

Alhamdulillah, butir pengembangan instrumen penilaian artikel jurnal ilmiah dikatakan valid oleh para rater

Syarief Fajaruddin ^{1,2 *}, Heri Retnawati ², Tommy Tanu Wijaya ³, Syahri Ramadhan ², Yuli Prihatni ¹

¹ Universitas Sarjanawiyata Tamansiswa. Jl. Kusumanegara No.157, Yogyakarta, 55165, Indonesia

² Universitas Negeri Yogyakarta. Jalan Colombo No. 1, Yogyakarta, 55281, Indonesia

³ Beijing Normal University. 19 Xinhai Ave, Beitaipingzhuang, Haidian District, Beijing, 100875, China

* Corresponding Author. E-mail: syarief.fa@outlook.com; heri_retnawati@uny.ac.id; 202139130001@mail.bnu.edu.cn; syahri.ramadan1453@gmail.com; yuliku7781@gmail.com

Received: 22 September 2021; Revised: 27 November 2021; Accepted: 30 December 2021

Abstrak: Penelitian ini bertujuan untuk mengetahui kelayakan instrumen penilaian artikel ilmiah yang telah dikembangkan oleh peneliti. Instrumen penilaian artikel ilmiah ini dikembangkan berdasarkan literatur dengan indikator organisasi artikel, format penulisan, komponen penting artikel, dan grammatical structure. Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif. Butir dianalisis menggunakan perhitungan formula Aiken's V. Data validitas isi diperoleh dari sebelas ahli dari berbagai bidang ilmu yang berpengalaman dalam penulisan artikel ilmiah. Hasil validasi ahli ditemukan dari 56 butir yang dibuat dinyatakan valid dengan indeks Aikens V > 0,75, dengan rincian 5 butir memiliki koefisien validitas sedang, dan 51 butir memiliki koefisien yang tinggi. Hal ini menunjukkan bahwa semua butir instrumen memiliki validitas isi yang baik dan layak digunakan.

Kata Kunci: Artikel Ilmiah, Aiken, Validitas Butir, Penulisan Artikel, Publikasi

Alhamdulillah, the points of developing scientific journal article assessment instruments are said to be valid by raters

Abstract: This study aims to determine the feasibility of the scientific article assessment instrument that has been developed by the researcher. This scientific article assessment instrument was developed based on the literature with indicators of article organization, writing format, important components of the article, and grammatical structure. This study uses a quantitative approach with a descriptive method. The items were analyzed using Aiken's V formula calculations. Content validity data were obtained from eleven experts from various fields of science who are experienced in writing scientific articles. The results of expert validation were found that 56 items made were declared valid with an Aikens V index > 0.75, with details of 5 items having a moderate validity coefficient, and 51 items having a high coefficient. This indicates that all items of the instrument have good content validity and are suitable for use.

Keywords: Scientific Article, Aiken, Item Validity, Article Writing, Publication

How to Cite: Fajaruddin, S., Retnawati, H., Wijaya, T. T., Ramadhan, S., & Prihatni, Y. (2021). Alhamdulillah, butir pengembangan instrumen penilaian artikel jurnal ilmiah dikatakan valid oleh para rater. *Measurement In Educational Research (Meter)*, 1(2), 89-96. doi:<http://dx.doi.org/10.33292/meter.v1i2.156>



PENDAHULUAN

Artikel ilmiah adalah sarana komunikasi ilmiah yang paling populer di kalangan para ilmuwan di berbagai bidang ilmu. Komunikasi ilmiah merupakan proses komunikasi antara para ilmuwan dengan tujuan menyebarkan hasil penelitian maupun pengetahuannya dalam bentuk tulisan dalam majalah atau jurnal (Miswan, 2002). Sebagai sarana komunikasi ilmiah, penulisan artikel yang berkualitas akan menjadikan sebuah publikasi yang sukses dan menarik bagi pembaca global (Hong, 2014). Akan tetapi penulisan artikel ilmiah seringkali merupakan tugas yang sulit dan berat bagi banyak peneliti muda (Turbek et al., 2016). Banyak peneliti muda merasa kesulitan untuk menghasilkan publikasi ilmiah untuk meningkatkan prospek karier mereka, mengajukan permintaan dana hibah, atau memenuhi persyaratan bagi kualifikasi universi-

tas seperti gelar Master atau disertasi Doktoral (Ecarnot et al., 2015). Selain itu, bagi yang tertarik sebagai reviewer, reviewer pemula juga kurang memiliki pengetahuan yang memadai tentang peran mereka sebagai peer-reviewed dan keterampilan yang dibutuhkan untuk melakukan review yang komprehensif dan adil (Azer et al., 2012). Apabila proses penulisan dilakukan dengan metode dan proses peer-reviewed yang baik, maka akan dihasilkan sebuah publikasi yang berkualitas dan valid.

