih 180 km dari Unismuh Makassar tanggal 20-22 April 2017 menunjukkan sebagian besar guru sudah memiliki

beberapa tes tipe soal pilihan ganda dan soal uraian. Jumlah guru sebanyak 35 orang terdiri dari 9 orang laki-laki dan 26 orang perempuan. Guru berstatus PNS sebanyak 5 orang, selebihnya guru berstatus honorer dan kontrak. Sekolah tersebut sudah memiliki beberapa tes baik tipe soal pilihan ganda maupun uraian namun umumnya menggunakan soal tipe pilihan ganda. Soal-soal tersebut dipakai saat ujian di kelas baik ujian tengah semester maupun ujian akhir (kenaikan kelas) selama beberapa tahun, bahkan sebagian soal tersebut diguna-kan saat ujian akhir sekolah.

Pendalaman lebih lanjut tentang karakteristik soal tersebut menggunakan tabel analisis kualitatif dengan melibatkan guru dan kepala sekolah menunjukkan bahwa soal buatan guru masih sangat sederhana. Guru memiliki keterbatasan untuk membuat dan meng-analisis soal sesuai prinsip, mekanisme, dan prosedur penilaian sehingga tidak sedikit soal tersebut berasal dari contoh soal yang sudah ada di buku teks. Demikian halnya, soal contoh pembahasan materi pembelajaran saat guru mengajar kadangkala guru

menggunakan lagi saat ujian semester. Jika demikian halnya, kemampuan siswa mengerjakan soal tersebut lebih fokus pada ingatan atau drill sehingga siswa mengerjakan soal tersebut minim proses penalaran.

Secara detail, guru memiliki keterbatasan kemampuan sebagai berikut. Pertama, keterbatasan memahami indikator untuk membuat soal. Kedua, keterbatasan membuat soal tipe pilihan ganda terutama pengecoh dan pokok soal. Ketiga, guru belum memperhatikan secara detail aspek materi dan keadaan kognitif siswa secara berimbang. Keempat, guru belum melakukan analisis soal secara kualitatif (materi, konstruksi maupun bahasa). Kelima, soal belum pernah diujicobakan sehingga karakteristik soal secara kuantitatif tidak ada. Kalau soal dibuat dengan keterbatasan kemampuan, tentu karakteristik soal tidak dapat dijadikan dasar pengambilan keputusan baik bagi siswa, guru, sekolah, maupun pihak lain untuk meningkatkan proses pembelajaran di kelas dan meningkatkan mutu pendidikan (Kerlinger, 1973; Cangelosi, 1990; Al- Qahtani & Higgins, 2013; Rukli, 2016). Artinya, soal tersebut tidak dapat dijadikan referensi kemampuan siswa yang sesungguhnya karena belum sepenuhnya menggambarkan tingkat pencapaian kompetensi siswa. Soal kredibel di samping melalui telaah kualitatif juga analisis secara kuantitatif agar soal memiliki karakteristik (daya beda, tingkat kesukaran, pengecoh, validitas, dan reliabilitas tes) (Allen & Yen, 1979; Mardapi, 2008; Azwar, 2008; Secolsky & Deniso, 2012; Wainer & Braun, 2013; Wu & Jen, 2017).

Penyelidikan selanjutnya tentang karakteristik soal, tidak ada soal buatan guru tersebut berkarakter kredibel. Saat ditanyakan kepada guru di sekolah tersebut tentang soal tersebut, sebagian guru mengata-kan bahwa sudah pernah mendengar dan membaca di media online namun belum memahami tahapan untuk membuat soal tersebut. Soal kredibel dapat me-motivasi siswa untuk belajar bernalar sesuai kondisi dan pemahaman awal dalam proses belajar mengajar di kelas (Savery, 2015; Smith, 2016; Geary et al., 2017). Karakter soal tersebut dibutuhkan siswa untuk

bernalair sesuai realitas ekosistem sekolah dan untuk memberikan solusi sesuai zamannya. Dengan adanya soal kredibel di sekolah, soal tersebut dapat digunakan guru saat ujian semester, kenaikan kelas, bahkan ujian akhir sekolah. Selanjutnya dapat tampil sukses dalam setiap perlombaan antar sekolah, daerah, nasional bahkan internasional. Menurut Wijaya et al. (2015) bahwa guru dapat menawarkan kesempatan untuk mempelajari tugas berbasis konteks kepada siswa jika ada waktu yang cukup.

