

## **Analisis Kualitas Penilaian Virtual Program *Wordwall Game Quiz* Materi Sistem Reproduksi Manusia**

**Dwi Ayu Rahmawati & Didi Nur Jamaludin**

Tadris Biologi, Institut Agama Islam Negeri Kudus

[dwirahma264@gmail.com](mailto:dwirahma264@gmail.com)

### **ABSTRACT**

*This research aims to analyze the quality of the items on the human reproductive system taken from The Wordwall Game Quiz program in terms of validity, reliability, level of difficulty, discriminating power, and effectiveness of the distractor. This research uses quantitative approach and the type is descriptive research. The population is all answer sheets belonging to class XI IPA. The samples were 3 types of human reproductive system questions taken from The Wordwall Game Quiz program and 100 student answer sheets. The variable is an analysis of the quality of the items on the human reproductive system in The Wordwall Game Quiz program. Data collection techniques using tests, interviews, and documentation. Quantitative analysis using the Anates 4.0.9 application and Microsoft Excel 2019. The results showed that the three types of questions had high content and construct validity. In terms of statistical validity, it has sufficient validity. The reliability is in the high category. From the level of difficulty, dominated by questions in the easy category. From the aspect of discriminating power, the item has a fairly good discriminating power and from the aspect of the effectiveness of the distractor, it has the effectiveness of the detractor which is not good.*

*Keywords: Test Quality, Human Reproductive, Wordwall*

### **ABSTRAK**

Penelitian ini bertujuan untuk menganalisis kualitas butir soal sistem reproduksi manusia yang diambil dari program *Wordwall Game Quiz* ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Penelitian ini menggunakan penelitian kuantitatif. Jenis penelitian ini merupakan penelitian deskriptif (*descriptive research*). Populasi pada penelitian ini adalah seluruh lembar jawaban soal milik kelas XI IPA. Sampel dalam penelitian ini adalah 3 tipe soal sistem reproduksi manusia yang diambil dari program *Wordwall Game Quiz* dan lembar jawaban siswa sebanyak 100 lembar. Variabel dalam penelitian ini merupakan analisis kualitas butir soal sistem reproduksi manusia pada program *Wordwall Game Quiz*. Teknik pengumpulan data menggunakan test, wawancara, dan dokumentasi. Analisis secara

kuantitatif ini dilakukan dengan bantuan aplikasi Anates 4.0.9 dan Microsoft Excel 2019. Hasil penelitian menunjukkan bahwa ketiga tipe soal memiliki validitas isi dan konstruk tinggi. Adapun dari segi validitas statistik, ketiga tipe soal memiliki validitas cukup. Reliabilitas ketiga tipe soal berkategori tinggi. Dari tingkat kesulitan, ketiga tipe soal didominasi oleh soal dengan kategori mudah. Sedangkan, dari aspek daya pembeda, soal sistem reproduksi manusia dari *Wordwall Game Quiz* memiliki daya pembeda yang cukup baik dan dari segi efektivitas pengecohnya, ketiga soal memiliki efektivitas pengecoh yang kurang baik.

**Kata kunci:** Kualitas Soal, Reproduksi Manusia, *Wordwall*

## PENDAHULUAN

Alat evaluasi memiliki kedudukan yang penting dalam pelaksanaan pendidikan. Tanpa melalui proses evaluasi, keberhasilan pendidikan tidak dapat terukur dengan baik. Proses evaluasi juga membantu guru untuk mengetahui apakah siswa telah mengikuti pembelajaran dengan baik atau tidak. Hasil evaluasi pembelajaran akan akurat jika instrumen evaluasi (soal) yang digunakan berkualitas serta mengikuti kaidah dan prinsip dengan baik. (Jamaludin, 2020). Maka penelitian tentang analisis kualitas butir soal menjadi perlu dilakukan.

Perkembangan teknologi yang sangat pesat memungkinkan guru untuk melakukan penilaian secara *online*. Penilaian *online* menjadi populer sejak diberlakukan pembelajaran jarak jauh karena pandemi Covid-19. Guru dapat memanfaatkan aplikasi atau *website* yang tersedia di internet untuk melakukan penilaian *online*. Salah satu *website* yang dapat digunakan adalah *Wordwall Game Quiz*.