Dalam penelitian di bidang pendidikan, istilah valid tidak bisa dilepaskan dari butir-butir instrumen, karena kevalidan butir instrumen adalah penentu ketepatan dalam proses pengukuran objek yang diteliti (Sugiharni & Setiasih, 2018). Butir-butir instrumen dapat dikatakan baik apabila sudah valid, karena Interpretasi sebuah hasil penelitian akan bermakna ketika disajikan dengan jelas masalah validitas penelitian (Cor, 2016). Para ahli juga mengatakan bahwa validitas suatu alat ukur adalah sejauh mana alat ukur tersebut mampu mengukur apa yang seharusnya diukur (Allen & Yen, 1979, 2001; Kerlinger & Lee, 2000; Nunnally, 1981; Retnawati, 2016).

Butir-butir instrumen merupakan bahan dasar yang secara bersama-sama membentuk sebuah instrumen, sehingga kualitas butir-butir instrumen akan menentukan kualitas instrumen secara keseluruhan (Muzaffar, 2016). Kualitas butir instrumen wajib mendapatkan perhatian yang tinggi dalam proses penulisan serta pengembangannya, walaupun kualitas butir tidak secara eksplisit disebutkan dalam penilaian validitas. Butir-butir yang berkualitas memerlukan perhatian, tahapan, proses, dan prosedur yang harus dilalui dengan pilihan dan kesulitan yang harus dipertimbangkan untuk sampai pada pilihan yang tepat. Apabila semua hal tersebut telah dilalui dengan benar dan menghasilkan butir-butir berkualitas, maka tepatlah penilaian validitas akan secara langsung fokus kepada kesimpulan, penafsiran dan bahkan tindakan yang diambil berdasarkan hasil dan bukti yang diperoleh dari sebuah tes dan studi validasi (Muzaffar, 2016).

Dalam sebuah penelitian evaluasi, instrumen yang valid merupakan sebuah kebutuhan dasar sebagai alat ukur untuk mendapatkan gambaran tingkat keefektifan dari objek yang dievaluasi (Sugiharni & Setiasih, 2018). Oleh karena itu, diperlukan suatu proses perhitungan yang tepat dalam menentukan kevalidan instrumen dengan menggunakan beberapa formula yang diusulkan oleh para ahli. Lawshe (1975) menyarankan menggunakan *Content Validity Ratio* (CVR) untuk mengukur derajat kesepakatan para ahli dari satu item dan yang dapat mengekspresikan tingkat validitas konten melalui indikator tunggal yang berkisar dari -1 sampai 1. Aiken (1980, 1985) mengusulkan koefisien validitas isi dan reliabilitas koefisien homogenitas yang dapat digunakan untuk mengukur peringkat validitas setiap item (V value). Martuza (1977) mengembangkan *Content Validity Index* (CVI) yang banyak digunakan pada penelitian keperawatan. Sedangkan Landis dan Koch (1977) memberikan pedoman penilaian menggunakan *Interrater reliability* (*Kappa Statistic*), yaitu ukuran yang digunakan untuk menguji kesepakatan antara dua orang (penilai/pengamat) pada variabel kategoris.

Beberapa penelitian dalam lima tahun terakhir ini mengenai pengukuran validitas butir instrumen yang telah dilakukan, antara lain pengukuran validitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM (Bashooir & Supahar, 2018). Penelitian ini menggunakan menggunakan 5 kategori rating dan 7 rater dengan mendapatkan skor Aiken di atas 0,75. Kurniawati (2021) menganalisis instrumen tes berpikir kritis IPS kelas V SD Kota Yogyakarta menggunakan 4 rater rincian 7 butir memiliki indeks Aiken sebesar 0,75 dan 33 butir > 0,80. Rusijono et al. (2020) menguji instrumen penilaian kompetensi sosial guru di era revolusi industri keempat dengan menggunakan 7 rater dengan hasil memiliki indeks Aiken di atas 0,60 (valid). Pandawa dan Ridwan (2021) mengukur validitas instrumen standar kompetensi kebencanaan menggunakan 5 rater dengan hasil dengan mendapatkan skor Aiken di atas 0,75. Berdasarkan hal tersebut, menurut Aiken (1985) semakin banyak rater, semakin kecil nilai V yang disyaratkan. Semakin banyak pakar yang menelaah instrumen maka akan semakin baik kualitas instrumen dalam aspek validitas isi (Bashooir & Supahar, 2018).