Salah satu contoh soal bertipe kredibel adalah soal *Tren in Mathematic and Science Study* (TIMSS). Wijaya et al. (2014) mengemukakan bahwa siswa membuat sebagian besar kesalahan dalam dua tahap pertama dari proses solusinya. Dari total jumlah kesalahan 38% yang ada kaitan pemahaman makna tugas berbasis konteks. Kesalahan pemahaman ini terutama mencakup pemilihan informasi yang relevan.

Selanjutnya, Stacey (2014) mengatakan bahwa hampir 70% siswa Indonesia berada di bawah level 2 dari semua topik di *Programme for International Student Assesment* (PISA). Jupri et al. (2014) mengemukakan bahwa kemampuan siswa untuk menterjemahkan bolak-balik antara dunia situasi masalah dan dunia matematika dan untuk menata ulang sistem matematis itu sendiri, merupakan kesulitan yang paling sering diamati baik dalam tes tulis maupun data wawancara. Jensen et al. (2014) mengemukakan bahwa siswa yang diuji saat semester menggunakan soal kredibel untuk memperoleh pemahaman konseptual secara mendalam tentang materi dan memori yang lebih baik dan memberi dukungan pada sifat hirarki yang diusulkan dari taksonomi bloom. Lebih lanjut Ganapathy (2017) mengungkapkan bahwa penting mengenai potensi peluang teknologi dalam memfasilitasi kredibel namun kesuksesannya terletak pada tugas yang dirancang secara tepat untuk mempromosikan konten.

Hasil demikian membuat posisi Indonesia selalu berada di sekitar posisi bawah pada kelompok bawah pada setiap perlombaan Internasional. Hasil kajian Rukli (2017) tentang karakter soal TIMSS menunjuk-kan bahwa soal tipe pilihan ganda lebih

digemari siswa Indonesia dibandingkan soal tipe konstruk respon (uraian) pada semua domain konten dan domain kognitif. Domain kognitif pengetahuan lebih digemari daripada aplikasi, aplikasi lebih digemari daripada penalaran pada kedua tipe soal. Mata pelajaran aljabar paling tidak dikuasai siswa terutama soal bertipe konstruk respon.

Hasil penelitian Stanford University yang dilaksanakan tahun 2008, menyimpulkan bahwa tes PISA ini bisa dijadikan sebagai acuan, karena peningkatan peringkat dalam tes PISA dengan peningkatan rata-rata skor akhir sebuah negara, berkorelasi positif dengan peningkatan GDP (*Gross Domestic Product*) negara tersebut. Sebagai sebuah sistem penilaian, PISA mengembangkan instrumen penilaian sains dan matematika, untuk sebahagiannya mengukur *critical thinking* dan *problem solving*. Dengan demikian, dalam soal tersebut ada beberapa indikator yang mengukur kemampuan dua kompetensi penting tersebut untuk kemajuan sebuah bangsa. Berdasarkan hal tersebut, peningkatan pemahaman dan keterampilan guru dibutuhkan untuk ditingkatkan agar karakteristik soal sesuai karakteristik soal kredibel pada sekolah mitra.

#### *B. Permasalahan Mitra*

1. Berdasarkan aspek pembuatan kisi-kisi soal kredibel, guru memiliki keterbatasan pemahaman tentang indikator dan kisi-kisi soal. Tahapan tersebut sebagai awal perancangan soal agar lebih terarah pembuatan soal sesuai tujuan tes.
2. Berdasarkan aspek penulisan soal kredibel tipe pilihan ganda, guru memiliki keterbatasan pemahaman dan keterampilan tentang pokok soal dan pengecoh sesuai indikator dan kisi-kisi soal.
3. Berdasarkan aspek analisis secara kualitatif, guru memiliki keterbatasan pemahaman dan keterampilan tentang aspek materi, konstruksi, dan bahasa soal secara detail. Analisis tersebut dibutuhkan untuk meningkatkan mutu soal agar lebih kredibel.