*Wordwall Game Quiz* memungkinkan guru untuk membuat dan menyimpan soal dalam bentuk *game*. Siswa dapat melihat skor nilai di akhir *game* serta dapat menuliskan namanya di papan *leaderboard*. Amany & Desire (2020) mengungkapkan bahwa hal tersebut dapat meningkatkan pola pikir logis dan memacu daya saing positif dalam pembelajaran. Hasil penelitian yang dilakukan Arimbawa (2021) juga menunjukkan bahwa penggunaan *Wordwall Game Quiz* dalam pembelajaran biologi dapat meningkatkan motivasi belajar siswa. Hal ini dikarenakan siswa merasa senang dapat mengeksplorasi konsepnya sendiri walaupun konsep tersebut belum sepenuhnya benar. Peningkatan motivasi belajar siswa ini diikuti juga dengan peningkatan prestasi belajar siswa.

*Wordwall* juga memiliki fitur komunitas yang memuat kumpulan soal dari berbagai mata pelajaran yang dibuat orang lain. Guru dan siswa dapat melihat serta mengerjakan latihan soal tersebut. Salah satu latihan soal yang dapat ditemukan adalah mengenai sistem reproduksi manusia.

Materi sistem reproduksi manusia memiliki daya tarik tersendiri dari materi lainnya. Siswa menunjukkan berbagai macam ekspresi ketertarikan baik secara verbal maupun nonverbal selama mempelajari sistem reproduksi. Macam-macam ekspresi tersebut dipengaruhi oleh persepsi siswa terhadap materi sistem reproduksi manusia yang dilihat berdasarkan perasaan pribadi (60%), nilai-nilai yang diyakininya (100%), kebutuhan dan minat siswa (80%), serta tingkat pengetahuan akan materi yang dipelajari (100%). (Mufidah dan Supriana, 2021)

Kualitas soal yang terdapat pada *Wordwall* memang belum diketahui. Soal-soal tersebut belum tentu melalui proses analisis sebelum diunggah. Untuk mengetahui kualitas soal yang ada pada program *Wordwall Game Quiz*, khususnya materi sistem reproduksi manusia, peneliti melakukan analisis terhadap butir soal yang meliputi aspek validitas (isi, konstruk, dan statistik), reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

## **METODE PENELITIAN**

Penelitian ini menggunakan penelitian kuantitatif. Jenis penelitian ini merupakan penelitian deskriptif (*descriptive research*). Penelitian deskriptif bertujuan untuk mendeskripsikan suatu keadaan dengan apa adanya tanpa melakukan manipulasi atau pemberian perlakuan-perlakuan tertentu terhadap objek penelitian. Penelitian deskriptif dapat menggunakan pendekatan kuantitatif dimana pengumpulan dan pengukuran data berbentuk angka (Sukmadinata, 2015).

Penelitian ini dilaksanakan di MA NU Al-Hidayah Kudus dan di MA NU Raudlatul Shiblyan. Populasi pada penelitian ini adalah seluruh lembar jawaban latihan soal sistem reproduksi dari program *Wordwall Game Quiz* milik siswa kelas XI IPA. Sampel dalam penelitian ini adalah 3 tipe soal sistem reproduksi manusia yang diambil dari program *Wordwall Game Quiz* dan lembar jawaban siswa sebanyak 100 lembar. Teknik pengumpulan data pada penelitian ini menggunakan test, wawancara, dan dokumentasi. Analisis data kuantitatif dilakukan dengan bantuan aplikasi Anates 4.0.9 dan Microsoft Excel 2019.

## **HASIL DAN PEMBAHASAN**

Analisis butir soal sistem reproduksi manusia pada program *Wordwall Game Quiz* ini menggunakan bantuan aplikasi Microsoft Excel 2019 dan Anates 4.0.9. Hasil analisis meliputi

validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan tingkat pengecoh akan dibahas sebagai berikut.

## 1. Validitas

Validitas merupakan ukuran untuk menentukan bahwa instrumen dapat mengukur apa yang seharusnya diukur dan dilakukan penilaian (Jamaludin, 2020). Analisis validitas ketiga tipe soal sistem reproduksi manusia ini meliputi validitas isi, validitas konstruk, dan validitas statistik. Perhitungan validitas isi dan konstruk dilakukan dengan aplikasi Microsoft Excel 2019 menggunakan rumus Aiken V sebagai berikut.