Terkait penelitian mengenai pengembangan instrumen penilaian artikel ilmiah, Clabough dan Clabough (2016) telah mengembangkan rubrik untuk mengevaluasi makalah ilmiah dan menilai berbagai metode penggunaan rubrik di kelas pengantar biologi berbasis inkuiri. Penelitian ini dilakukan untuk menanamkan penulisan ilmiah ke dalam kelas sains berbasis inkuiri, dan memberikan jalan yang jelas untuk meningkatkan komunikasi dan keterampilan menulis ilmiah. Selain itu Supriyadi (2017) mengembangkan instrumen evaluasi pembelajaran menulis ilmiah dengan pendekatan konstruktivisme. Akan tetapi kedua penelitian tersebut merupakan penelitian pengembangan produk yang tidak diketahui validitasnya. Berdasarkan

pemikiran dan pertimbangan permasalahan yang dihadapi, peneliti memandang perlu mengembangkan instrumen penilaian artikel ilmiah. Instrumen yang dikembangkan tersebut perlu diuji validitas butirnya. Sehingga dari penelitian ini diharapkan instrumen penilaian artikel ilmiah yang dikembangkan oleh peneliti dapat diketahui validitas tiap butir instrumennya, sehingga dapat digunakan dalam penilaian artikel ilmiah.

METODE

Penelitian ini menggunakan pendekatan kuantitatif karena hasil yang diharapkan dalam penelitian ini adalah menunjukkan proses perhitungan validasi butir instrumen (Sugiharni & Setiasih, 2018). Metode penelitian yang digunakan adalah metode deskriptif.

Perancangan instrumen dan validasi, dilakukan dalam tiga tahap: (a) tinjauan pustaka, (b) perancangan draf pertama instrumen, (c) tinjauan kualitatif dan kuantitatif rater terhadap instrumen. Pada tahap pertama dilakukan review terhadap database sebagai berikut: Web of Science (WOS), Scopus, dan Google Scholar. Kata kunci pencarian adalah "*quality article*", "*structure article*", "instrumen penilaian artikel". Setelah pemilihan artikel dengan instrumen observasional, peneliti meninjau komponen, indikator, dan sub-indikatornya. Pada tahap kedua, draf indikator dan subindikator dibuat dari literatur ilmiah terkait. Instrumen akhir terdiri dari dua komponen, 4 indikator, dan 22 sub-indikator. Pada tahap ketiga, Validitas butir isi ditetapkan oleh 11 orang rater yang berasal dari 10 perguruan tinggi yang memiliki pendidikan S3. Rater ini merupakan orang yang berpengalaman dalam penulisan artikel dengan rata-rata h-index adalah 9.

Para ahli melakukan evaluasi kuantitatif (skala 1-5) dan kualitatif dengan memberikan saran untuk perbaikan instrumen. Validitas isi dihitung dengan koefisien V Aikens (Aiken, 1985). Untuk menghitung koefisien validitas isi yang didasarkan pada hasil penilaian dari para ahli sebanyak n orang terhadap suatu butir dari segi sejauhmana butir tersebut mewakili konstruk yang diukur. Penilaian dilakukan dengan cara memberikan angka antara 1 (tidak relevan) sampai dengan 5 (sangat relevan). Bila terdapat sebanyak n ahli yang menilai sebuah butir melalui rating (r) dengan pilihan 1 (tidak relevan) sampai dengan 5 (sangat relevan), berarti kategori tertinggi adalah 5 dan kategori terendah (l) adalah 1, maka dapat ditulis dengan pada Persamaan Rumus 1.

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

(Aiken, 1980, 1985; Kumaidi, 2014; Retnawati, 2016)

Persamaan Rumus 1 tersebut diketahui bahwa s merupakan skor yang ditetapkan ahli dikurangi skor terendah dalam kategori yang dipakai, dan c adalah banyaknya kategori yang dapat dipilih oleh ahli.

$$s = r - l \dots\dots\dots 2)$$

(Retnawati, 2016)

Dengan menggunakan Persamaan Rumus 1, peneliti mengkategorikan hasil penghitungan para rater dengan indeks kesepakatan ahli mengenai validitas. Indeks kesepakatan apabila kurang dari 0,4, maka dikatakan validitasnya rendah, diantara 0,4 - 0,8 dikatakan validitasnya sedang, dan jika lebih dari 0,8 dikatakan tinggi (Retnawati, 2016).