#### *C. Solusi Menyelesaikan Permasalahan*

1. Guru mengikuti pendampingan mengenai perancangan dan pembuatan indikator dan kisi-kisi soal.
2. Guru mengikuti pendampingan mengenai pembuatan pokok soal dan pengecoh pada soal tipe pilihan ganda sesuai indikator dan kisi-kisi soal.
3. Guru mengikuti pendampingan mengenai analisis soal secara kualitatif.

#### *D. Target luaran*

1. Guru memiliki kisi-kisi soal berdasarkan materi dan tingkatan kognitif tertentu sesuai mata pelajaran masing-masing
2. Guru memiliki peningkatan pemahaman dan keterampilan untuk membuat pokok soal dan pengecoh pada soal tipe pilihan ganda sesuai indikator dan kisi-kisi soal.
3. Guru memiliki peningkatan pemahaman dan keterampilan analisis soal secara kualitatif untuk diterapkan di sekolahnya.

## **II. METODE PELAKSANAAN**

### *A. Tahapan Pelaksanaan*

Pendampingan perancangan pembuatan soal kredibel dilakukan di sekolah mitra. Pendampingan dilakukan sesuai kesepakatan antara guru dan kepala sekolah dengan Tim Pengusul. Tahapan pendamping-an adalah sebagai berikut.

1. Tahap identifikasi
  - a. Identifikasi lokasi sasaran secara terfokus pada sekolah mitra.
  - b. Membentuk kepanitian pendampingan yang terdiri atas pihak guru dari sekolah mitra dan kepala sekolah sebagai koordinator.
  - c. Melakukan pengecekan dan sosialisasi ke sekolah mitra tentang pendampingan analisis soal kredibel kepada guru tentang jadwal pendampingan supaya tidak mengganggu proses belajar di sekolah mitra.
2. Tahap persiapan materi soal kredibel
  - a. Setiap guru membawa soal masing-masing sesuai mata pelajarannya ke tempat klinik soal.

- b. Identifikasi materi sesuai dengan tingkatan kelas dimana guru mengajar yakni kelas VII, VIII, dan IX.
- c. Identifikasi ranah kognitif sesuai karakteristik soal kredibel
- d. Stimulus pada pokok soal atau pertanyaan dasar menggunakan data teks bacaan, paragraf, teks drama, penggalan novel/cerita/dongeng, puisi, kasus, gambar, grafik, foto, rumus, tabel, daftar kata/symbol, contoh, peta, film, atau suara yang direkam, dianalisis, dievaluasi, dan dikreasikan khusus soal tipe uraian.
- e. Pembuatan indikator terbagi dalam beberapa bagian yakni menfokuskan pada pertanyaan, menganalisis argumen, mempertimbangkan laporan observasi, membandingkan kesimpulan, menentukan kesimpulan, mempertimbangkan kemampuan induksi, menilai, mendefinisikan konsep, mendefinisikan asumsi, dan mendeskripsikan.

3. Tahap pelaksanaan

- a. Guru dikelompokkan sesuai mata pelajaran.
- b. Perancangan kisi-kisi soal memperhatikan kompetensi khusus dan pengembangan kognitif siswa untuk mendorong motivasi, kreatifitas, inspiratif, kemandirian, dan semangat belajar.
- c. Pengenalan dan pembuatan soal tipe pilihan ganda. Soal pilihan ganda ditekankan pada pembuatan stem dan pengecoh.
- d. Analisis soal secara kualitatif menggunakan tabel penelaahan soal baik dari segi materi, konstruksi, maupun bahasa secara silang antar guru mata pelajaran yang sejenis.
- e. Finalisasi dan pelaporan jumlah soal yang baik secara kualitatif.

