

Research Article



## Analisis Butir Soal Kemampuan *Tree Thinking* Pada Instrumen Uji Coba Materi Spermatophyta

(*Analysis of tree thinking ability problem ability in spermatophyta trial material instruments*)

Utami Khofifah\*, Jujun Ratnasari, Aa Juhanda

Universitas Muhammadiyah Sukabumi

Jl. R. Syamsudin, S.H. No. 50, Cikole, Kec. Cikole, Kota Sukabumi, 43113, Indonesia

\*Corresponding Author: [tamimada999@gmail.com](mailto:tamimada999@gmail.com)

Informasi Artikel	ABSTRACT
Submit: 11 – 06 – 2020 Diterima: 27 – 08 – 2020 Dipublikasikan: 10 – 09 – 2020	<p><i>This study aims to analyze the ability of Tree Thinking in students. This type of research is quantitative descriptive, the instrument used in the form of test items on Tree Thinking items on Spermatophyta material. This item is a multiple choice with 5 options. The results of this study are information in the form of item validity, distinguishing features, and level of difficulty of the questions. The results of the analysis show the quality of the items, valid at 95% while those that are not valid at 5%. Based on the ability to distinguish the questions, 10% is considered satisfactory, 50% is categorized as good and 40% is categorized as very good. Difficulties about the data obtained are 85% in the medium category, and 15% are in the difficult category. Data analysis shows that the item has a high validity while the distinguishing power is classified as good and the level of difficulty of the item is classified as moderate.</i></p> <p><b>Keywords:</b> <i>Item bullet, Tree Thinking, Problem validity, Power of differentiation, and Difficulty of matter</i></p>
Penerbit	ABSTRAK
Program Studi Pendidikan Biologi, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Jambi	<p>Penelitian ini bertujuan untuk menganalisis kemampuan <i>Tree Thinking</i> pada siswa. Jenis penelitian ini adalah deskriptif kuantitatif, instrumen yang digunakan berupa hasil tes butir soal <i>Tree Thinking</i> pada materi Spermatophyta. Butir soal ini berupa pilihan ganda dengan 5 opsi. Hasil dari penelitian ini adalah informasi berupa validitas butir soal, daya pembeda, dan tingkat kesukaran soal. Hasil analisis menunjukkan kualitas butir soal, yang valid sebesar 95% sedangkan yang tidak valid sebesar 5%. Berdasarkan daya pembeda soal menunjukkan 10% dikategorikan memuaskan, 50% dikategorikan baik dan 40% dikategorikan sangat baik. Kesukaran soal data yang di dapat yaitu 85% berada pada kategori sedang, dan 15% berada pada kategori sukar. Analisis data menunjukkan bahwa butir soal memiliki validitas tinggi sedangkan daya pembeda tergolong baik dan tingkat kesukaran soal tergolong sedang.</p> <p><b>Katakunci:</b> <i>Butir soal, Tree Thinking, Validitas soal, Daya pembeda, dan Kesukaran soal</i></p>



This BIODIK : Jurnal Ilmiah Pendidikan Biologi is licensed under a [CC BY-NC-SA \(Creative Commons Attribution-ShareAlike 4.0 International License\)](https://creativecommons.org/licenses/by-nc-sa/4.0/)

## PENDAHULUAN

Kegiatan pembelajaran yang dilakukan di sekolah salah satunya adalah penilaian pembelajaran. Penilaian pembelajaran tersebut menggunakan alat ukur yang tepat agar menghasilkan data yang lebih akurat. Bagi para guru maupun calon guru perlu mengetahui cara menggunakan penilaian, pertama pengumpulan informasi, penyusunan instrument penilaian, dan pengolahan hasil tes. Penilaian yang dilakukan. Dalam penggunaan instrumen sendiri harus berkualitas baik. Azmi (2019) cara memperoleh instrumen yang berkualitas adalah dengan mengukur validitas, reabilitas, daya pembeda dan indeks kesukaran terhadap soal yang disusun guru. Berdasarkan pendapat Azmi di atas, bahwa alat ukur yang menunjukkan hasil valid dan reliabel merupakan kriteria minimal suatu instrumen yang baik. Selain kriteria di atas suatu tes dianggap baik apabila daya pembeda, tingkat kesukaran dan pengecoh pada soal pilihan gandanya tergolong baik.

