

KUALITAS PEER ASSESSMENT SEBAGAI ASSESSMENT FORMATIF

Siti Rabiatul Adawiyah^{1*} & Nofisulastri²

¹Program Studi Pendidikan Olahraga dan Kesehatan, FIKKM,
Universitas Pendidikan Mandalika, Indonesia

²Program Studi Pendidikan Biologi, FSTT,
Universitas Pendidikan Mandalika, Indonesia

E-mail : siti_ra@yahoo.co.id

ABSTRAK: Penelitian ini bertujuan untuk menghasilkan instrumen *peer assessment* yang layak sebagai asesmen formatif. Jenis penelitian ini adalah penelitian pengembangan, karena mengembangkan instrumen *peer assessment*. Pengembangan instrumen *peer assessment* mengacu pada model pengembangan Borg & Gall yang meliputi 10 tahapan, yaitu: 1) studi pendahuluan; 2) perencanaan; 3) pengembangan model permulaan produk; 4) penelaahan model permulaan produk; 5) revisi; 6) uji coba terbatas; 7) revisi hasil uji coba; 8) uji coba skala luas; 9) revisi produk akhir; dan 10) penyebarluasan. Namun, dalam penelitian ini hanya sampai pada tahap 8, sedangkan tahap 9 dan tahap 10 tidak dilakukan disebabkan waktu dan kondisi yang tidak memungkinkan. Validasi oleh pakar dilakukan terlebih dahulu sebelum diujicoba terbatas dan ujicoba skala luas di Kelas X Madrasah Aliyah An-Najah, Yayasan Pendidikan Al-Halimy, Sesela, Kecamatan Gunungsari, Kabupaten Lombok Barat. Hasil penelitian menunjukkan bahwa, validitas logis instrumen *peer assessment* tergolong sangat baik, dengan skor rata-rata 3,7. Hasil analisis dengan uji korelasi *product moment* untuk mengetahui validitas empiris *peer assessment* menunjukkan validitas empiris *peer assessment* tergolong dalam kategori tinggi (0,94; 0,96; 0,97; dan 0,97). Berdasarkan hasil penelitian disimpulkan bahwa, *peer assessment* memiliki kualitas yang baik sebagai asesmen formatif.

Kata Kunci: Kualitas, *Peer Assessment*, *Assessment* Formatif.

ABSTRACT: This study aims to produce a suitable *peer assessment* instrument as a formative assessment. This type of research is development research, because it develops a *peer assessment* instrument. The development of the *peer assessment* instrument refers to the Borg & Gall development model which includes 10 stages, namely: 1) a preliminary study; 2) planning; 3) development of a product initiation model; 4) reviewing the initiation model of the product; 5) revision; 6) limited trials; 7) revised trial results; 8) wide-scale trials; 9) final product revision; and 10) dissemination. However, in this study only up to stage 8, while stage 9 and stage 10 were not carried out due to time and conditions that did not allow. Expert validation was carried out beforehand before limited testing and wide-scale trials in Class X Madrasah Aliyah An-Najah, Al-Halimy Education Foundation, Sesela, Gunungsari District, West Lombok Regency. The results showed that the logical validity of the *peer assessment* instrument was very good, with an average score of 3.7. The results of the analysis using the product moment correlation test to determine the empirical validity of the *peer assessment* showed that the empirical validity of the *peer assessment* was in the high category (0.94; 0.96; 0.97; and 0.97). Based on the results of the study, it was concluded that *peer assessment* was of good quality as a formative assessment.

Keywords: Quality, *Peer Assessment*, *Formative Assessment*.

PENDAHULUAN

Salah satu perubahan pada implementasi pembelajaran adalah perubahan pada proses asesmen hasil belajar. Asesmen hasil belajar pada kurikulum 2013 tidak hanya menekankan pada hasil belajar, namun juga pada proses pembelajaran (Direktorat Jenderal Guru dan Tenaga Kependidikan, 2018). Oleh karena itu, pada kurikulum 2013, asesmen formatif menjadi sangat penting untuk dilakukan oleh



guru, dimana asesmen ini bertujuan sebagai umpan balik terhadap proses pembelajaran yang telah dilakukan.

