

## **Evaluasi Kualitas Soal Penilaian Akhir Semester (PAS) Matematika Kelas VIII di SMPS Advent Singkawang: Analisis Validitas, Reliabilitas, Indeks Kesukaran, dan Daya Pembeda**

Yosef Imanuel \*, B. Buyung, Rika Wahyuni, Resy Nirawati  
Institut Sains dan Bisnis Internasional Singkawang, Indonesia

\*Corresponding Author e-mail: [yosefimanuel02@gmail.com](mailto:yosefimanuel02@gmail.com)

### **Abstract**

*Assessment is an integral part of the educational process that aims to measure learner competencies. One form of assessment that is often used in schools is the End of Semester Assessment (PAS). In ensuring that the questions used in the PAS can measure student competence validly and reliably, it is important to analyze the items. It aims to provide an overview of the quality of even PAS question items in Mathematics at SMPS Advent Singkawang and provide suggestions for future question improvements. The method used is descriptive quantitative with a documentation study approach. Data were collected from PAS question and student answer sheets for the 2023/2024 school year. Validity analysis shows that 27% (8 questions) are valid. Reliability is moderate (0,586). Difficulty index indicates 3% (1 question) easy, 77% (23 questions) moderate, and 20% (6 questions) difficult. Distinguishing power analysis shows 23% (7 questions) good, 30% (9 questions) sufficient, 20% (6 questions) poor, and 27% (8 questions) very poor. This study found that the validity of even PAS questions in mathematics class VIII at SMPS Advent Singkawang still needs to be improved, because most of the items are not valid. The reliability of the questions is in the sufficient category, but still requires improvement to reach a higher category. The difficulty index shows a balanced distribution of questions, while the distinguishing power shows that there needs to be an improvement in the preparation of questions to ensure good distinguishing power. This study provides an important overview of the quality of PAS questions and provides a basis for future improvements to improve the quality of assessment at SMPS Advent Singkawang.*

**Keywords:** *Analysis, Question Item Quality, The End of Semester Assessment, Validity, Reliability, Difficulty Index, Distinguishing Power*

### **1. Pendahuluan**

Penilaian merupakan bagian integral dari proses pendidikan yang bertujuan untuk mengukur pencapaian kompetensi siswa. Salah satu bentuk penilaian yang sering digunakan di sekolah adalah Penilaian Akhir Semester (PAS). Dalam memastikan bahwa soal-soal yang digunakan dalam PAS dapat mengukur kompetensi siswa secara valid dan reliabel maka penting untuk dilakukan analisis pada butir soal. Validitas dan reliabilitas soal merupakan dua aspek yang sangat penting yang menentukan kualitas instrumen penilaian. Validitas merujuk pada sejauh mana suatu tes dapat mengukur apa yang semestinya diukur, yaitu sejauh mana ketepatan dan

kecermatan alat ukur dalam menjalankan fungsinya (Sugiono *et al.*, 2020). Adapun reliabilitas berhubungan dengan konsistensi hasil yang diperoleh dari instrumen tersebut (Ndiung dan Jediut, 2020).

Beberapa penelitian sebelumnya telah menganalisis validitas dan reliabilitas butir soal PAS di berbagai mata pelajaran dan tingkat pendidikan. Hasil penelitian Anshari *et al.* (2024) menunjukkan bahwa pilihan ganda Ujian Akhir Semester (UAS) Pendidikan Agama Islam kelas X tahun ajaran 2022/2023 dinyatakan valid dan reliabel. Sementara itu, penelitian dari Novia *et al.* (2020) menemukan bahwa dari keseluruhan 15 butir soal Ujian Tengah Semester (UTS) Fisika kelas X tahun ajaran 2019/2020, 10 di antaranya valid, sedangkan 5 soal lainnya tidak valid. Meskipun demikian, soal-soal tersebut memiliki reliabilitas tinggi dengan nilai 0,72, sehingga dianggap baik serta layak digunakan.

