

PENGGUNAAN *EFFECT SIZE* SEBAGAI MEDIASI DALAM KOREKSI EFEK SUATU PENELITIAN

Khairunnisa¹, Fenty Fitriani Sari², Mega Anggelena³,
Deka Agustika⁴, Euis Nursa'adah⁵
Universitas Bengkulu^{1,2,3,4,5}
mayanisa0411@gmail.com¹

ABSTRAK

Penelitian ini bertujuan untuk mengkaji studi-studi tentang penggunaan *effect size* dalam memperkirakan ukuran efek suatu percobaan di berbagai bidang ilmu. Metode yang digunakan yaitu *Literature Review* dengan menelusuri artikel-artikel relevan pada *database Google Scholar*, *National Center or Biotechnology Information (NCBI)*, *Mendeley* dan *ScienceDirect*. Hasil penelitian menunjukkan bahwa secara keseluruhan artikel yang direview memperlihatkan penggunaan *effect size* sebagai uji statistik untuk menentukan ukuran efek dari suatu perlakuan dalam studi riset. Namun dalam penggunaannya, *effect size* perlu disertakan juga dengan uji statistik lainnya guna memperoleh indeks bias yang kecil agar data lebih akurat. Simpulan, *effect size* adalah bagian dari uji statistik yang berperan dalam membantu pengguna memahami besarnya perbedaan yang ditemukan dalam suatu percobaan penelitian.

Kata Kunci: Akurasi, *Effect Size*, Percobaan, Statistik

ABSTRACT

This study aims to examine studies on the use of effect size in estimating the effect size of an experiment in various fields of science. The method used is Literature Review by searching for relevant articles in the databases of Google Scholar, National Center or Biotechnology Information (NCBI), Mendeley and ScienceDirect. The results of the study indicate that overall the articles reviewed show the use of effect size as a statistical test to determine the effect size of treatment in research studies. However, in its use, the effect size needs to be included with other statistical tests to obtain a small refractive index so that the data is more accurate. In conclusion, the effect size is part of a statistical test that plays a role in helping users understand the magnitude of the differences found in a research experiment.

Keywords: Accuracy, *Effect Size*, Experiment, Statistics

PENDAHULUAN

Statistika merupakan istilah yang berasal dari Bahasa Italia *statistita* yang berarti negarawan. Pada mulanya statistika digunakan dalam mengelola data kenegaraan. Statistika digunakan dalam segala bidang ilmu dan kehidupan. Statistika digunakan dalam pemecahan masalah secara nyata dalam kehidupan manusia sehari-hari. Tidak dapat dipungkiri bahwa perkembangan ilmu pengetahuan dan teknologi sebagai peran dari statistika. Statistika digunakan sebagai alat bantu dalam pengumpulan data, melakukan analisis atas gejala-gejala yang diamati (Anam, 2020). Salah satu uji statistika yang dapat digunakan untuk analisis tersebut adalah *effect size*.

Effect size adalah perbedaan kejadian efek antara kelas kontrol dan kelas eksperimen merupakan. Menurut Cohen (1988) *effect size* digunakan untuk untuk mengetahui besarnya pengaruh setelah diberikan perlakuan. *Effect size* umumnya digunakan pada penelitian berpopulasi besar serta memiliki variabel dan pembandingan yang jelas (Kirby & Gerlanc, 2013; Schäfer & Schwarz, 2019). Sejauh ini informasi terkait *effect size* masih dalam spesifikasi kajian penggunaannya sebagai alat ukur statistik dalam suatu penelitian, namun belum ada studi yang membahas secara mendalam perkembangan penelitian-penelitian yang memanfaatkan *effect size* dan kriteria penggunaannya dalam riset studi. Penelitian ini bertujuan untuk mengumpulkan informasi-informasi terkait penggunaan *effect size* dalam memperkirakan ukuran efek suatu percobaan di berbagai bidang diantaranya bidang kesehatan dan pendidikan.