Asesmen formatif dilakukan selama proses pembelajaran berlangsung yang bertujuan untuk mengecek apakah proses pembelajaran dapat mengarahkan peserta didik mencapai tujuan pembelajaran. Rahmawati, Hartono, & Nugroho (2015) menjelaskan bahwa, asesmen formatif tidak dibuat untuk menggantikan tes tertulis (penilaian sumatif), melainkan merupakan upaya untuk melengkapi keterbatasan tes tertulis yang hanya mengukur hasil akhir tanpa melihat proses belajar peserta didik. Dalam hal ini, Guru dituntut terampil dalam memilih dan menggunakan strategi asesmen, agar asesmen formatif yang diberikan lebih efektif. Salah satu strategi asesmen yang dapat digunakan guru untuk melaksanakan asesmen formatif adalah *peer assessment*.

Peer assessment merupakan strategi penilaian dimana seorang peserta didik menilai kinerja atau kesuksesan belajar yang dicapai oleh peserta didik lainnya. Noonan & Duncan dalam Sumarti, Harjito, & Purwinarko (2015) mengungkapkan bahwa, *peer assessment* dideskripsikan sebagai suatu strategi penilaian yang melibatkan peserta didik dalam menyimpulkan kinerja temannya ketika mereka dilibatkan dalam kerja kelompok. Hal yang sama diungkapkan oleh White dalam Adawiyah (2013), *peer assessment* didefinisikan sebagai suatu penilaian dimana seorang peserta didik dilibatkan dalam mempertimbangkan jumlah, level, nilai, kualitas, atau kesuksesan produk atau hasil belajar temannya yang berada dalam satu tingkatan kelas yang sama. Berdasarkan kedua definisi di atas dapat disimpulkan bahwa, *peer assessment* merupakan sebuah cara untuk menilai kinerja peserta didik dengan melibatkan seorang peserta didik untuk menilai kinerja atau kesuksesan peserta didik lainnya yang memiliki tingkatan kelas yang sama ketika mereka diorganisasikan dalam kerja kelompok.

Lebih lanjut Zulharman dalam Siswaningsih, Dwiyaniti, & Gumilar (2013) menjelaskan bahwa, *peer assessment* dapat diterapkan untuk menilai kemampuan kognitif maupun kemampuan non kognitif peserta didik apabila dilihat dari kemampuan yang ingin diuji dan dapat sebagai alat penilaian formatif dan sumatif apabila dilihat dari tujuan penilaian. Namun, terjadi perdebatan mengenai penggunaan *peer assessment* sebagai alat penilaian, yaitu meliputi masalah kualitas *peer assessment* yang mencakup validitas dan reliabilitasnya sehingga metode ini lebih sering diterapkan sebagai penilaian formatif daripada sumatif.

Perdebatan mengenai permasalahan penerapan *peer assessment* sebagai salah satu strategi penilaian dalam proses pembelajaran yaitu yang berkaitan dengan kualitas *peer assessment* harus diselesaikan. Oleh karena itu, perlu dilakukan penelitian untuk mengetahui kualitas *peer assessment* yang mencakup validitas dan reliabilitasnya. Hasil penelitian ini dapat dijadikan sebagai dasar untuk menyimpulkan bahwa, strategi penilaian *peer assessment* dapat menghasilkan hasil penilaian yang baik dan dapat dijadikan acuan untuk melakukan evaluasi pembelajaran baik evaluasi formatif ataupun evaluasi sumatif.



METODE

Jenis Penelitian

Penelitian ini merupakan penelitian pengembangan, karena diawali dengan pengembangan instrumen *peer assessment*. Pengembangan instrumen *peer assessment* dalam penelitian ini berdasarkan pada model pengembangan Borg & Gall dalam Wijayanti, Sukestiyarno, & Masrukhi (2014), yang meliputi 10 tahapan, yaitu: 1) studi pendahuluan; 2) perencanaan; 3) pengembangan model permulaan produk; 4) penelaahan model permulaan produk; 5) revisi; 6) uji coba terbatas; 7) revisi hasil uji coba; 8) uji coba skala luas; 9) revisi produk akhir; dan 10) penyebarluasan. Namun, karena waktu dan kondisi yang tidak memungkinkan maka dalam penelitian ini tahap pengembangan hanya dilakukan sampai tahap ke-8. Lembar *peer assessment* yang dikembangkan digunakan sebagai asesmen formatif untuk menilai ranah afektif, yaitu keterampilan sosial peserta didik. Instrumen *peer assessment* yang dikembangkan disebut lembar *peer assessment*. Lembar penilaian ini terdiri atas 12 aspek penilaian yang dibagi ke dalam empat komponen, yaitu: 1) keterampilan sikap kooperatif; 2) keterampilan kepemimpinan; 3) keterampilan memfasilitasi pemahaman; dan 4) keterampilan tantangan intelektual. Aspek-aspek yang dinilai dalam lembar *peer assessment* dipertegas oleh Johnson & Johnson dalam Syahrul (2015).