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

Keterangan :

s = r – lo

lo = angka penilaian validitas terendah

c = angka penilaian validitas tertinggi

r = angka yang diberikan oleh penilai

n = jumlah ahli yang menilai

Hasil perhitungan validitas isi dan konstruk disajikan dalam tabel berikut.

Tabel 1. Hasil Perhitungan Validitas Isi dan Konstruk

Type Soal	Validitas	Kategori
A	0.69	Tinggi
B	0.66	Tinggi
C	0.68	Tinggi

Adapun analisis validitas statistik dilakukan dengan bantuan program Anates 4.0.9. hasil analisis validitas masing-masing tipe soal disajikan dalam tabel berikut.

Tabel 2. Persentase Analisis Validitas Soal Tipe A

No	Kategori Validitas	Persentase		
		Type A	Type B	Type C
1	Sangat rendah (0,00 – 0,20)	13.3%	15%	5.9%
2	Rendah (0,21 – 0,40)	33,3%	25%	29.4%
3	Cukup (0,41 – 0,60)	53.4%	50%	64.7%
4	Tinggi (0,61 – 0,80)	0%	10%	0%
5	Sangat tinggi (0,81 – 1)	0%	0%	0%

Hasil perhitungan validitas isi dan konstruk soal tipe A menunjukkan nilai 0.69 yang berarti soal tersebut memiliki validitas tinggi. Dari 15 soal yang ada, hanya satu soal yang memiliki validitas dengan kategori cukup. Sedangkan 14 soal lainnya termasuk soal dengan kategori tinggi. Soal tipe B memiliki nilai validitas 0.66 yang berarti soal tersebut termasuk soal dengan kategori validitas tinggi. Empat dari 20 soal tipe B memiliki validitas yang cukup, sedangkan 16 soal sisanya memiliki validitas yang tinggi. Nilai validitas soal tipe C adalah sebesar 0.68 dengan kategori validitas soal tinggi. Dibandingkan dua tipe lainnya, soal tipe C memiliki validitas butir soal dengan kategori yang lebih beragam. Dari keseluruhan 17 soal, satu soal memiliki validitas yang rendah, dua soal bervaliditas cukup, dan 14 soal berkategori tinggi.

Adapun perhitungan validitas statistik dilakukan berdasarkan hasil yang didapat dari uji coba soal terhadap 100 siswa. Hasil perhitungan dikategorikan ke dalam batas signifikansi dengan derajat kebebasan  $N-2$ . Hasil analisis validitas soal tipe A menunjukkan 13.3% soal memiliki validitas yang sangat rendah, 33.3% soal memiliki validitas rendah, dan 53.4% soal memiliki validitas yang cukup. Soal tipe B memiliki 15% soal dengan kategori sangat rendah, 25% soal berkategori rendah, 50% soal berkategori cukup, 10% berkategori tinggi. Sedangkan soal tipe C memiliki 5.9% soal dengan kategori sangat rendah, 39.4% dengan kategori rendah, dan 64.7% soal.



Gambar 1. Soal dengan Validitas Tinggi

## 2. Reliabilitas

Reliabilitas berasal dari bahasa Inggris "*reliability*" yang berarti dapat dipercaya. Instrumen tes dikatakan reliabel jika hasil penilaian yang didapat memberikan hasil yang tetap atau konsisten (Widoyoko, 2018). Validitas dan reliabilitas memiliki keterkaitan. Jika suatu

instrumen soal dinyatakan valid, maka instrumen soal tersebut sudah pasti reliabel. Akan tetapi, instrumen soal yang reliabel belum tentu valid (Ghazali, 2016). Hasil perhitungan reliabilitas ketiga tipe soal disajikan dalam tabel berikut.

Tabel 3. Hasil Analisis Reliabilitas Tes

<b>Tipe Soal</b>	<b>Reliabilitas Tes</b>	<b>Kategori</b>
A	0.64	Tinggi
B	0.75	Tinggi
C	0.69	Tinggi

Koefisien reliabilitas soal sistem reproduksi manusia tipe A adalah sebesar 0.64. Hal itu berarti soal tipe A memiliki tingkat reliabilitas yang tinggi. Pada soal tipe B, didapat koefisien reliabilitas soal sebesar 0.75. Artinya, soal tipe B memiliki reliabilitas yang tinggi. Koefisien reliabilitas soal tipe C adalah sebesar 0.69, dimana nilai tersebut dapat diinterpretasikan bahwa soal tipe C memiliki tingkat reliabilitas yang tinggi. Dengan demikian, dapat dilihat bahwa ketiga tipe soal sistem reproduksi manusia yang diambil dari program *Wordwall Game Quiz* adalah reliabel. Hal ini juga membuktikan bahwa ketiga tipe soal sistem reproduksi yang valid, juga sudah tentu reliabel.