B. Uraian Materi

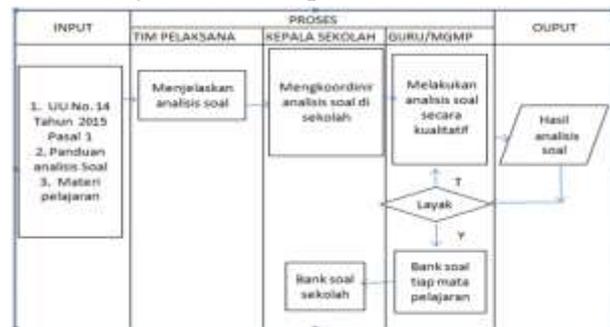
Uraian materi, tujuan dan waktu pendampingan pada Tabel 1. Kegiatan pendampingan dilakukan dua hari yakni hari pertama terdiri dari tahap I, II, dan III, sedangkan hari kedua terdiri dari tahap IV dan V.

Tabel 1. Uraian materi, tujuan dan waktu

Tahap	Materi	Tujuan	Waktu
I	Identifikasi kurikulum, standar kompetensi lulusan (skl), kompetensi inti, kompetensi dasar, dan indikator	memberikan pengetahuan kepada guru tentang kurikulum, skl, ki, kd, dan indikator dalam pembuatan soal	1 jam
II	Perancangan kisi-kisi soal	melatih guru merancang kisi-kisi soal	1 jam
	Penulisan soal	melatih guru menulis soal pilihan ganda dan uraian	1 jam
III	Analisis soal secara kualitatif	melatih guru menganalisis soal berdasarkan materi, konstruksi, dan bahasa	1 jam
IV	Finalisasi	melatih guru mengoreksi sendiri soal secara kualitatif atau kelompok guru mata pelajaran yang sejenis	2 jam
V	Pelaporan	memberikan informasi tata cara penyimpanan soal kredibel sebagai soal inti dan soal cadangan serta bank soal secara kualitatif	2 jam

C. Alur Analisis Butir Soal

Alur analisis butir soal melibatkan mitra dalam bentuk diagram aktivitas pada Gambar 1.



Gambar 1. Alur analisis soal

D. Keahlian Personal

Jenis kepakaran yang diperlukan dalam melakukan pendampingan perancangan dan pembuatan soal kredibel yakni pakar evaluasi dan bahasa. Pakar tersebut yakni Ketua Tim Dr. Drs. Rukli, M.Pd., M.Cs. merupakan pakar dibidang evaluasi dimana S2 dan S3 di Universitas Negeri Yogyakarta menempuh program studi Penelitian dan Evaluasi Pendidikan. Sedangkan anggota tim adalah Dr. Munirah, M.Pd. merupakan lulusan S2 Pendidikan Bahasa di PPS Universitas Negeri Makassar dan S3 Linguistik PPS Universitas Hasanuddin. Keduanya bersinergi untuk menyelesaikan seluruh persoalan sesuai kebutuhan mitra.

### III. HASIL DAN PEMBAHASAN

Hasil kegiatan pengabdian masyarakat secara garis besar mencakup beberapa komponen sebagai berikut: 1) Keberhasilan target jumlah peserta pelatihan; 2) Ketercapaian tujuan pelatihan; 3) Ketercapaian target materi yang telah direncanakan; dan 4) Kemampuan peserta dalam membuat soal kredibel.

Target peserta pendampingan seperti direncanakan sebelumnya kurang lebih 30 orang guru. Dalam pelaksanaannya, kegiatan ini diikuti oleh 30 orang peserta. Dengan demikian dapat dikatakan bahwa target peserta tercapai 100%. Angka tersebut menunjukkan bahwa kegiatan pengabdian dilihat dari jumlah peserta yang mengikuti pendampingan dalam pembuatan soal cukup sukses.