Penelitian ini bertujuan untuk mengisi kekurangan penelitian sebelumnya dengan fokus pada analisis butir soal PAS genap pada mata pelajaran matematika kelas VIII di SMPS Advent Singkawang. Beberapa peneliti telah memfokuskan pada validitas dan reliabilitas soal secara umum, tetapi tidak memperhatikan secara khusus butir soal. Oleh karena itu, Penelitian ini memperhatikan analisis yang menyeluruh terhadap butir soal, yang belum banyak dibahas dalam penelitian sebelumnya. Selain itu, penelitian ini memperkenalkan metodologi yang lebih komprehensif dalam mengukur validitas dan reliabilitas butir soal dengan menambahkan analisis pada indeks kesukaran dan daya pembeda.

Tujuan dari penelitian ini adalah untuk memberikan gambaran tentang kualitas pada butir soal PAS genap mata pelajaran Matematika di SMPS Advent Singkawang dan memberikan saran untuk perbaikan soal di masa mendatang. Dengan demikian, penelitian ini diharapkan dapat memberikan manfaat untuk meningkatkan mutu penilaian di SMPS Advent Singkawang dan sebagai acuan bagi sekolah lain dalam melakukan evaluasi kualitas soal.

## **2. Metode**

### **2.1. Partisipan/Sampel Penelitian**

Menurut Sugiyono (2017), sampel adalah bagian dari jumlah dan karakteristik pada populasi. Sampel yang digunakan dalam penelitian ini adalah 12 siswa kelas VIII di SMPS Advent Singkawang.

### **2.2. Instrumen Penelitian**

Instrumen penelitian adalah alat yang digunakan untuk mengumpulkan berbagai data dalam penelitian, atau teknik yang digunakan dalam penelitian (Elan *et al.*, 2022). Instrumen yang digunakan dalam penelitian ini berupa soal PAS genap mata pelajaran matematika kelas VIII tahun ajaran 2023/2024.

### **2.3. Pengumpulan Data dan Analisis Data**

Penelitian ini menggunakan metode analisis data deskriptif kuantitatif dalam bentuk studi dokumentasi untuk menganalisis butir soal Penilaian Akhir Semester

(PAS) genap mata pelajaran Matematika kelas VIII di SMPS Advent Singkawang. Metode analisis data kuantitatif deskriptif merupakan cara yang digunakan untuk menggambarkan, menunjukkan, ataupun meringkas data secara konstruktif. Metode ini menitikberatkan pada penyajian statistik untuk membantu memahami rincian data melalui peringkasan serta identifikasi pola dari sampel data tertentu. (Sudirman *et al.*, 2023). Tujuan dari penelitian deskriptif bukan untuk menguji suatu hipotesis, melainkan hanya menggambarkan kondisi suatu variabel, gejala, serta keadaan sebagaimana adanya. (Zellatifanny dan Mudjiyanto, 2018). Adapun studi dokumentasi merupakan proses pengumpulan catatan tentang peristiwa yang telah terjadi berupa tulisan, gambar atau foto, maupun karya-karya monumental dari individu atau instansi (Sugiyono, 2020). Metode deskriptif digunakan untuk mendeskripsikan validitas butir soal, sementara metode dokumentasi digunakan untuk mengumpulkan data mengenai soal dan jawaban siswa.

Langkah awal yang dilakukan adalah memperoleh data berupa lembar soal PAS Genap mata pelajaran Matematika tahun ajaran 2023/2024 dan lembar jawaban siswa. Data dikumpulkan melalui dokumentasi lembar soal PAS genap mata pelajaran matematika, kunci jawaban, serta lembar jawaban siswa setelah pelaksanaan tes PAS Genap kepada seluruh siswa kelas VIII di sekolah tersebut. Soal-soal PAS yang digunakan dalam penelitian ini terdiri dari 30 butir soal pilihan ganda, dengan setiap soal memiliki 4 opsi jawaban: a, b, c, dan d. Adapun materi yang digunakan dalam soal PAS adalah Teorema Pythagoras, Lingkaran, Bangun Datar dan Bangun Ruang. Setelah data dikumpulkan, maka dilakukan analisis pada lembar soal PAS. Analisis yang dilakukan meliputi uji validitas, reliabilitas, indeks kesukaran, serta daya pembeda.