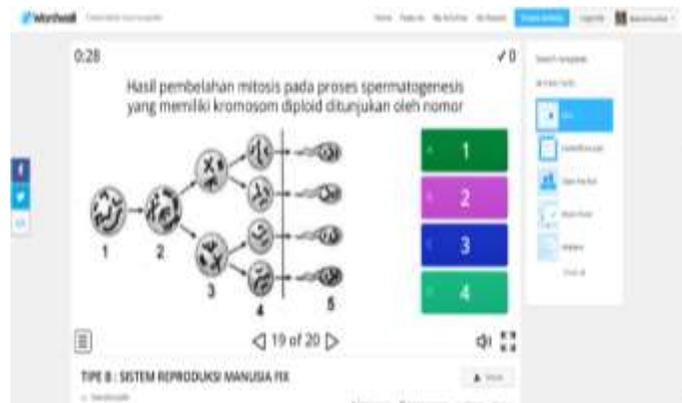
### 3. Tingkat Kesulitan

Soal yang baik merupakan soal yang tidak terlalu mudah sehingga dapat merangsang kemampuan berpikir tinggi, serta tidak terlalu sulit sehingga tidak menyebabkan siswa berputus asa. Keseimbangan tingkat kesulitan soal mengacu pada kurva normal dimana instrumen soal terdiri dari 25% mudah, 50% sedang, dan 25% sulit (Jamaludin, 2020). Persentase tingkat kesulitan masing-masing tipe soal disajikan dalam tabel berikut.

Tabel 4. Persentase Tingkat Kesulitan Tiap Tipe Soal

No	Kategori	Persentase		
		Tipe A	Tipe B	Tipe C
1	Sulit (0,0 – 0,30)	0%	0%	0%
2	Sedang (0,31 – 0,70)	26,7%	20%	17,6%
3	Mudah (0,71 – 1,00)	73,3%	80%	82,4%

Ketiga tipe soal sistem reproduksi manusia yang diambil dari program *Wordwall Game Quiz* memiliki tingkat kesukaran soal yang kurang baik. Hal ini ditunjukkan dengan tidak adanya soal yang berkategori sulit dan didominasi soal dengan kategori mudah. Soal tipe A hanya memiliki 26,7% soal dengan kategori sedang dan 73,3% berkategori mudah. Pada tipe B, 20% soal memiliki kategori sedang dan 80% berkategori mudah. Begitu pula dengan soal tipe C yang hanya memiliki 17,6% berkategori sedang dan 82,4% berkategori mudah. Persentasi diatas menunjukkan bahwa keseimbangan tingkat kesulitan ketiga tipe soal tidak mengacu pada kurva normal.



Gambar 2. Soal dengan Tingkat Kesulitan Sedang

#### 4. Daya Pembeda

Daya pembeda dalam suatu analisis butir soal adalah kemampuan suatu butir soal untuk membedakan kelompok siswa yang memiliki kemampuan tinggi dan kelompok siswa yang berkemampuan rendah (Bagiyono, 2017). Persentase daya pembeda masing-masing tipe soal disajikan dalam tabel berikut.

Tabel 5. Persentase Daya Pembeda Tiap Tipe Soal

No	Kategori	Persentase		
		Tipe A	Tipe B	Tipe C
1	Jelek (0,00 – 0,20)	6.7%	10%	23.5%
2	Cukup (0,21 – 0,40)	60%	50%	47.1%
3	Baik (0,41 – 0,70)	33.3%	40%	29.4%
4	Baik Sekali (0,71 – 1,00)	0%	0%	0%
5	Jelek (0,00 – 0,20)	6.7%	10%	23.5%

Hasil analisis daya pembeda soal sistem reproduksi manusia tipe A menunjukkan bahwa 6.7% soal memiliki daya pembeda yang jelek, 60% berkategori cukup, dan 33.3% berkategori baik. Soal sistem reproduksi tipe B memiliki 10% berdaya pembeda jelek, 50% berkategori cukup, dan 40% berkategori baik. Sedangkan soal tipe C memiliki 23.5% soal dengan kategori daya pembeda jelek, 47.1% berkategori cukup, dan 29.4% berkategori baik. Ketiga tipe soal yang diambil dari program *Wordwall Game Quiz* tidak memiliki soal yang berdaya pembeda baik sekali. Soal-soal tersebut hanya memiliki daya pembeda dengan kategori jelek, cukup, dan baik.