Apabila butir instrumen tergolong pada kategori tinggi, maka butir tersebut digunakan. Jika butir instrumen evaluasi tergolong pada kategori rendah, maka butir instrumen tersebut dibuang/tidak digunakan. Apabila butir pada kategori sedang, maka perlu dilakukan kajian ulang pada butir instrumen tersebut.

HASIL DAN PEMBAHASAN

Hasil

Butir instrumen yang dikembangkan akan digunakan untuk mengukur kualitas artikel yang telah dibuat oleh penulis. Hasil validitas butir instrumen pengembangan instrumen penilaian artikel yang menggunakan perhitungan formula Aiken dapat ditunjukkan pada Tabel 1.

Nilai s_1 , s_2 , s_3 , sampai dengan s_{11} yang hasilnya ditunjukkan pada Tabel 1 didapatkan dari penilaian 11 rater dengan menggunakan Persamaan Rumus 2. Berdasarkan Persamaan Rumus 2, maka s_1 untuk butir ke-1 didapatkan dari penilaian rater 1, yaitu $= 5-1 = 4$, rater 2 (s_2) untuk butir ke-1 $= 5-1 = 4$; s_3 untuk butir ke-1 $= 3-1 = 2$; dan seterusnya sampai rater ke sebelas dan sampai butir ke-56 dengan perhitungan yang

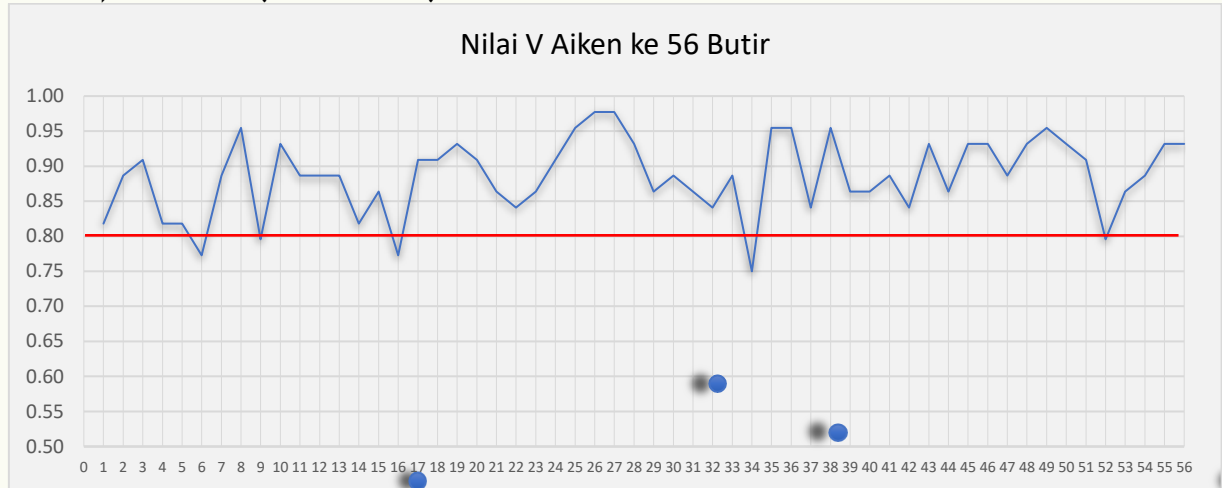
sama menggunakan Persamaan Rumus 2. Sedangkan untuk memperoleh nilai V yang hasilnya telah ditunjukkan pada Tabel 1 diperoleh menggunakan Persamaan Rumus 1.