Untuk analisis data, peneliti menghitung validitas, reliabilitas, indeks kesukaran, dan daya pembeda dengan menggunakan bantuan aplikasi Microsoft Excel. Excel melakukan perhitungan statistik yang diperlukan dan untuk menampilkan data dalam format tabel serta grafik guna mempermudah interpretasi hasil. Penggunaan Excel membantu memastikan keakuratan perhitungan dan penyajian data yang terstruktur, yang mendukung analisis yang lebih detail dan menyeluruh mengenai kualitas butir soal PAS genap mata pelajaran Matematika kelas VIII di SMPS Advent Singkawang.

## **2.4. Pertimbangan Etis**

Penelitian ini memperhatikan berbagai pertimbangan etis untuk memastikan integritas dan kerahasiaan data yang diperoleh. Peneliti menjamin kerahasiaan informasi pribadi siswa dengan memberikan kode pada lembar jawaban dan tidak menyertakan nama atau data identitas lainnya yang dapat membocorkan identitas siswa. Data-data yang dikumpulkan hanya digunakan untuk keperluan penelitian ini serta disimpan dengan aman untuk mencegah akses dari pihak yang tidak bertanggung jawab. Selain itu, peneliti juga menjelaskan tujuan penelitian kepada pihak sekolah. Dengan demikian, penelitian ini berusaha mematuhi prinsip-prinsip etika penelitian dalam pendidikan, seperti menghormati privasi peserta, mendapatkan persetujuan yang tepat, dan menjaga integritas data.

## **2.5. Keterbatasan Studi**

Meskipun penelitian ini telah dilakukan dengan cermat, terdapat beberapa keterbatasan yang perlu diperhatikan yaitu ukuran sampel yang relatif kecil, yaitu hanya terdiri dari 12 siswa. Jumlah sampel yang terbatas ini dapat mempengaruhi kemampuan untuk menggeneralisasi hasil penelitian ke populasi yang lebih luas. Selain itu, data yang digunakan hanya mencakup satu mata pelajaran dan satu kelas saja, sehingga hasil analisis mungkin tidak mencerminkan kualitas soal pada mata pelajaran lain atau kelas lain di sekolah yang sama. Adapun keterbatasan lainnya adalah kemungkinan adanya kekeliruan dalam penghitungan yang dilakukan oleh peneliti, yang dapat mempengaruhi keakuratan hasil analisis.

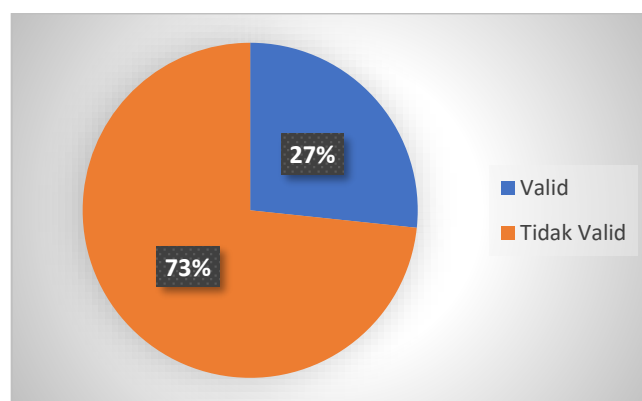
## **3. Hasil dan Pembahasan**

### **3.1. Hasil**

Setelah pengumpulan data oleh peneliti, data tersebut kemudian diolah melalui berbagai teknik analisis untuk menghasilkan informasi yang dapat diinterpretasikan dan digunakan dalam pengambilan keputusan atau untuk tujuan penelitian lebih lanjut. Adapun data yang diolah adalah validitas butir soal, reliabilitas, indeks kesukaran, serta daya pembeda.