Ketercapaian tujuan pendampingan peningkatan pengetahuan tentang soal kredibel secara umum sudah baik, namun keterbatasan waktu yang disediakan mengakibatkan tidak semua materi tentang uji coba soal dan analisis soal secara kuantitatif dapat disampaikan dan dilakukan di lapangan. Namun dilihat dari hasil latihan para peserta yaitu kualitas soal secara kualitatif telah dibuat sangat baik, maka dapat disimpulkan bahwa tujuan kegiatan ini dapat tercapai.

Ketercapaian target materi pada kegiatan pengabdian ini cukup baik, karena materi pendamping-an telah dapat disampaikan secara keseluruhan walau-pun soal belum diuji coba di lapangan. Materi pendampingan yang telah disampaikan sebagai berikut: 1) Pengantar pembuatan soal yang baik. Teori pengukuran skala klasik plus skala modern; 2) Tata cara pembuatan indikator penilaian kaitan indikator pembelajaran; 3) Prosedur dan format analisis soal secara kualitatif; 4) Praktek pembuatan indikator dan soalnya; dan 5) Praktek analisis soal secara kualitatif.

Kemampuan peserta dilihat dari penguasaan materi masih kurang dikarenakan waktu yang singkat dalam penyampaian materi dan kemampuan para peserta yang berbeda-beda terkait pengalaman guru dan keterlibatan dalam pelatihan

sebelumnya. Hal ini disebabkan jumlah materi yang banyak hanya disampaikan dalam waktu dua hari sehingga tidak cukup waktu bagi para peserta untuk memahami dan mempraktekkan secara optimal.

Secara keseluruhan kegiatan pendampingan pembuatan soal kredibel mendapatkan minimal 20 soal tiap guru. Keberhasilan ini selain diukur dari keenam komponen di atas, juga dapat dilihat dari kepuasan peserta setelah mengikuti kegiatan. Manfaat yang diperoleh guru yakni guru dapat membuat soal kredibel yang diharapkan mengikuti standar untuk dapat dipakai sebagai contoh lebih lanjut untuk dikembangkan menjadi bank soal. Contoh soal kredibel melalui pembuatan soal sebelum dan setelah pendampingan sebagai berikut.

<p>Soal buatan guru pada sekolah mitra  <b>Berapa nilai dari <math>\frac{3}{2} - \frac{4}{3}</math> ?</b></p> <p>a. <math>\frac{1}{-1}</math></p> <p>b. <math>\frac{1}{5}</math></p> <p>c. <math>\frac{-1}{6}</math></p> <p>d. <math>\frac{1}{6}</math></p> <p>Soal buatan guru berorientasi hasil dan tahap berpikir rendah serta soal tersebut sudah ada di beberapa buku cetak dan online. Disamping itu, soal tersebut belum melalui analisis kualitatif sehingga belum kredibel untuk dijadikan referensi kemampuan siswa.</p>
<p>Soal buatan guru di atas direvisi optionnya melalui pendampingan menjadi soal kredibel.</p> <p><b>Metode apa yang tepat untuk menyelesaikan <math>\frac{3}{2} - \frac{4}{3}</math> ?</b></p> <p>a. <math>\frac{3-4}{2-3}</math></p> <p>b. <math>\frac{4-3}{2-3}</math></p> <p>c. <math>\frac{6-9}{2 \times 3}</math></p> <p>d. <math>\frac{9-8}{2 \times 3}</math></p> <p>Soal di atas berorientasi proses dan nalar tingkat tinggi. Disamping itu, soal tersebut melalui proses analisis kualitatif yakni materi, konstruk, dan bahasa sesuai indikator.</p>

### IV. KESIMPULAN

Kegiatan pendampingan pengabdian masyarakat telah membantu guru meningkatkan pengetahuan tentang pembuatan soal kredibel melalui proses

pembuatan indikator soal ujian, pembuatan pokok soal dan option serta analisis soal secara kualitatif. Analisis soal secara kualitatif dilakukan oleh guru mata pelajaran yang sama baik segi materi, konstruksi, dan bahasa. Ada banyak soal diproses (dibuang, diedit, diterima) oleh guru sehingga jumlah soal kredibel didapatkan guru tidak sama. Walaupun demikian, setiap guru mampu membuat soal kredibel minimal 20 butir dimana soal tersebut berkategori soal cukup baik untuk dipakai, disimpan, atau dikembangkan sebagai bank soal di sekolah.