Tabel 1. Penghitungan indeks kesepakatan rater mengenai validitas

Butir	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	$\sum s$	V	Validitas
1.	4	4	2	4	4	3	4	4	1	2	4	36	0.818	Tinggi
2.	3	4	3	4	4	4	4	2	4	3	4	39	0.886	Tinggi
3.	1	4	4	4	4	3	4	4	4	4	4	40	0.909	Tinggi
4.	1	3	3	3	4	3	4	3	4	4	4	36	0.818	Tinggi
5.	1	4	3	3	4	3	4	3	4	4	3	36	0.818	Tinggi
6.	0	4	3	4	3	3	3	3	4	4	3	34	0.773	Sedang
7.	4	4	3	4	4	4	4	4	1	4	3	39	0.886	Tinggi
8.	4	4	3	4	4	4	4	4	4	4	3	42	0.955	Tinggi
9.	1	4	3	3	4	4	2	2	4	4	4	35	0.795	Sedang
10.	3	3	3	4	4	4	4	4	4	4	4	41	0.932	Tinggi
11.	4	3	3	4	4	3	4	4	3	3	4	39	0.886	Tinggi
12.	3	4	3	4	4	4	4	4	4	2	3	39	0.886	Tinggi
13.	3	4	3	4	4	4	4	4	3	3	3	39	0.886	Tinggi
14.	1	4	3	4	4	4	4	2	4	4	2	36	0.818	Tinggi
15.	3	4	3	4	4	4	4	2	4	4	2	38	0.864	Tinggi
16.	3	4	3	4	3	3	4	3	2	3	2	34	0.773	Sedang
17.	4	4	3	4	4	4	4	4	4	3	2	40	0.909	Tinggi
18.	4	4	3	4	4	3	4	4	4	3	3	40	0.909	Tinggi
19.	4	4	3	4	4	4	4	4	4	3	3	41	0.932	Tinggi
20.	4	4	3	4	4	4	4	4	3	3	3	40	0.909	Tinggi
21.	4	4	3	4	4	3	4	1	4	4	3	38	0.864	Tinggi
22.	1	3	3	4	4	3	4	4	4	4	3	37	0.841	Tinggi
23.	2	4	3	4	4	4	3	3	4	3	4	38	0.864	Tinggi
24.	4	4	3	4	4	3	4	3	3	4	4	40	0.909	Tinggi
25.	4	4	3	4	4	3	4	4	4	4	4	42	0.955	Tinggi
26.	4	4	3	4	4	4	4	4	4	4	4	43	0.977	Tinggi
27.	4	4	3	4	4	4	4	4	4	4	4	43	0.977	Tinggi
28.	3	4	3	4	4	3	4	4	4	4	4	41	0.932	Tinggi
29.	3	3	3	4	4	3	4	2	4	4	4	38	0.864	Tinggi
30.	4	4	3	4	4	3	4	2	3	4	4	39	0.886	Tinggi
31.	4	4	3	4	4	3	4	1	4	3	4	38	0.864	Tinggi
32.	0	4	3	4	4	3	4	3	4	4	4	37	0.841	Tinggi
33.	1	4	3	4	4	3	4	4	4	4	4	39	0.886	Tinggi
34.	1	4	3	3	4	3	2	1	4	4	4	33	0.75	Sedang
35.	4	4	3	4	4	3	4	4	4	4	4	42	0.955	Tinggi
36.	4	4	3	4	4	3	4	4	4	4	4	42	0.955	Tinggi
37.	0	3	3	4	4	3	4	4	4	4	4	37	0.841	Tinggi
38.	4	4	3	4	4	3	4	4	4	4	4	42	0.955	Tinggi
39.	1	4	3	4	4	3	4	4	4	4	3	38	0.864	Tinggi
40.	1	4	3	4	4	3	4	4	4	4	3	38	0.864	Tinggi
41.	4	4	3	4	3	3	4	3	4	4	3	39	0.886	Tinggi
42.	3	4	3	4	4	3	4	1	4	4	3	37	0.841	Tinggi
43.	3	4	3	4	4	3	4	4	4	4	4	41	0.932	Tinggi
44.	0	4	3	4	4	3	4	4	4	4	4	38	0.864	Tinggi
45.	3	4	3	4	4	3	4	4	4	4	4	41	0.932	Tinggi
46.	3	4	3	4	4	3	4	4	4	4	4	41	0.932	Tinggi
47.	1	4	3	4	4	3	4	4	4	4	4	39	0.886	Tinggi
48.	4	4	3	3	4	3	4	4	4	4	4	41	0.932	Tinggi
49.	4	4	3	4	4	3	4	4	4	4	4	42	0.955	Tinggi
50.	3	4	3	4	4	3	4	4	4	4	4	41	0.932	Tinggi
51.	4	3	3	4	4	3	4	4	4	4	3	40	0.909	Tinggi
52.	1	3	3	4	4	4	3	4	4	2	3	35	0.795	Sedang
53.	3	3	3	4	4	3	4	3	4	4	3	38	0.864	Tinggi
54.	4	4	3	4	4	3	4	4	4	2	3	39	0.886	Tinggi
55.	4	4	3	4	4	4	4	4	4	3	3	41	0.932	Tinggi
56.	4	4	3	4	4	4	4	4	4	3	3	41	0.932	Tinggi

Penentuan kategori validitas tiap butir instrumen yang ditunjukkan pada Tabel 1 dilakukan dengan membandingkan nilai V tiap butir instrumen dengan indeks kesepakatan ahli mengenai validitas. Kategori tinggi apabila nilai V-nya lebih dari 0,8, kategori sedang apabila nilainya diantara 0,4 - 0,8, dan kategori kurang apabila nilainya kurang dari 0,4.