#### **a. Analisis Validitas Butir Soal**

Validitas butir soal dihitung menggunakan indeks poin biserial dengan taraf signifikansi 5%. Soal dinyatakan valid jika  $r_{\text{hitung}} > r_{\text{tabel}}$ . Sebaliknya, soal dinyatakan tidak valid jika  $r_{\text{hitung}} \leq r_{\text{tabel}}$ . Hasil penghitungan validitas soal disajikan pada Gambar 1. berikut.



Gambar 1. Hasil Validitas Butir Soal

Dari Gambar 1 yang disajikan di atas diperoleh 27% atau 8 soal yang valid, sementara 73% atau 22 soal lainnya tidak valid.

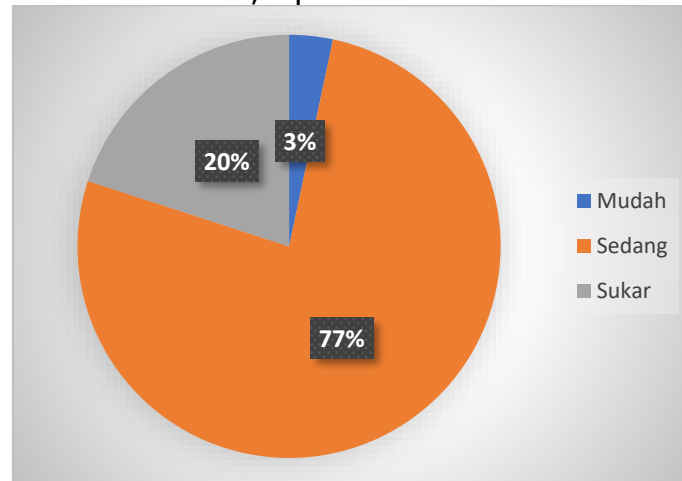
#### **b. Analisis Reliabilitas**

Berdasarkan analisa data terhadap soal PAS Genap Mata Pelajaran Matematika Kelas VIII dengan menggunakan rumus *Cronbach Alpha*, diperoleh hasil bahwa nilai

reliabilitas soal  $r_{11}$  sebesar 0,586. Berdasarkan hasil tersebut, nilai ini termasuk dalam kategori cukup.

**c. Analisis Indeks Kesukaran**

Setelah dilakukan analisis soal, diperoleh hasil indeks kesukaran sebagai berikut.

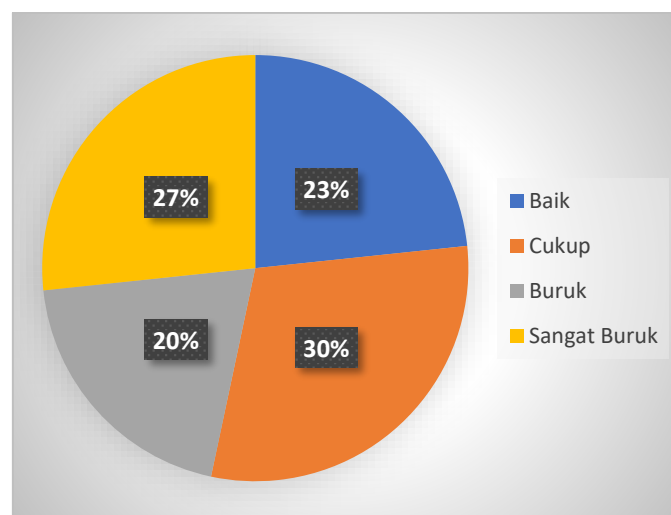


Gambar 2. Hasil Indeks Kesukaran

Gambar 2 yang disajikan di atas menunjukkan bahwa butir soal yang tergolong mudah sebesar 3% atau sebanyak 1 soal, sedang sebesar 77% atau sebanyak 23 soal, dan sukar sebesar 20% atau sebanyak 6 soal.