Subjek Penelitian

Subjek dalam penelitian ini adalah peserta didik kelas X Madrasah Aliyah An-Najah Yayasan Pendidikan Al-Halimy, Sesela, Kecamatan Gunungsari, Kabupaten Lombok Barat. Pemilihan sampel ujicoba terbatas dan ujicoba skala luas dalam penelitian ini dilakukan secara acak (*random sampling*). Sampel ujicoba terbatas terdiri dari 12 orang peserta didik, sedangkan sampel pada ujicoba skala lebih luas terdiri dari satu kelas yaitu sebanyak 25 orang peserta didik.

Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan di Madrasah Aliyah An-Najah Yayasan Pendidikan Al-Halimy, Sesela, Kecamatan Gunungsari, Kabupaten Lombok Barat pada tahun ajaran 2019/2020, mulai bulan Januari 2019-Juni 2019.

Teknik Pengambilan Data

Lembar *peer assessment* yang dikembangkan terdiri atas lembar *peer assessment* untuk peserta didik dan lembar *peer assessment* yang digunakan oleh guru, keduanya berisi aspek dan skala penilaian yang sama. Lembar *peer assessment* untuk peserta didik digunakan oleh peserta didik untuk menilai keterampilan sosial teman sekelompoknya, sedangkan lembar *peer assessment* yang digunakan oleh guru diisi oleh guru dan asesor untuk menilai keterampilan peserta didik. Skala penilaian pada lembar *peer assessment* menggunakan 5 skala penilaian, yaitu skala 5, 4, 3, 2, dan 1 yang ditentukan berdasarkan rubrik *peer assessment*. Lembar *peer assessment* yang dikembangkan pada tahap 3, selanjutnya ditelaah oleh para pakar (tahap 4). Penelaahan ini dimaksudkan untuk memperoleh validitas logis dari 3 orang pakar pendidikan.

Lembar *peer assessment* yang telah direvisi berdasarkan saran dan masukan pakar (tahap 5), kemudian diujicoba dalam kelompok terbatas (tahap 6).



Ujicoba terbatas dilaksanakan dalam 4 kali pertemuan di mata pelajaran biologi yang menerapkan metode pembelajaran kooperatif tipe NHT (*Numbered Head Together*). Pada ujicoba terbatas, lembar *peer assessment* digunakan oleh guru atau asesor untuk menilai masing-masing peserta didik selama proses kerja kelompok berlangsung. Sementara itu, peserta didik menggunakan lembar *peer assessment* untuk menilai seluruh anggota kelompoknya. Penilaian baik oleh guru dan peserta didik dilaksanakan pada setiap kali pertemuan.

Hasil evaluasi berdasarkan ujicoba secara terbatas dijadikan bahan pertimbangan untuk melakukan revisi terhadap lembar *peer assessment* dan implementasinya (tahap 7), sebelum digunakan pada ujicoba skala luas (tahap 8). Pelaksanaan ujicoba skala luas sama dengan ujicoba terbatas, yaitu dilaksanakan dalam 4 kali pertemuan, dimana di setiap kali pertemuan dilakukan asesmen formatif menggunakan lembar *peer assessment*, baik dilakukan oleh guru maupun peserta didik. Namun sebelum ujicoba skala luas dilaksanakan, terlebih dahulu diberikan pelatihan kepada peserta didik tentang tata cara penerapan *peer assessment*, terutama tentang kriteria penilaian dan prosedur penilaian.