Alat pengukuran pertama yaitu validitas, menurut (Supandi & Farikhah, 2016) validitas digunakan untuk mengetahui valid tidaknya item tes. Instrument pengukuran dikatakan memiliki validitas tinggi apabila alat ukur tersebut menerapkan fungsi ukurnya sesuai dengan pengukuran tersebut. Kedua yaitu reabilitas, menurut Masidjo (1995: 208) reabilitas adalah taraf kemampuan tes dalam menunjukkan konsisten hasil pengukuran yang diperlihatkan dalam taraf ketepatan dan ketelitian hasil. Ketiga yaitu daya pembeda, daya pembeda adalah taraf jumlah jawaban benar siswa yang tergolong kelompok (pandai = *upper group*) berbeda dari siswa yang tergolong kelompok bawah (kurang pandai = *lower group*) untuk beberapa item. Keempat yaitu tingkat kesukaran, menurut Sulistyorini (2009: 176) tingkat kesulitan merupakan kemampuan siswa untuk menjawab soal dengan kriteria soal mudah, sedang dan sukar. Tingkat kesukaran dikatakan baik apabila suatu tes dikategorikan 25% mudah, 50% sedang, dan 25% sukar. Kelima yaitu analisis pengecoh, menurut Arikunto (2012: 234) pengecoh dapat berfungsi dengan baik apabila pengecoh tersebut mempunyai daya tarik bagi peserta tes yang kurang memahami materi. Pada penelitian ini, alat ukur yang akan digunakan yaitu validitas soal, daya pembeda, dan kesukaran soal.

Sama halnya dengan penelitian sebelumnya yang dilakukan oleh (Azmi, 2019), soal kemampuan *tree thinking* merupakan soal yang paling jarang muncul di buku pelajaran. Selain dari itu, soal *tree thinking* bukanlah soal yang biasa di buat oleh guru, maka perlu dikembangkan soal *tree thinking* yang valid sehingga siap dipakai oleh guru dan diuji cobakan pada siswa, sekurang-kurangnya dapat dijadikan acuan bagi guru dalam membuat soal yang mengasah kemampuan *tree thinking*. Terdapat dua kategori tes yang dapat digunakan yaitu tes objektif dan tes subjektif. Tes objektif memiliki bentuk tes penilaian jawaban atau tanggapan siswa tes sepenuhnya dilakukan secara objektif oleh pemeriksa. Penilaiannya sendiri di bantu menggunakan aplikasi tertentu. Tes yang digunakan berupa tes pilihan ganda yang terdiri dari 5 opsi pada materi Spermatophyta untuk menguji kemampuan *Tree Thinking*.

Abad ke-21 ini, kurikulum pada mata pelajaran biologi SMA menyertakan kompetensi dasar siswa untuk dapat memahami kladogram/pohon filogenik. Salah

satu kompetensi inti 4.8 “*Menyajikan data hasil pengamatan dan analisis fenetik dan filogenetik tumbuhan serta peran tumbuhan dalam kelangsungan hidup di bumi.*” (Permendikbud, 2016). Di Indonesia, hasil belajar siswa dilihat di akhir periode pembelajaran. Nilai tersebut didapatkan dari ujian ataupun tes yang diberikan oleh guru kepada siswanya dalam bentuk tes pembuatan cladogram. Pada penelitian ini konsep tumbuhan biji (Spermatophyta) dipilih sebagaimana yang diketahui bahwa Indonesia memiliki keanekaragaman flora yang tinggi. Sebab pada praktiknya tumbuhan yang diperlukan mudah diperoleh sehingga siswa dapat mengamati bentuk dan cirinya secara langsung (Sari, 2019).

Mempelajari klasifikasi tumbuhan dengan menggunakan pendekatan filogenetik tentunya akan lebih menarik perhatian siswa dibandingkan menggunakan metode ceramah. Menurut Rustaman (1991), pembelajaran klasifikasi tumbuhan seharusnya tidak menuntut siswa untuk menghafal ciri-ciri tumbuhan tanpa melakukan observasi secara langsung. Pengalaman belajar utama yang sebaiknya didapatkan siswa adalah pengalaman belajar secara langsung dengan tumbuhan di lingkungan sekitar agar pembelajaran siswa mengenai klasifikasi ini dapat terbentuk.

Analisis bentuk soal dapat diunakan sebagai alat evaluasi tahap awal. Butir soal yang dijadikan sebagai alat evaluasi harus mencakup seluruh kompetensi dasar ,seluruh kurikulum dan tercapainya standar kompetensi. Beberapa penyebab kesulitan belajar siswa pada waktu mengikuti ujian terutama pada tiga aspek, yaitu; (1) materi belum tuntas diajarkan, (2) materi sulit dipahami oleh siswa, dan (3) siswa belum bisa menjawab model soal yang di ujikan.