**d. Analisis Daya Pembeda**

Setelah dilakukan pengujian data, diperoleh hasil daya pembeda soal sebagai berikut.



Gambar 3. Hasil Indeks Kesukaran

Dari Gambar 3 yang disajikan di atas, diperoleh hasil bahwa 23% atau 7 soal memiliki daya pembeda baik, 30% atau 9 soal cukup, 20% atau 6 soal buruk, serta 27% atau 8 soal sangat buruk.

### **3.2. Pembahasan**

Penelitian ini menganalisis kualitas butir soal PAS Genap mata pelajaran Matematika kelas VIII di SMPS Advent Singkawang menggunakan metode deskriptif kuantitatif melalui studi dokumentasi. Adapun data yang diolah berupa validitas butir soal, reliabilitas, indeks kesukaran, serta daya pembeda. Berikut adalah pembahasan hasil analisis yang telah dilakukan:

#### **a. Analisis Validitas Butir Soal**

Validitas butir soal merupakan ketepatan pengukuran yang dimiliki oleh suatu butir soal, yang merupakan bagian integral dari keseluruhan tes (Sukiman dalam Rais *et al.*, 2022). Validitas butir soal dihitung menggunakan indeks poin biserial dengan taraf signifikansi sebesar 5%. Berdasarkan hasil analisis, diperoleh bahwa dari 30 butir soal, hanya 27% atau 8 soal yang valid yang terdiri dari butir soal nomor 4, 5, 6, 11, 12, 16, 22, dan 29. Sementara 73% atau 22 soal tidak valid yang terdiri dari butir soal nomor 1, 2, 3, 7, 8, 9, 10, 13, 14, 15, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, dan 30.

Dari hasil tersebut dapat disimpulkan bahwa sebagian besar soal PAS genap pada mata pelajaran matematika kelas VIII di SMPS Advent Singkawang tidak memenuhi kriteria validitas yang diharapkan. Hal ini dapat disebabkan oleh berbagai faktor salah satunya jarang dilakukan peninjauan maupun uji coba soal sebelum digunakan dalam tes. Adapun cara untuk mengatasinya salah satunya yaitu merevisi soal dengan menata ulang kalimat pada soal. Hasil ini juga mengindikasikan bahwa perlu dilakukan pelatihan lebih lanjut bagi guru dalam menyusun soal yang valid untuk memastikan bahwa setiap butir soal benar-benar mengukur kompetensi yang diharapkan.

#### **b. Analisis Reliabilitas**

Reliabilitas merupakan keakuratan atau konsistensi suatu alat dalam menilai apa yang dinilainya (Ramadhan *et al.*, 2024). Reliabilitas soal diukur menggunakan rumus Cronbach Alpha. Nilai reliabilitas yang diperoleh dari analisis adalah 0,586, yang termasuk dalam kategori cukup. Meskipun nilai tersebut menunjukkan konsistensi yang cukup baik, masih ada ruang untuk meningkatkan reliabilitas agar mencapai tingkat yang lebih optimal. Peningkatan reliabilitas dapat dilakukan dengan memperbaiki butir soal yang tidak valid dan memastikan setiap butir soal secara konsisten mengukur kemampuan yang sama. Proses ini melibatkan peninjauan kembali serta pengembangan soal baru yang lebih sesuai. Selain itu, guru perlu mempertimbangkan variasi dalam penyusunan soal untuk menghindari bias dan meningkatkan konsistensi hasil pengukuran.