Berdasarkan penilaian kesebelas rater yang ditunjukkan pada Tabel 1, diketahui bahwa Sebagian besar butir instrument penilaian artikel yang dikembangkan dikatakan memiliki validitas tinggi (91%), sedangkan 9% memiliki validitas sedang. Butir instrumen yang memiliki validitas sedang yaitu Butir 6, Butir 9, Butir 16, Butir 34, dan butir 52 (lihat Gambar 1).



Gambar 1. Nilai V Aiken ke 56 Butir

Pembahasan

Mengembangkan dan memvalidasi instrumen penilaian artikel ilmiah ini akan berkontribusi pada alat inovatif yang digunakan untuk memvalidasi artikel. Hal ini akan memberikan manfaat bagi para peneliti muda dan para pengelola jurnal dalam mempersiapkan publikasinya. Dari instrumen ini diharapkan penulis semakin mudah menyusun makalah dan pengelola jurnal dapat mengembangkan kualitas artikelnya dengan baik.

Penggunaan banyak rater dalam validasi butir instrument ini menguatkan penemuan sebelumnya yang menyatakan bahwa semakin banyak rater yang menelaah instrument, maka akan semakin baik kualitas instrumen dalam aspek validitas isi (Bashooir & Supahar, 2018). Peneliti menggunakan sebelas rater dari berbagai bidang ilmu dengan pendidikan rata-rata S3. Para rater ini merupakan peneliti muda dengan banyak pengalaman menulis pada berbagai jurnal nasional maupun internasional.

Butir instrument yang divalidasi terdiri atas dua komponen, empat indikator, dua puluh dua sub-indikator, dengan total 56 butir. Semua butir dikatakan valid karena mendapatkan skor Aiken di atas 0,75. Akan tetapi dari 56 butir tersebut ada 5 butir yang memiliki validitas sedang yaitu Butir 6, Butir 9, Butir 16, Butir 34, dan butir 52.

Pada butir 6 disebutkan bahwa Judul memberikan penjelasan, pembaca dapat menyampaikan apa yang dilakukan oleh peneliti tanpa membaca seluruh makalahnya. Hal ini bertentangan dengan pernyataan butir 3, bahwa judul maksimal 14 kata. Sehingga beberapa rater meminta mengganti butir ini dan menyarankan pernyataan pada butir ini menjadi judul mudah dipahami, menarik, dan unik. Hal tersebut sesuai dengan pernyataan (Cuschieri et al., 2019; Hong, 2014) bahwa judul artikel harus menarik bagi pembaca, spesifik, dan pendek, tetapi komprehensif dan cukup deskriptif (Cuschieri et al., 2019; Hong, 2014). Judul umumnya tidak boleh melebihi 150 karakter atau 12-16 kata, meskipun ini harus disesuaikan dengan instruksi dari jurnal tertentu (Prem, 2017).

Butir 9 menyebutkan bahwa panjang abstrak antara 200-350 kata. Akan tetapi para rater menyarankan Panjang abstrak antara 200-250 kata. Hal ini sejalan dengan pernyataan Cuschieri et al. (2019) yang mengatakan bahwa abstrak abstrak harus ringkas dengan batas kata, sekitar 250 kata. Akan tetapi Sauaia et al. (2014) menyarankan 300 kata.

Pernyataan mengandung semua informasi yang diperlukan pada Butir 16 terasa ambigu. Pada pernyataan ini ada kata yang hilang, yaitu kata Tabel/Gambar. Sebaiknya pernyataan ini menjelaskan informasi mengenai tabel dan gambar yang dituangkan pada artikel. Tabel dan Gambar mewakili data dalam format yang mudah dimengerti, dan paling baik digunakan ketika kumpulan data yang berbeda dibandingkan satu sama lain (Pai et al., 2017).