Teknik Analisis Data

Instrumen pengumpulan data dalam penelitian ini adalah lembar validasi lembar *peer assessment* dan lembar *peer assessment*. Lembar validasi lembar *peer assessment* digunakan untuk mengumpulkan data kevalidan lembar *peer assessment* berdasarkan pendapat pakar. Sementara itu, lembar *peer assessment* digunakan untuk memperoleh validitas empiris *peer assessment*. Validitas empiris dianalisis dengan rumus korelasi *product moment*.

Data hasil validasi pakar untuk lembar *peer assessment* dianalisis secara deskriptif kualitatif. Dalam penelitian ini, *passing grade* (batas penerimaan) adalah skor rerata (\bar{X}) dari hasil penilaian para validator, yang kemudian ditentukan tingkat validitasnya dengan mencocokkan skor rerata (\bar{X}) dengan kriteria pengkategorian seperti yang tersaji pada Tabel 1.

Tabel 1. Kriteria Pengkategorian Validasi Pakar.

Interval Skor	Kategori Penilaian
$3.6 \leq \bar{X} < 4.0$	Sangat Baik
$2.6 \leq \bar{X} < 3.5$	Baik
$1.6 \leq \bar{X} < 2.5$	Kurang Baik
$1.0 \leq \bar{X} < 1.5$	Tidak Baik

(Sumber: Ratumanan & Laurends (2006) dalam Sudiarman, Soegimin, & Susantini, 2015).

Validitas empiris dianalisis dengan menggunakan rumus korelasi *product moment*. Ada atau tidaknya kesesuaian antara hasil penilaian oleh guru dan peserta didik dinyatakan dalam angka pada koefisien korelasi. Interpretasi tinggi rendahnya kesesuaian dapat diketahui dari besar kecilnya angka dalam indeks korelasi. Makin besar angka dalam indeks korelasi, maka semakin tinggi kesesuaian antara hasil penilaian oleh guru dan peserta didik. Besarnya koefisien korelasi menjadi indikator tinggi atau rendahnya validitas *peer assessment*.

HASIL DAN PEMBAHASAN

Lembar *peer assessment* yang telah dikembangkan kemudian divalidasi oleh dua orang pakar yaitu validator 1 dan validator 2 menggunakan instrumen lembar validasi lembar *peer assessment*. Hasil validitas logis oleh pakar disajikan pada Tabel 2. Lembar *peer assessment* yang telah divalidasi oleh pakar, kemudian direvisi berdasarkan saran dan masukan dari validator. Lembar *peer assessment* yang telah direvisi, kemudian diujicoba dalam kelompok terbatas dengan sampel sebanyak 12 orang. Hasil penilaian oleh *peer* (peserta didik), selanjutnya dianalisis untuk memperoleh validitas empiris *peer assessment*, yang dapat dilihat pada Tabel 3.

Tabel 2. Hasil Validasi Lembar *Peer Assessment* oleh Pakar.

Aspek	Komponen	Penilaian Validator				Rerata Aspek	Kategori
		V1	V2	ΣV	\bar{x}		
Format	1	3	4	7	3.5	3.7	Sangat Baik
	2	4	4	8	4		
	3	4	4	8	4		
	4	4	4	8	4		
	5	3	3	6	3		
Isi	1	3	4	7	3.5	3.7	Sangat Baik
	2	4	4	8	4		
	3	3	4	7	3.5		
Bahasa	1	4	4	8	4	3.8	Sangat Baik
	2	4	4	8	4		
	3	4	4	8	4		
	4	3	3	6	3		
Skor Rata-rata Seluruh Aspek						3.7	Sangat Baik
Reliabilitas						0.75	

Keterangan:

V1 = Validator 1;

V2 = Validator 2.

Tabel 3. Hasil Validitas Empiris *Peer Assessment* pada Ujicoba Terbatas.