#### **c. Analisis Indeks Kesukaran**

Indeks kesukaran merupakan angka yang menunjukkan tingkat kesulitan dari suatu butir soal (Hutasuhut, 2022). Apabila suatu soal memiliki tingkat kesukaran yang seimbang, maka soal tersebut dapat dianggap memiliki kualitas yang baik (Arifin, 2017). Dari hasil analisis, diperoleh bahwa 3% atau 1 soal tergolong mudah, 77%

atau 23 soal tergolong sedang, dan 20% atau 6 soal tergolong sukar. Dalam hal ini, mayoritas soal berada dalam kategori sedang, yang menunjukkan tingkat kesulitan yang seimbang dan sesuai dengan kemampuan rata-rata siswa. Namun, distribusi soal yang ideal sebaiknya mencakup lebih banyak variasi untuk menantang berbagai tingkat kemampuan siswa. Oleh karena itu, dalam pembuatan soal berikutnya, guru disarankan untuk menyusun soal dengan tingkat kesukaran yang bervariasi untuk mengukur kemampuan siswa dengan lebih komprehensif. Soal yang terlalu mudah atau terlalu sulit mungkin tidak efektif dalam mengevaluasi seluruh kemampuan siswa, sehingga perlu ada keseimbangan yang lebih baik pada soal.

#### **d. Analisis Daya Pembeda**

Daya pembeda merupakan ukuran yang menunjukkan seberapa jauh suatu butir soal dapat membedakan antara siswa yang telah maupun belum/kurang menguasai kompetensi menurut kriteria tertentu (Sardia *et al.*, 2021). Hasil analisis menunjukkan bahwa 23% atau 7 soal memiliki daya pembeda yang baik, 30% atau 9 soal cukup, 20% atau 6 soal buruk, dan 27% atau 8 soal sangat buruk. Mengetahui daya pembeda pada setiap butir soal sangatlah penting, sebab pedoman dalam menyusun soal salah satunya ialah mengetahui bahwa kemampuan siswa beraneka ragam dan butir-butir soal harus dapat memberikan hasil yang menunjukkan perbedaan kemampuan di kalangan siswa tersebut (Mayasari, 2023). Daya pembeda yang buruk dapat disebabkan oleh beberapa hal, salah satunya adalah soal yang tidak jelas dan ambigu. Soal yang memiliki daya pembeda yang buruk atau sangat buruk perlu direvisi atau diganti untuk meningkatkan efektivitas tes sebagai alat evaluasi. Implikasi dari hasil ini menunjukkan bahwa guru harus lebih selektif dalam memilih dan menyusun butir soal untuk memastikan bahwa setiap soal mampu membedakan dengan jelas antara siswa yang memiliki pemahaman yang baik dengan yang kurang.

### **4. Kesimpulan**

Penelitian ini menemukan bahwa validitas soal PAS genap mata pelajaran matematika kelas VIII di SMPS Advent Singkawang masih perlu ditingkatkan, karena sebagian besar butir soal tidak valid. Reliabilitas soal berada dalam kategori cukup, namun masih memerlukan perbaikan untuk mencapai kategori yang lebih tinggi. Indeks kesukaran menunjukkan distribusi soal yang seimbang, sementara daya pembeda menunjukkan bahwa perlu ada peningkatan dalam penyusunan soal untuk memastikan daya pembeda baik. Penelitian ini memberikan gambaran penting tentang kualitas soal PAS dan memberikan dasar bagi perbaikan ke depan untuk meningkatkan kualitas penilaian di SMPS Advent Singkawang.

### **5. Ucapan Terima kasih**

Peneliti mengucapkan terima kasih kepada seluruh pihak yang terlibat dalam penelitian ini terutama kepada para dosen pembimbing atas masukan dan arahan yang begitu berharga. Selain itu, peneliti juga berterima kasih kepada pihak SMPS Advent Singkawang atas izin dan bantuan dalam pengumpulan data untuk penelitian ini.