Gambar 3. Soal dengan Daya Pembeda Baik

## 5. Efektivitas Pengecoh

Tingkat pengecoh merupakan distribusi jawaban siswa peserta tes. Butir soal yang baik memiliki pengecoh yang dipilih secara merata oleh siswa yang menjawab salah (Jamaludin, 2020). Hasil analisis pada setiap butir soal tipe A dan tipe B yang memiliki 4 pilihan jawaban diinterpretasikan berdasarkan adaptasi dari skala *Likert* berikut.

- a. Apabila ketiga jawaban memiliki pengecoh yang baik, maka soal dikategorikan memiliki efektivitas pengecoh yang baik.
- b. Apabila terdapat dua jawaban yang memiliki pengecoh yang baik, maka soal tersebut dikategorikan memiliki efektivitas pengecoh yang cukup baik.
- c. Apabila hanya ada satu jawaban yang memiliki pengecoh yang baik, maka soal dikategorikan memiliki efektivitas pengecoh yang kurang baik.
- d. Apabila semua jawaban tidak berfungsi sebagai pengecoh, maka soal dikategorikan memiliki efektivitas yang tidak baik.

Hasil analisis tingkat pengecoh tipe soal A dan tipe B disajikan dalam tabel berikut.

Tabel 6. Efektivitas Pengecoh Soal Tipe A dan B

No	Kategori	Persentase	
		Tipe A	Tipe B
1	Baik	13.3%	20%
2	Cukup Baik	20%	25%
3	Kurang Baik	66.7%	45%
4	Tidak Baik	0%	10%
5	Baik	13.3%	20%

Sedangkan hasil analisis tingkat pengecoh soal tipe C diinterpretasikan berdasarkan adaptasi dari skala *Likert* berikut.

- a. Apabila keempat jawaban memiliki pengecoh yang baik, maka soal dikategorikan memiliki efektivitas pengecoh yang sangat baik.
- b. Apabila ketiga jawaban memiliki pengecoh yang baik, maka soal tersebut dikategorikan memiliki efektivitas pengecoh yang baik.
- c. Apabila terdapat dua jawaban yang memiliki pengecoh yang baik, maka soal dikategorikan memiliki efektivitas pengecoh yang cukup baik.
- d. Apabila hanya satu jawaban yang memiliki pengecoh yang baik, maka soal dikategorikan memiliki efektivitas pengecoh yang kurang baik.
- e. Apabila semua jawaban tidak berfungsi sebagai pengecoh, maka soal dikategorikan

Hasil analisis tingkat pengecoh tipe soal C disajikan dalam tabel berikut.

Tabel 7. Persentase Efektivitas Pengecoh Soal Tipe C

No.	Kategori	Persentase
1	Sangat Baik	0%
2	Baik	5.9%
3	Cukup Baik	52.9%
4	Kurang Baik	29.4%
5	Tidak Baik	11.8%

Analisis efektivitas pengecoh pada soal tipe A dan tipe B dibagi menjadi empat kategori, yaitu baik, cukup baik, kurang baik, dan tidak baik. Hal ini dikarenakan kedua tipe soal tersebut hanya memiliki empat pilihan jawaban. Hasil analisis menunjukkan 13.3% soal tipe A memiliki efektivitas pengecoh yang baik, 20% berkategori cukup baik, dan 66.7% berkategori kurang baik. Pada soal tipe B, 20% soal memiliki efektivitas pengecoh yang baik, 25% dalam kategori cukup baik, 45% berkategori kurang baik, dan 10% memiliki efektivitas pengecoh yang tidak baik.