Kode Peserta Didik	Pertemuan			
	1	2	3	4
1	0.9	0.9	0.9	0.99
2	0.92	0.94	0.95	0.98
3	0.9	0.96	0.94	0.97
4	0.94	0.97	0.95	0.94
5	0.92	0.97	0.98	0.99
6	0.9	0.98	0.97	0.96
7	0.9	0.99	0.95	0.98
8	0.9	0.98	0.96	0.99
9	0.87	0.97	0.96	0.96
10	0.8	0.97	0.98	0.99
11	0.87	0.97	0.96	0.97
12	0.87	0.96	0.98	0.93
\bar{x}	0.89	0.96	0.96	0.97

Hasil dan proses ujicoba terbatas menjadi bahan pertimbangan dalam melakukan revisi terhadap *peer assessment*. Hasil revisi yang selanjutnya



diimplementasikan pada ujicoba skala luas. Berdasarkan proses dan hasil ujicoba terbatas, validitas *peer assessment* terus meningkat dari pertemuan ke pertemuan. Peningkatan ini disebabkan karena peserta didik semakin memahami kriteria dan prosedur penilaian dari pertemuan ke pertemuan. Temuan ini sesuai dengan yang dikemukakan oleh Smith, Cooper, & Lancaster dalam Adawiyah (2013) bahwa, *peer assessment* sebagai teknik penilaian sebenarnya dapat dikatakan valid dan reliabel, baik sebagai penilaian formatif dan penilaian sumatif, asalkan kriteria yang digunakan untuk menilai dijelaskan secara jelas oleh guru. Lebih lanjut Zulharman dalam Siswaningsih, Dwiyaniti, & Gumilar (2013) menyatakan bahwa, proses *peer assessment* dimulai dengan mendiskusikan item dan kriteria penilaian oleh dosen dan peserta didik. Kemudian masing-masing peserta didik menilai teman mereka dan juga memberikan *feedback*. Namun pada penelitian ini, kriteria penilaian ditentukan sepenuhnya oleh peneliti, dengan pertimbangan bahwa peserta didik belum pernah melaksanakan teknik *peer assessment* sebelumnya. Oleh karena itu, agar hasil dan proses *peer assessment* menjadi lebih baik diujicoba skala luas, maka dilakukan pelatihan tentang kriteria penilaian dan prosedur *peer assessment* kepada peserta didik sebelum kegiatan ujicoba skala luas dilaksanakan.

Hasil penilaian guru dan *peer* menggunakan lembar *peer assessment* dianalisis untuk mengetahui validitas empiris *peer assessment*. Validitas empiris *peer assessment* dapat diketahui dengan menghitung koefisien korelasinya dengan menggunakan uji korelasi *product moment*. Hasil analisis penilaian antara guru dengan peserta didik (validitas empiris *peer assessment*) ujicoba skala luas tersaji pada Tabel 4.

Tabel 4. Validitas Empiris *Peer Assessment* Ujicoba Skala Luas.

Kode Peserta Didik	Pertemuan			
	1	2	3	4
1	0.96	0.98	0.9	0.9
2	0.98	0.99	0.96	0.96
3	0.97	0.99	0.97	0.97
4	0.95	0.98	0.98	0.95
5	0.96	0.97	0.99	0.98
6	0	0.96	0.99	0.98
7	0.97	0.94	0.97	0.98
8	0.97	0.97	0.98	0.96
9	0.95	0.99	0.99	0.97
10	0.95	0.96	0.98	0.96
11	0.94	0.97	0.96	0.98
12	0.97	0.98	0.98	0.97
13	0.94	0.98	0.97	0.98
14	0.93	0.96	0.97	0.97
15	0.93	0.93	0.98	0.97
16	0.93	0.91	0.97	0.98
17	0.95	0.98	0.98	0.98
18	0.92	0.98	0.98	0.98
19	0.89	0.95	0.98	0.98
20	0.9	0.97	0.98	0.99
21	0.95	0.98	0.98	0.98



22	0.96	0.97	0.98	0.97
23	0.83	0.93	0.98	0.97
24	0.87	0.86	0.98	0.96
25	0.94	0.91	0.97	0.97
\bar{x}	0.94	0.96	0.97	0.97

Berdasarkan Tabel 4 dapat dilihat bahwa, kesesuaian antara penilaian yang diberikan oleh guru dan asesor dengan penilaian yang diberikan oleh *peer* (peserta didik) tergolong dalam kategori tinggi, dimana rata-rata koefisien korelasi secara berturut-turut yaitu 0,94; 0,96; 0,97; dan 0,97. Oleh karena itu, dapat dikatakan bahwa *peer assessment* memiliki validitas yang baik sebagai asesmen formatif. Hal ini sejalan dengan yang dikemukakan oleh Smith, Cooper, & Lancaster dalam Adawiyah (2013), dimana mereka merangkum beberapa penelitian yang merupakan bukti bahwa validitas dan reliabilitas *peer assessment* masih menjadi perdebatan, dan berdasarkan rangkuman beberapa penelitian tersebut, disimpulkan bahwa sebenarnya *peer assessment* dapat dikatakan valid dan reliabel baik sebagai penilaian formatif dan penilaian sumatif asalkan kriteria yang digunakan untuk menilai dijelaskan secara jelas oleh guru.