## **6. Daftar Pustaka**

- Anshari, M. I., Nasution, R., Irsyad, M., Alifa, A. Z., & Zuhriyah, I. A. (2024). Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI. *EDUKATIF: JURNAL ILMU PENDIDIKAN*, 6(1), 964-975. DOI: <https://doi.org/10.31004/edukatif.v6i1.5931>
- Arifin, Z. (2017). *Evaluasi Pembelajaran*. Bandung: PT. Remaja Rosdakarya.
- Elan, E., Sumardi, S., & Juandi, A. S. (2022). Penyusunan Instrumen Penelitian Tindakan Kelas dalam Upaya Peningkatan Keterampilan Sosial. *Jurnal PAUD Agapedia*, 6(1), 91-98. DOI: <https://doi.org/10.17509/jpa.v6i1.51339>
- Hutasuhut, S. H. (2022). Peranan Statistika dalam Penelitian Pendidikan Matematika. *MES: Journal of Mathematics Education and Science*, 7(2), 60-69. DOI: <https://doi.org/10.30743/mes.v7i2.5186>
- Mayasari, E. (2023). Instrument Tes Sebagai Alat Evaluasi (Analisis Soal, Indeks Kesukaran, Daya Pembeda Dan Fungsi Distractor). *Jurnal Pendidikan dan Pengajaran*, 2(1), 2964-2272.
- Ndiung, S., & Jediut, M. (2020). Pengembangan Instrumen Tes Hasil Belajar matematika peserta didik sekolah dasar berorientasi pada Berpikir Tingkat Tinggi. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), 94. DOI: [10.25273/pe.v10i1.6274](https://doi.org/10.25273/pe.v10i1.6274)
- Novia, T., Wardani, A., Canda, C., Nurdi, N., & Nurmasyitah, N. (2020). Analisis Validitas dan Reliabilitas Butir Soal UTS Fisika Kelas X SMA Swasta Muhammadiyah 4 Langsa. *GRAVITASI: Jurnal Pendidikan Fisika dan Sains*, 3(01), 19-22. DOI: <https://doi.org/10.33059/gravitasi.jpfs.v3i01.2256>
- Rais, F., Karnia, N., & Amilosa, P. (2022). Analisis Validitas Tes Mata Pelajaran Al Qur'an Hadis di SMP Muhamamdiyah Boarding School Yogyakarta. *HAWARI : Jurnal Pendidikan Agama Dan Keagamaan Islam*, 3(1). DOI: <https://doi.org/10.35706/hw.v3i1.6791>
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas dan Reliabilitas. *Journal on Education*, 6(2), 10967-10975. DOI: <https://doi.org/10.31004/joe.v6i2.4885>
- Sardia, Y., Fitriati, F., & Salmina, M. (2021). Analisis Kualitas Butir Soal Try Out Bidang Studi Matematika SMA N 1 Simeulue Tengah. *Jurnal Ilmiah Mahasiswa Pendidikan*, 2(1).
- Sudirman, Kondolayuk, M, L., Sriwahyuningrum, A., Cahaya, I. M. E., Astuti, N. L., Setiawan, J., Tandirerung, W. Y., Rahmi., S., Nusantara, D. O., Indrawati, F., Fitriya, N. L., Aziza, N., Kurniawati, N., Wardhana, A., Hasanah, T. (2023) *Metodologi Penelitian 1*. Edited By S. Haryanti. Bandung: Media Sains Indonesia.
- Sugiono, S., Noerdjanah, N., & Wahyu, A. (2020). Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation. *Jurnal Keterampilan Fisik*, 5(1), 55-61. DOI: <https://doi.org/10.37341/jkf.v5i1.167>

- Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sugiyono. (2020). *Metode Penelitian Kualitatif Untuk Penelitian Yang Bersifat: Eksploratif, Enterpretif Dan Konstruktif*. Edited By Y. Suryandari. Bandung: Alfabeta.
- Zellatifanny, C. M., & Mudjiyanto, B. (2018). Tipe Penelitian Deskripsi dalam Ilmu Komunikasi. *Diakom: Jurnal Media Dan Komunikasi*, 1(2), 83-90. DOI: [10.17933/diakom.v1i2.20](https://doi.org/10.17933/diakom.v1i2.20)