Pernyataan bagian hasil dipresentasikan dalam bentuk paragraph pada Butir 34 memerlukan revisi. Sebagian besar rater menyarankan bahwa bagian hasil dipresentasikan dalam bentuk paragraf dan hasil dipresentasikan dalam bentuk tabel dan atau diagram. Sebab Tabel dan Gambar biasanya digunakan di bagian hasil (Pai et al., 2017). Hal ini sesuai dengan pernyataan Ramesh dan Ananthakrishnan (2017) bahwa bagian hasil dimulai dengan paragraf singkat dengan setiap paragraf mengacu pada satu tabel atau gambar atau satu parameter studi.

Butir 52 pada instrument ini mendapatkan skor 0,795 (Kategori Sedang). Rater menilai similarity maksimal sebuah artikel ilmiah adalah 15%-20%. Menurut Memon (2020), similarity text 5% atau kurang dapat diterima oleh beberapa editor jurnal, sementara itu apabila lebih dari 20% naskah tersebut perlu diawasi. Pendapat lain mengatakan bahwa editor jurnal cenderung menolak naskah jika kesamaan teksnya di atas 10% (Peh & Arokiasamy, 2008; Swaan, 2010). Akan tetapi beberapa institusi menganggap kesamaan teks kurang dari 20% sebagai hal yang dapat diterima, bahkan pedoman pada University Grants Commission of India memungkinkan kesamaan hingga 10% sebagai dapat diterima atau minor (Level 0).

Studi kami memiliki implikasi untuk penelitian dan praktik publikasi. Sepengetahuan kami, ini adalah studi validasi pertama mengenai instrumen penilaian artikel jurnal ilmiah. Temuan ini bisa menjadi alat dalam penilaian artikel dan sebagai panduan mempersiapkan artikel untuk publikasi. Di masa depan, harapannya instrumen ini dapat dikembangkan lebih baik lagi.

SIMPULAN

Studi ini membangun dan memvalidasi instrument penilaian artikel. Peneliti telah mengembangkan 56 butir instrumen penilaian artikel ilmiah yang didasarkan pada dua komponen, empat indikator, dan dua puluh dua sub-indikator. Setelah dilakukan uji validitas butir instrumen dengan koefisien V Aikens diperoleh 5 butir dengan nilai V antara 0,75-0,80 dan 51 butir memiliki koefisien di atas 0,80. Berdasarkan hal tersebut maka semua butir instrumen valid, dan dapat digunakan.

DAFTAR PUSTAKA

- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40(4), 955-959. <https://doi.org/10.1177/001316448004000419>
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131-142. <https://doi.org/10.1177/0013164485451012>
- Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Brooks/Cole Publishing Company.
- Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Waveland Press, Inc.
- Azer, S. A., Ramani, S., & Peterson, R. (2012). Becoming a peer reviewer to medical education journals. *Medical Teacher*, 34(9), 698-704. <https://doi.org/10.3109/0142159X.2012.687488>
- Bashoor, K., & Supahar, S. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(2), 219-230. <https://doi.org/10.21831/pep.v22i2.19590>
- Clabough, E. B. D., & Clabough, S. W. (2016). Using rubrics as a scientific writing instructional method in early stage undergraduate neuroscience study. *Journal of Undergraduate Neuroscience Education: JUNE: A Publication of FUN, Faculty for Undergraduate Neuroscience*, 15(1), A85-A93.
- Cor, M. K. (2016). Trust me, it is valid: Research validity in pharmacy education research. *Currents in Pharmacy Teaching and Learning*, 8(3), 391-400. <https://doi.org/10.1016/j.cptl.2016.02.014>
- Cuschieri, S., Grech, V., & Savona-Ventura, C. (2019). WASP (Write a Scientific Paper): Structuring a scientific paper. *Early Human Development*, 128, 114-117. <https://doi.org/10.1016/j.earlhumdev.2018.09.011>