Jika dibandingkan dari pertemuan ke pertemuan, kualitas penilaian yang diberikan oleh peserta didik terus meningkat. Hal ini terlihat dari validitas empiris *peer assessment* yang meningkat dari pertemuan ke pertemuan. Peningkatan kualitas penilaian *peer* ini menunjukkan bahwa, dengan terus menerus mengimplementasikan *peer assessment* dalam proses pembelajaran, peserta didik akan lebih memahami prosedur penilaian yang kemudian akan meningkatkan kualitas hasil penilaian *peer*. Hal ini dipertegas oleh Nurhardini (2017) bahwa, sebanyak 88% siswa menyatakan bahwa penerapan *self* dan *peer assessment* memacu untuk berpikir, bersikap aktif dan memacu kejujuran, dan aspek menghargai pendapat teman, melatih tanggung jawab dan mengakui kepandaian teman yang dinyatakan oleh 29 orang dari 32 siswa.

Topping dalam Indasari, Purwati, & Anam (2019) menyatakan bahwa, meskipun penilai sejawat (*peer*) memiliki kemampuan menilai yang rendah, namun dengan terus latihan dalam menilai maka dia akan bisa menilai lebih valid dan reliabel bahkan dari gurunya. Oleh karena itu, untuk menerapkan *peer assessment* ini, guru harus melatih peserta didik terlebih dahulu sehingga dapat memahami kriteria penilaian dengan baik, sekaligus melatih peserta didik dalam melakukan penilaian secara objektif. Pada peserta didik yang sebelumnya tidak pernah melaksanakan *peer assessment*, penilaian dari peserta didik harus dibandingkan dengan penilaian yang diberikan oleh guru. Alzaid (2017) menyatakan bahwa, dalam implementasi *peer assessment*, terdapat dua hal yang sangat penting yaitu melatih keterampilan *peer assessment* dan membandingkan penilaian peserta didik dengan penilaian guru. Hal yang sama juga ditegaskan oleh Speyer, *et. al.* (2011) bahwa, membandingkan penilaian guru dengan penilaian yang diberikan oleh peserta didik sangat diperlukan agar penilaian menjadi valid dan reliabel. Oleh karena itu, sebelum menentukan nilai peserta didik, guru hendaknya membandingkan penilaian yang dilakukannya dengan penilaian peserta didik.

Untuk menentukan apakah penilaian *peer* dapat digunakan, maka harus dilakukan analisis. Zulharman dalam Siswaningsih, Dwiyantri, & Gumilar (2013) menegaskan bahwa, penilaian yang diberikan oleh peserta didik dapat diterima jika selisih antara penilaian yang diberikan oleh guru dan peserta didik kurang dari 10%. Berdasarkan koefisien korelasi yang diperoleh dari perhitungan uji korelasi *product moment* pada Tabel 4, terlihat bahwa terdapat beberapa peserta didik yang selisih penilaian antara guru dan *peer* $\geq 10\%$, yaitu peserta didik 19, 23, dan 24 pada pertemuan pertama serta pada pertemuan kedua yaitu peserta didik 24. Oleh karena itu, maka penilaian yang diberikan oleh *peer* untuk peserta didik 19, 23, dan 24 pada pertemuan pertama dan untuk peserta didik 24 pada pertemuan kedua tidak dapat diterima, sehingga yang digunakan untuk menentukan nilai keterampilan sosial peserta didik ini adalah skor yang diberikan oleh guru.

SIMPULAN

Teknik *peer assessment* merupakan penilaian yang memiliki kualitas yang baik sebagai asesmen formatif. Hal ini ditunjukkan dari hasil perhitungan korelasi yang tinggi, yang menjadi indikator bahwa validitas *peer assessment* termasuk dalam kategori tinggi.