## **METODE PENELITIAN**

Penelitian ini dilakukan di salah satu SMA di kabupaten Sukabumi pada minggu pertama di bulan April 2020. Metode penelitian yang digunakan pada penelitian ini yaitu deskriptif kualitatif. Data penelitian yang digunakan merupakan soal penelitian pada materi Spermatophyta yang dilakukan di kelas XI IPA 1 tahun ajaran 2019/2020, dan lembar kerja siswa , dan silabus materi Spermathophyta. Teknik pengumpulsn data pada penelitian ini dilakukan dengan menggunakan metode daring. Uji validitas data terkait dengan analisis kualitatif butir soal penilaian pada materi Spermatophyta kelas XI IPA 1 di SMA di kabupaten Sukabumi tahun ajaran 2019/2020 yang erat hubungannya dengan kemampuan *Tree Thinking*.

*Tree thinking* merupakan salah satu kemampuan dalam memahami diagram, menginterpretasikan, dan menggunakannya sebagai suatu media Catley, Philips, & Novick, (2013). Kemampuan *tree thinking* merupakan kemampuan memahami dengan menggunakan gambar pohon filogenetik untuk membantu memudahkan siswa memahami materi Baum & Smith dikutip dalam Smith *et al.*, (2013). Dalam penggunaan *tree thinking* dilakukan dengan menggunakan dua metode pendekatan untuk merekonstruksi hubungan dari sebuah kelompok organisme biologi, yaitu fenetik dan filogenetik (Mahbubah,2017). Metode fenetik menitik beratkan persamaan karakter yang dimiliki oleh organisme untuk melihat kekerabatannya, sedangkan filogenetik di artikan sebagai model atau mempresentasikan hubungan

antar nenek moyang organisme, sekuen molekul atau keduanya (Brinkman & Leipe, 2001).

Sugiono (2012: 89) analisis data ialah proses mencari dan menyusun secara sistematis data yang didapat dari hasil wawancara, hasil observasi dan hasil dokumentasi, dengan cara mengorganisasikan data kedalam kategori, menjabarkan data ke dalam unit-unit, melakukan sintesa, menyusun pola, menentukan hal yang penting dan yang akan dipelajari, dan membuat kesimpulan sehingga mudah dipahami oleh diri sendiri maupun orang lain. Teknik pengumpulan data pada penelitian ini menggunakan instrument berupa lembar soal kemampuan *Tree Thinking*. Teknik pengambilan data ini dilakukan dengan menguji kemampuan *Tree Thinking* pada siswa.

## HASIL DAN PEMBAHASAN

Tes ini bertujuan untuk mengetahui kemampuan *tree thinking* pada siswa SMA kelas X tentang materi klasifikasi tumbuhan berbiji melalui instrumen *Tree Thinking Concept Inventory* (TTCI) menggunakan indikator kemampuan *tree thinking* nya mengacu pada Novick & Catley (2012), pada jenjang C2-C6 (memahami, mengaplikasikan, menganalisis, mengevaluasi, dan membuat) dalam taksonomi bloom revisi (domain kognitif) berupa soal pilihan ganda. Tes menggunakan indikator kemampuan *tree thinking* yang berkaitan dengan pemahaman dan alasan memakai cladogram berdasarkan Novick & Catley (2012). Tes diberikan pada akhir kegiatan pembelajaran pada materi tumbuhan berbiji, dengan tipe soal pilihan ganda, setiap soal terdapat empat *option* pilihan jawaban. Hasil analisis dengan menggunakan aplikasi ANATES Pilihan Ganda versi 4.0.9 soal pilihan ganda sebanyak 30 butir soal, pengumpulan data yang akan digunakan meliputi korelasi skor soal, daya pembeda, tingkat kesukaran. Rangkuman analisis butir soal pilihan ganda dengan memakai aplikasi ANATES versi 4.0.9 dapat dilihat dari tabel 1.