Gambar 4. Soal dengan *Distractor* Baik

Berbeda dengan soal tipe A dan B, soal tipe C memiliki 5 kategori efektivitas pengecoh, yaitu sangat baik, baik, cukup baik, kurang baik, dan tidak baik. Kategori ini dihitung berdasarkan jumlah pilihan jawaban pada soal tipe C, yaitu lima. Hasil analisis menunjukkan tidak ditemukan soal yang memiliki efektivitas pengecoh dengan kategori sangat baik. Soal tipe C memiliki 5.9% dengan pengecoh yang berkategori baik, 52.9% memiliki pengecoh cukup baik, 29.4% berkategori kurang baik, dan 11.8% memiliki pengecoh yang tidak baik.

Penilaian wordwall quiz diterapkan sebagai media pembelajaran, dengan respon siswa lebih tertarik dan aktif dalam mengikuti kegiatan pembelajaran. Siswa lebih berani dalam

mengungkapkan pendapatnya baik di depan kelas maupun saat berdiskusi dengan teman kelompok, serta meningkatnya interaksi antar sesama sehingga terjadinya kefahaman terhadap materi yang diajarkan yang berdampak pada peningkatan hasil belajar (Sukma & Handayani, 2022). Penggunaan Quizizz juga sebagai penilaian formatif online untuk media pembelajaran berpengaruh terhadap hasil belajar siswa pada materi sistem periodik unsur (Dewi dkk, 2021). Berdasarkan hal tersebut penilaian online seperti *Wordwall Game Quiz* selain digunakan sebagai penilaian yang memiliki kualitas butir soal yang baik, juga dapat digunakan sebagai media pembelajaran.

## SIMPULAN

Secara umum, soal sistem reproduksi manusia dari *Wordwall Game Quiz* adalah valid dari segi isi, konstruk, dan statistik. Dari segi reliabilitas, soal sistem reproduksi manusia dari *Wordwall Game Quiz* adalah reliabel. Ditinjau dari aspek tingkat kesulitan, ketiga tipe soal didominasi oleh soal dengan kategori mudah. Sedangkan, dari aspek daya pembeda, soal sistem reproduksi manusia dari *Wordwall Game Quiz* memiliki daya pembeda yang cukup baik dan dari segi efektivitas pengecohnya, ketiga soal memiliki efektivitas pengecoh yang kurang baik.

## DAFTAR PUSTAKA

- Amany, D. dan Desire A. (2000). Pembelajaran Interaktif berbasis Gamifikasi guna Mendukung Program WFH pada saat Pandemic Covid-19. *ADI Bisnis Digital Interdisiplin Jurnal* Volume 1, No. 1 2020: 48-55.
- Arimbawa, I. G. P. A. (2021). Penerapan Word Wall Game Quis Berpadukan Classroom untuk Meningkatkan Motivasi dan Prestasi Belajar Biologi, *Indonesian Journal of Educational Development* l Volume 2, No 2 2021: 324-332 .
- Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat I, *Jurnal Widyanuklida* Volume 16 No 1 2017: 1-12
- Dewi, L. U., Irwandi, D & Bahriah, E. S. (2021) Pengaruh Media Penilaian Formatif Online Quizizz terhadap Hasil Belajar Siswa pada Materi Sistem Periodik Unsur. *Jurnal Riset Pendidikan Kimia*, 2021, l Volume 11, No 1, 2021.
- Ghazali, Nor Hasnida M D. A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System: A Pilot Study. *International Journal of Evaluation and Research in Education (IJERE)* 5, No 2 2016: 148-157
- Jamaludin, D B. (2020). *Pengembangan Evaluasi Pembelajaran*. Kudus: IAIN Kudus.
- Mufidah, Lailatul dan Supriana D. N. (2021). Students Perceptions Toward Humans Reproductive System Material (Case Study of Ninth Grade Students at SMP Negeri 1 Paciran). *Proceeding of International Conference on Education, Science, and Technology (ICNEST)*, Tuban, 19-20 Maret, 2021

- Sukma, K. I. & Handayani. T. (2022). Pengaruh Penggunaan Media Interaktif Berbasis Wordwall Quiz Terhadap Hasil Belajar IPA di Sekolah Dasar. *Jurnal Cakrawala Pendas*. Vol. 8 No. 4, 2022
- Sukmadinata, Nana Syaodih. (2015). *Metode Penelitian Pendidikan*. Bandung: PT Remaja Rosdakarya.
- Widoyoko, Eko Putro. (2018). *Penilaian Hasil Pembelajaran di Sekolah*. Yogyakarta: Pustaka Belajar.