METODE PENELITIAN

Penelitian ini menggunakan metode *Literature Review* yang bertujuan untuk mengumpulkan dan mengambil intisari serta menganalisa beberapa artikel ilmiah guna memberikan pemahaman tentang perkembangan pengetahuan dan menciptakan ide baru terkait

penggunaan *Effect Size*. Database yang digunakan dalam pencarian literatur adalah *Google Scholar*, *National Center or Biotechnology Information (NCBI)*, *Mendeley* dan *ScienceDirect* dengan kata kunci "*effect size*". Semua artikel terkait *effect size* dikumpulkan kemudian direview berdasarkan kriteria inklusi dari referensi terpilih sebagai sumber yang relevan. Kriteria inklusi dalam penelitian ini adalah *full text* artikel ilmiah yang dipublikasikan pada tahun 2011-2021.

Berdasarkan uraian yang telah disampaikan sebelumnya, maka penulis merasa perlu untuk melakukan kajian tentang *effect size*, serta manfaatnya pada berbagai disiplin ilmu, adapun tujuan untuk mengkaji studi-studi tentang penggunaan *effect size* dalam memperkirakan ukuran efek suatu percobaan di berbagai bidang ilmu.

HASIL PENELITIAN

Hasil penelusuran literatur pada studi ini memperoleh sebanyak 19 artikel ilmiah terkait penggunaan *effect size* dalam berbagai bidang ilmu, seperti di bidang kesehatan dan pendidikan. Detail hasil review disajikan pada tabel 1, dibawah ini.

Tabel 1. Hasil Review Literatur

Tahun	Identitas article	Metode	Hasil	Kesimpulan
2020	Mattan S., dkk. <i>effects size: Estimation of Effect Size Indices and Standardized Parameters</i> . The Journal of Open Source Software. 2020.	<ul style="list-style-type: none"> • Standarisasi Parameter dan Model • <i>Effect size</i> untuk ANOVA • Konversi <i>Effect size</i> dari uji Statistik 	<i>Effect size</i> adalah paket R (Tim Inti R, 2020) yang mengisi celah penting ini, menyediakan utilitas untuk dengan mudah memperkirakan berbagai <i>Effect size</i> standar (yaitu, <i>Effect size</i> yang tidak terikat dengan unit pengukuran variabel yang diinginkan) dan interval kepercayaan (CI), dari berbagai dari model statistik. <i>Effect size</i> menyediakan fungsi yang mudah digunakan, dengan dokumentasi lengkap dan penjelasan tentang berbagai <i>Effect size</i> yang ditawarkan, dan juga digunakan oleh pengembang paket R lainnya sebagai back-end untuk perhitungan <i>Effect size</i>	<i>Effect size</i> adalah bagian dari statistik mudah, proyek kolaboratif yang dibuat untuk memfasilitasi penggunaan nilai r untuk analisis statistik.
2020	Funder, D. C., & Ozer, D. J. (2019). <i>Evaluating effect size in psychological research: Sense and nonsense</i> . <i>Advances in Methods and Practices in Psychological Science</i> , 2, 156–168. doi:10.1177/2515245919847202	Standak korelai dan kuadrat korelasi	menjelaskan masalah dengan evaluasi sederhana dengan <i>effect size</i> , termasuk cara umum yang disalahartikan — kesalahan paling umum adalah menggambarkannya dengan cara yang tidak menyampaikan informasi yang berguna atau secara aktif menyesatkan. Selanjutnya, menguraikan beberapa cara (berdasarkan proposal oleh penulis sebelumnya) untuk mengilhami angka <i>effect size</i> dengan makna	menawarkan beberapa rekomendasi untuk cara yang paling berguna untuk mengevaluasi <i>Effect size</i> dan bahkan, dengan berani, menyarankan seperangkat standar baru. Harapan kami adalah bahwa artikel ini mungkin memainkan peran kecil dalam membantu memajukan pengobatan <i>Effect size</i> sehingga daripada menjadi angka yang dilaporkan tanpa Merevisi pedoman Cohen interpretasi, atau ditafsirkan secara dangkal atau salah,
2021	Zubaidah, S. dkk., <i>The Effect Size of Different Learning on Critical and Creative Thinking Skills of Biology Students</i> . <i>International Journal of Instruction</i> 14(3):187-206. 2021 DOI:10.29333/iji.2021.14311a	Interpretasi <i>effect size</i> yang digunakan dalam penelitian ini mengacu pada standarisasi), yang meliputi 0,10 sebagai interpretasi rendah, 0,30 sebagai interpretasi sedang, dan 0,50 sebagai interpretasi besar	Hasil uji <i>effect size</i> yang dilakukan pada kedua pembelajaran menunjukkan bahwa nilai <i>effect size</i> memiliki kriteria tinggi dan rendah.	Uji efek ukuran memiliki beberapa keunggulan karena informasi yang diperoleh dari pengujian semacam ini berkaitan erat dengan kualifikasi hasil analisis data, atau yang biasa dikenal dengan interpretasi hasil uji signifikansi statistik. Sedangkan uji ANCOVA LSD tidak membahas kualifikasi signifikansi statistik, tetapi menjelaskan kuantifikasi hasil uji signifikansi statistik. Dengan mengetahui <i>Effect size</i> suatu model pembelajaran, pendidik dapat memilih model pembelajaran yang paling tepat untuk kegiatan belajar mengajarnya. Oleh karena itu, melalui penelitian semacam ini, diharapkan peneliti selanjutnya tidak hanya menyajikan hasil uji signifikansi statistik atau hanya menyajikan hasil uji <i>Effect size</i> , namun kedua jenis tes tersebut perlu dilakukan secara bersamaan agar informasi yang