- Ecarnot, F., Seronde, M.-F., Chopard, R., Schiele, F., & Meneveau, N. (2015). Writing a scientific article: A step-by-step guide for beginners. *European Geriatric Medicine*, 4(6), 573–579. <https://doi.org/10.1016/j.eurger.2015.08.005>
- Hong, S.-T. (2014). Ten tips for authors of scientific articles. *Journal of Korean Medical Science*, 29(8), 1035. <https://doi.org/10.3346/jkms.2014.29.8.1035>
- Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. Wadsworth Publishing.
- Kumaidi, K. (2014). Validitas dan pemvalidasian instrumen penilaian karakter. *Seminar Psikometri Fakultas Psikologi Universitas Muhammadiyah Surakarta*.
- Kurniawati, K. (2021). Analisis validitas isi instrumen tes berpikir kritis IPS kelas V SD Kota Yogyakarta. *Pelita: Jurnal Penelitian Dan Karya Ilmiah*, 21(1), 130–140. <https://doi.org/10.33592/pelita.v21i1.1396>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Martuza, V. R. (1977). *Applying norm-referenced and criterion-referenced measurement in education*. Allyn & Bacon.
- Memon, A. R. (2020). Similarity and plagiarism in scholarly journal submissions: Bringing clarity to the concept for authors, reviewers and editors. *Journal of Korean Medical Science*, 35(27). <https://doi.org/10.3346/jkms.2020.35.e217>
- Miswan, M. (2002). Jurnal elektronik sebagai sarana komunikasi ilmiah. *AL-MAKTABAH*, 4(1). <https://doi.org/10.15408/almaktabah.v4i1.1643>
- Muzaffar, A. (2016). Validitas tes dan kualitas butir soal. □□□□□□□□ (LISANUNA): *Jurnal Ilmu Bahasa Arab Dan Pembelajarannya*, 5(1), 128–143. <https://doi.org/10.22373/l.v5i1.859>
- Nunnally, J. C. (1981). *Psychometric theory*. Tata Mcgraw Hill Publishing Company Limited.
- Pai, D., Chua, S. K., & Sood, S. (2017). Figures, tables and supporting material. In *Writing and Publishing a Scientific Research Paper* (pp. 95–105). Springer. https://doi.org/10.1007/978-981-10-4720-6_10
- Pandawa, R. M., & Ridwan, A. (2021). Analysis of the instrument content validity using the aiken index on disaster nursing competency assessment. *DEGRES*, 20(1), 298–308. <https://doi.org/10.1877/degres.v20i1.90>
- Peh, W. C. G., & Arokiasamy, J. (2008). Plagiarism: a joint statement from the Singapore Medical Journal and the Medical Journal of Malaysia. *Singapore Medical Journal*, 49(12), 965. <http://www.smj.org.sg/sites/default/files/4912/4912e1.pdf>
- Prem, S. S. (2017). Title. In *Writing and Publishing a Scientific Research Paper* (p. 17). Springer.
- Ramesh, R., & Ananthakrishnan, N. (2017). Results. In *Writing and Publishing a Scientific Research Paper* (pp. 61–74). Springer Singapore. https://doi.org/10.1007/978-981-10-4720-6_7
- Retnawati, H. (2016). *Validitas reliabilitas dan karakteristik butir*. Parama Publishing.
- Rusijono, R., Jaedun, A., Kartowagiran, B., Ahman, A., Laliyo, L. A. R., & Mam, S. (2020). Developing the teacher's social competency assessment instrument in the fourth industrial revolution era. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 24(2). <https://doi.org/10.21831/pep.v24i2.29482>
- Sauaia, A., Moore, E. E., Crebs, J. L., Maier, R. V., Hoyt, D. B., & Shackford, S. R. (2014). The anatomy of an article. *Journal of Trauma and Acute Care Surgery*, 76(5), 1322–1327. <https://doi.org/10.1097/TA.0000000000000210>
- Sugiharni, G. A. D., & Setiasih, N. W. (2018). Validasi butir instrumen evaluasi model Alkin menggunakan formula Aiken. *Prosiding Seminar Nasional Pendidikan Teknik Informatika, (September)*.
- Supriyadi, S. (2017). Developing scientific writing instruction evaluation instrument by using a constructivism approach. *Proceedings of the 1st Yogyakarta International Conference on Educational Management/Administration and Pedagogy (YICEMAP 2017)*. <https://doi.org/10.2991/yicemap-17.2017.13>

- Swaan, P. W. (2010). Publication ethics—A guide for submitting manuscripts to Pharmaceutical Research. *Pharmaceutical Research*, 27(9), 1757–1758. <https://doi.org/10.1007/s11095-010-0188-5>
- Turbek, S. P., Chock, T. M., Donahue, K., Havrilla, C. A., Oliverio, A. M., Polutchko, S. K., Shoemaker, L. G., & Vimercati, L. (2016). Scientific writing made easy: A step-by-step guide to undergraduate writing in the biological sciences. *The Bulletin of the Ecological Society of America*, 97(4), 417–426. <https://doi.org/10.1002/bes2.1258>

Conflict of Interest Statement: The Author(s) declares that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Copyright: ©Measurement in Educational Research. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Measurement in Educational Research is an open access and peer-reviewed journal published by Research and Social Study Institute, Indonesia

Open Access 