**Tabel 1. Rekapitulasi Analisis Butir Soal Pilihan Ganda**

No. Soal	Butir Asli	D. Pembeda (%)	T. Kesukaran	Validitas	Indeks	Kualitas Soal
1	1	0,70	Sedang	0.564	Valid	Sangat Baik
2	2	0,80	Sedang	0.515	Valid	Baik
3	3	0,60	Sedang	0.555	Valid	Sangat Baik
4	4	0,80	Sedang	0.668	Valid	Sangat Baik
5	5	0,90	Sedang	0.670	Valid	Sangat Baik
6	6	0,60	Sukar	0.579	Valid	Sangat Baik
7	7	0,90	Sedang	0.755	Valid	Sangat Baik
8	8	0,80	Sedang	0.639	Valid	Sangat Baik
9	9	0,60	Sukar	0.688	Valid	Sangat Baik
10	10	0,90	Sedang	0.735	Valid	Sangat Baik
11	11	0,60	Sedang	0.602	Valid	Sangat Baik
12	12	0,50	Sedang	0.449	Valid	Baik
13	13	0,90	Sedang	0.726	Valid	Sangat Baik
14	14	0,50	Sedang	0.531	Valid	Baik
15	15	0,70	Sedang	0.621	Valid	Sangat Baik
16	16	0,70	Sedang	0.572	Valid	Sangat Baik
17	17	0,50	Sedang	0.548	Valid	Baik

18	18	0,90	Sedang	0.668	Valid	Sangat Baik
19	19	0,30	Sukar	0.434	Valid	Baik
20	20	0,30	Sedang	0.219	-	-

Uji validitas ini digunakan untuk mengetahui valid atau tidaknya suatu item tes. Berdasarkan hasil analisis soal yang valid berjumlah 19 butir soal (95%) sementara soal invalid berjumlah 1 butir soal (5%). Berdasarkan data di atas, dapat disimpulkan soal penilaian pada materi Spermatophita untuk kelas IX IPA 1 masuk dalam golongan dengan kualitas soal yang sangat baik dilihat dari segi validitas soal.

Surapranata (2005: 50) validitas merupakan suatu konsep yang berkaitan dengan sejauh mana tes telah mengukur apa yang harus diukur. Perlunya mempertinggi validitas sebelum membuat butir-butir soal dilakukan beberapa langkah diantaranya seperti yang dikemukakan oleh Budiyo (2003) yaitu diantaranya pertama mengidentifikasi kan bahan-bahan yang telah diberikan beserta tujuan instruksionalnya, kedua membuat kisi-kisi dari soal tes yang diujikan, ketiga menyusun soal tes beserta kunci jawabannya, dan keempat menelaah soal tes sebelum dicetak dan digandakan. Budiyo (2003) indikator yang dijadikan pedoman untuk mengukur validitas isi pada penelitian ini adalah pertama menyesuaikan dengan kisi-kisi tes, kedua kesesuaian dengan tujuan penelitian, ketiga butir soal merupakan sampel yang representatif dari sebuah populasi atau sub kompetensi dasar, keempat butir soal tidak memerlukan pengetahuan lain dalam menjawab pertanyaannya, dan kelima soal telah menggunakan bahasa yang sesuai dengan kaidah Bahasa Indonesia. Berdasarkan uji validitas yang dilakukan di kelas XI IPA 1 pada butir soal yang terdiri dari 20 soal pilihan ganda dapat dilihat pada tabel diatas pada butir soal 1 sampai 19 dinyatakan valid, sedangkan pada butir soal nomor 20 tidak valid.

Dalam suatu penelitian diperlukan penentuan butir soal yang digunakan memiliki daya pembeda yang baik terhadap siswa yang berbeda, tempat yang berbeda dan waktu yang berbeda. Menurut Craker & Algina dalam Kusaeri dan Surapranata (2012:177) dapat diterima jika daya pembeda 0,30-1,00, ditolak/direvisi jika daya pembeda 0,2 butir soal merupakan kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa berkemampuan rendah berdasarkan kriteria tertentu. Untuk mengetahui daya pembeda dapat digunakan korelasi product moment dari pearson (Budiyo, 2003). Berdasarkan penelitian yang dilakukan Supandi dan Fahrikah (2016) sebagian besar pola pemikiran berpikirnya masih sederhana belum bisa berpikir tingkat tinggi sehingga dapat menyebabkan soal tersebut tidak valid. Sebab pada setiap soal butir soal yang diberikan hampr semua jawaban yang mendapatkan nilai di atas rata-rata dan di bawah rata-rata belum ada yang menjawab dengan tepat.