### UCAPAN TERIMA KASIH

Saya mengucapkan banyak terima kasih kepada Rektor Universitas Muhammadiyah Makassar yang telah memberikan bantuan dana internal pengabdian lewat Lembaga Penelitian dan Pengabdian Masyarakat sehingga kegiatan pendampingan pembuatan soal dapat berjalan sesuai rencana, laporan hasil pengabdian rampung sesuai target, dan publikasi hasil tercapai.

### DAFTAR PUSTAKA

- Allen, M. J., & Yen, W. M., *Introduction to measurement theory*. Waveland Press, 2001.
- Al- Qahtani, A. A., & Higgins, S. E., *Effects of traditional, blended and e- learning on students' achievement in higher education*. Journal of Computer Assisted Learning, 29(3), 220-234, 2013.
- Azwar, *Dasar-dasar psikometri*. Yogyakarta: Pustaka Pelajar, 1999.
- Cangelosi, J. S., *Designing tests for evaluating student achievement*. Longman Pub Group, 1990
- Ganapathy, M., Singh, M. K. M., Kaur, S., & Kit, L. W., *Promoting Higher Order Thinking Skills via Teaching Practices*. 3L: Language, Linguistics, Literature®, 23(1), 2017.
- Geary, D. C., Berch, D. B., Ochsendorf, R., & Koepke, K. M. (Eds.), *Acquisition of Complex Arithmetic Skills and Higher-Order Mathematics Concepts* (Vol. 3). Academic Press., 2017.
- Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A., *Teaching to the test or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding*. Educational Psychology Review, 26(2), 307-329, 2014.
- Jupri, A., Drijvers, P., & van den Heuvel-Panhuizen, M., *Difficulties in initial algebra learning in Indonesia*. Mathematics Education Research Journal, 26(4), 683-710., 2014.
- Kerlinger, F. N., & Lee, H. B., *Foundations of behavioral research*, 1999.
- Mardapi, D., *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitrs Cendikia Offset, 2018.
- Rukli, *Analisis Item Information Function on the Test of Mathematics*. Proceeding International Conference on Educational Research and Evaluation (ICERE).29-31 Mei 2016, Yogyakarta, Indonesia.Hal.348-354. 2016.
- Rukli, *Characteristics Analysis Of Math Items Type On Data Trend In International Mathematics And Science Study (TIMSS)*. Proceeding Internasional Conference ADRI, 20-21 Januari 2017, Makassar, Indonesia.Hal.235-245, 2017.
- Savery, J. R., *Overview of problem-based learning: Definitions and distinctions. Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows*, 5-15, 2015
- Secolsky, C., & Denison, D. B. (Eds.), *Handbook on measurement, assessment, and evaluation in higher education*. Routledge, 2012
- Smith, S. H., *Instructional Strategies to Build Higher-Order Thinking Skills and Develop Language Ability in the Secondary Spanish Classroom*. Studies in Teaching 2016 Research Digest, 43, 2016.
- Stacey, K., *The PISA view of mathematical literacy in Indonesia*. Journal on Mathematics Education, 2(2), 95-126, 2014.
- Wainer, H., & Braun, H. I., *Test validity*. Routledge, 2013
- Wijaya, A., van den Heuvel-Panhuizen, M., Doorman, M., & Robitzsch, A., *Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors*. The Mathematics Enthusiast, 11(3), 555, 2014.
- Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M., *Teachers' teaching practices and beliefs regarding context-based tasks and their relation with students' difficulties in solving these tasks*. Mathematics Education Research Journal, 27(4), 637-662, 2015.
- Wu, M., Tam, H. P., & Jen, T. H., *Educational Measurement for Applied Researchers: Theory into Practice*. Springer, 2017.