SARAN

Bagi guru yang ingin menerapkan *peer assessment* sebagai asesmen formatif hendaknya menjelaskan kriteria penilaian dengan baik dan melatih keterampilan *peer assessment* sehingga *peer assessment* dapat berjalan lebih efektif. Bagi peneliti yang ingin melakukan kajian lebih lanjut tentang *peer assessment*, hendaknya mengimplementasikan *peer assessment* pada ranah hasil belajar yang lain (ranah kognitif dan psikomotorik).

UCAPAN TERIMA KASIH

Tim peneliti mengucapkan terima kasih yang sebesar-besarnya kepada pihak-pihak yang telah banyak membantu terlaksananya penelitian ini, khususnya keluarga besar Madrasah Aliyah An-Najah Yayasan Pendidikan Al-Halimy, Sesela, Kecamatan Gunungsari, Kabupaten Lombok Barat yang dengan keramahannya telah memfasilitasi kami selama kegiatan penelitian berlangsung.

DAFTAR RUJUKAN

- Adawiyah, S. R. (2013). Melatihkan Keterampilan Sosial Siswa melalui Model Pembelajaran Kooperatif Tipe NHT yang Mengintegrasikan *Peer Assessment*. *Seminar Nasional IPA IV* (pp. 219-225). Semarang, Indonesia: Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang.
- Alzaid, M. J. (2017). The Effect of Peer Assessment on the Evaluation Process of Students. *International Education Studies*, 10(6), 159-169.



- Direktorat Jenderal Guru dan Tenaga Kependidikan. (2018). *Modul Manajemen Implementasi Kurikulum 2013*. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia.
- Indasari, N. L., Purwati, O., & Anam, S. (2019). The Impact of Peer Assessment on Students Recount Text. *ETERNAL (English, Teaching, Learning, and Research Journal)*, 5(2), 221-238.
- Nurhardini, R. (2017). Pengaruh Self dan Peer Assessment pada Materi Ekosistem terhadap Berpikir Aplikatif dan Kritis Siswa SMA. *Jurnal Pendidikan Matematika dan Sains*, 5(1), 69-76.
- Rahmawati, I., Hartono, & Nugroho, S. E. (2015). Pengembangan Asesmen Formatif untuk Meningkatkan Kemampuan *Self Regulation* Siswa pada Tema Suhu dan Perubahannya. *Unnes Science Education Journal*, 4(2), 842-850.
- Siswaningsih, W., Dwiyaniti, G., & Gumilar, C. (2013). Penerapan Peer Assessment dan Self Assessment pada Tes Formatif Hidrokarbon untuk Feedback Siswa SMA Kelas X. *Jurnal Pengajaran MIPA*, 18(1), 107-115.
- Speyer, R., Pilz, W., Kruis, J. V. D., & Brunings, J. W. (2011). Reliability and Validity of Student Peer Assessment in Medical Education: A Systematic Review. *Medical Education*, 33(11), 572-585.
- Sudiarman, Soegimin, W. W., & Susantini, E. (2015). Pengembangan Perangkat Pembelajaran Fisika Berbasis Inkuiri Terbimbing untuk Melatihkan Keterampilan Proses Sains dan Meningkatkan Hasil Belajar pada Topik Suhu dan Perubahannya. *Jurnal Penelitian Pendidikan Sains (JPPS)*, 4(2), 658-671.
- Sumarti, S. S., Harjito, & Purwinarko, A. (2015). Design A Project Based Learning and Authentic Assessment Management System. *International Conference on Mathematics, Science, and Education (ICMSE 2015)* (pp. 29-36). Semarang, Indonesia.
- Syahrul. (2015). Analisis Konsistensi dan Validitas Asesmen Teman Sejawat oleh Mahasiswa dan Asesmen Unjuk Kerja oleh Dosen dalam Pembelajaran Praktikum pada Fakultas Teknik Universitas Negeri Makassar. *Jurnal Mekom*, 2(1), 73-87.
- Wijayanti, T., Sukestiyarno, & Masrukhi. (2014). Pengembangan Instrumen Penilaian Pembelajaran (Implementasi Penilaian Kemampuan Berpikir Kritis dan Karakter Demokratis pada Materi Sistem Politik Indonesia dengan Metode Pembelajaran *Role Playing* Berbasis Konservasi. *Innovative Journal of Curriculum and Educational Technology*, 3(2), 30-37.