Penentuan besarnya indeks daya pembeda (DB) yang digunakan yaitu berkisar antara -1,00 sampai dengan 1,00. Apabila diperoleh indeks daya pembeda semakin besar maka butir soal semakin baik dalam membedakan kemampuan siswa yang berkemampuan tinggi dan yang berkemampuan rendah dan begitupun sebaliknya apabila daya pembedanya berindeks kecil maka akan menunjukkan

bahwa butir soal semakin buruk untuk membedakan kemampuan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Adapun klasifikasi dan interpretasi daya pembeda setiap butir soal untuk setiap interval dapat disajikan pada tabel 2.

**Tabel 2. Klasifikasi dan interpretasi daya pembeda butir untuk setiap interval**

Interval daya beda butir	Klasifikasi	Interpretasi
$-1,00 \leq DB < 0,20$	Jelek	Daya pembeda jelek
$0,20 \leq DB < 0,40$	Memuaskan	Memiliki daya pembeda yang cukup
$0,40 \leq DB < 0,70$	Baik	Memiliki daya pembeda yang baik
$0,70 \leq DB \leq 1,00$	Sangat baik	Memiliki daya pembeda yang sangat baik

(Suwarto, 2007; Arikunto, 2012)

Klasifikasi dan interpretasi daya pembeda dapat dilihat pada tabel 3. Pada Tabel 3 terdapat 8 soal dengan kategori daya pembeda sangat baik, 10 soal dengan kategori baik dan 2 soal dengan kategori memuaskan. Pada data tersebut setiap butir soal dapat di pakai sebab soal tersebut memiliki daya pembeda yang dapat membedakan siswa berkemampuan tinggi dan siswa berkemampuan rendah, sama hal nya dengan penelitian sebelumnya Supandi dan Fahrikhah (2016) karena harus digunakan soal-soal kategori daya pembeda yang tinggi sehingga terdapat kategori memuaskan yang memiliki indeks diskriminasi 0,2 sampai 0,4 masih bisa dipakai.

**Tabel 3. Klasifikasi dan interpretasi daya pembeda**

Interval daya beda butir	Klasifikasi	Frekuensi
$-1,00 \leq DB < 0,20$	Jelek	0
$0,20 \leq DB < 0,40$	Memuaskan	2
$0,40 \leq DB < 0,70$	Baik	10
$0,70 \leq DB \leq 1,00$	Sangat baik	8
Jumlah		20

Pada Tabel 3 terdapat 8 soal dengan kategori daya pembeda sangat baik, 10 soal dengan kategori baik dan 2 soal dengan kategori memuaskan. Pada data tersebut setiap butir soal dapat di pakai sebab soal tersebut memiliki daya pembeda yang dapat membedakan siswa berkemampuan tinggi dan siswa berkemampuan rendah, sama hal nya dengan penelitian sebelumnya Supandi dan Fahrikhah (2016) karena harus digunakan soal-soal kategori daya pembeda yang tinggi sehingga terdapat kategori memuaskan yang memiliki indeks diskriminasi 0,2 sampai 0,4 masih bisa dipakai.

Penelitian sebelumnya pun telah diperkuat dengan adanya teori menurut Anas Sudijono (2011: 386) bahwa mengetahui daya pembeda item sangat penting, sebab salah satu dasar pegangan untuk menyusun butir hasil belajar adalah adanya anggapan bahwa kemampuan antara peserta didik adalah berbeda maka salah satu

analisis yang harus dilakukan untuk mengetahui apakah butir soal dapat dikatakan baik sebagai alat evaluasi adalah analisis terhadap daya pembeda.

Semakin sukar butir soal tersebut maka semakin kecil indeks kesukarannya. Kategori sukar berada pada rentang nilai 0,00-0,30, kategori sedang berada pada rentang nilai 0,32-0,75 dan kategori mudah berada pada pada rentang 0,71-1,00. Distribusi tingkat kesukaran, dari 20 uji coba soal yang digunakan yaitu: 25% mudah (5 soal), 50% sedang (10 soal), dan 25% sukar (5 soal). Berdasarkan hasil perhitungan tingkat kesukaran soal bisa dilihat dari Tabel 4.

**Tabel 4. Tingkat Kesukaran**

Tingkat Kesukaran (P)	Kategori	Frekuensi
< 0,25	Sukar	3
0,25 - 0,75	Sedang	17
> 0,75	Mudah	0
Jumlah		20

Pada tabel 4 didapatkan hasil nol soal (0%) dengan tingkat kesukaran pada kategori mudah, 17 soal (85%) dengan tingkat kesukaran pada kategori sedang dan 3 soal (20%) dengan tingkat kesukaran pada kategori sukar. Berdasarkan penelitian yang sebelumnya dilakukan Agustina Maria (2016) menunjukkan bahwa distribusi soal belum memenuhi kriteria soal yang baik. Sebab untuk mendapatkan distribusi soal yang baik, maka butir-butir soal pada kategori mudah perlu ditambahkan 4 soal, pada kategori sedang perlu dikurangi 2 soal, dan pada kategori sukar perlu di tambah 2 soal.

## KESIMPULAN

Berdasarkan hasil analisis, kualitas butir soal kemampuan *Tree Thinking* pada soal uji coba materi spermatophyta yang pertama ,hasil uji validitas butir soal yang dikelompokkan valid sebesar 95%. Sedangkan soal yang dikelompokkan tidak valid sebesar 5%. Kedua, hasil uji daya pembeda pada semua butir soal yang di uji dapat diterima dan tidak perlu direvisi kerana sudah dapat membedakan peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah. Ketiga, hasil uji analisis tingkat kesukaran pada butir soal terdiri dari nol soal (0%) memiliki tingkat kesukaran pada kategori mudah, 17 butir soal (85%) memiliki tingkat kesukaran pada kategori sedang dan 3 soal (15%) memiliki tingkat kesukaran pada kategori sukar.

## DAFTAR PUSTAKA

- Anwar, Suroyo. (2009). *Pemahaman Individu, Observasi, Checklist, Interview, Kuesioner dan Sosiometri*. Yogyakarta: Pustaka Pelajar.
- Arikunto, S. (2008). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara
- Arikunto, S., (2009). *Prosedur Penelitian Suatu Pendekatan Praktik*. Edisi Revisi 6. Jakarta : Rineka Cipta.
- Arikunto, Suharsimi. (2012). *Prosedur Penelitian*. Jakarta: Rineka Cipta.

- Azmi, M. P. (2019). Analisis Pengembangan Tes Kemampuan Analogi Matematis pada Materi Segi Empat. *JURING (Journal for Research in Mathematics Learning)*, 2(2), 099. <https://doi.org/10.24014/juring.v2i2.7490>.
- Baum, D.A. & Smith, S.D. (2013). *Tree Thinking An Introduction To Phylogenetic Biology*. United states: Roberts and Company Publisher, Inc.
- Brinkman, F. & Leipe D. (2001). "Phylogenetic Analysis, A Practical Guide to the Analisis of Gene and Protein". *Baxevanis*, 5 (2), hlm. 323-358.
- Budiyono. 2003. Metodologi Penelitian Pendidikan. Surakarta: UNS Press.
- Mahbubah, H. G., Hidayat, T., & Supriatno, B. (2017). Plant vs. Animal, Which is the Most Prefer Understanding of Evolution? *International Journal of Science and Applied Science: Conference Series*, 2(1), 156.
- Mardapi, Djemari. (2012). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Yogyakarta. Nuha Medika.
- Mudjijo, (1995). *Tes Hasil Belajar*, Jakarta: Bumi Aksara.
- Novick, L. & Catley, K. (2012). Reasoning about Evolution's Grad Pptes: College Students' Understanding of The Tree of life. *American Educational Research Journal*, 20(10), hlm 1-40.
- Novick, L. & Catley, K. (2013). Reasoning about Evolution's Grad Pptes: College Students' Understanding of The Tree of life. *American Educational Research Journal*, 40(1), 138-177.
- Permendikbud. (2016). *Kompetensi Inti dan Kompetensi Dasar Sekolah Menengah Atas/Madrasah Aliyah (SMA/MA)*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Rustaman, N. (1991). *Klasifikasi Makhluk Hidup dalam materi pokok Pendidikan IPA 1*. Jakarta: Depdikbud.
- Sari, A. P. (2019). *Analisis Kemampuan Tree Thinking dan Hubungan dengan Penguasaan Konsep Siswa pada Materi Tumbuhan Biji*. Skripsi. Tidak Diterbitkan.
- Sugiyono. (2012). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta.
- Sulistyorini. 2009. *Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan*, Yogyakarta: TERAS.
- Supandi, S., & Farikhah, L. (2016). Analisis Butir Soal Matematika Pada Instrumen Uji Coba Materi Segitiga. *JIPMat*, 1(1), 71–78.
- Surapranata Sumarna. (2005). Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes, Implementasi Kurikulum 2004. Bandung: Remaja Rosdakarya.