				diperoleh dapat saling melengkapi karena hasil dari masing-masing jenis tes tersebut memiliki beberapa kelebihan dan kekurangan. Informasi yang diperoleh dari hasil uji Effect size dapat memperkuat hasil uji signifikansi statistik dalam setiap penelitian kuantitatif.
2000	Christopher J. Ferguson. <i>An Effect Size Primer: A Guide for Clinicians and Researchers</i> . Professional Psychology: Research and Practic. 2020		Effect size dibagi ke dalam empat kategori umum 1. Indeks perbedaan kelompok. 2. Kekuatan indeks asosiasi 3. Perkiraan yang dikoreksi. 4. Perkiraan risiko Panduan untuk membantu dalam pemilihan effect size dan interpretasinya : 1. Pedoman disarankan sebagai batas minimum, tidak menjamin bahwa effect size yang melebihi batas tersebut berarti. Keterbatasan studi, kegagalan untuk mengontrol prediktor lain yang relevan atau ancaman terhadap validitas internal, reliabilitas dan validitas tanggapan terhadap ukuran yang digunakan, dll., harus dipertimbangkan ketika menginterpretasikan effect size. 2. Effect size yang dikoreksi lebih disukai daripada Effect size yang tidak dikoreksi 3. Untuk desain korelasional, parsial r dan koefisien regresi standar lebih unggul daripada bivariat r karena mereka memperkirakan varians unik yang disebabkan oleh prediktor yang mengendalikan variabel lain 4. Meskipun effect size dimaksudkan untuk mewakili efek "benar" dalam populasi, penting untuk dipahami bahwa perkiraan Effect size dapat dipengaruhi oleh pengambilan sampel dan pengukuran.	Tidak ada Effect size yang sempurna untuk semua situasi, dan beberapa ketidaksepakatan tetap ada atas keuntungan beberapa dari yang lain. Peneliti didorong untuk memilih dengan bijak (dan konservatif) dari antara pilihan yang disajikan. Interpretasi Effect size akan selalu tetap spesifik konteks. Menerima Effect size dari besarnya apa pun sebagai signifikan secara klinis atau praktis menjadikan kegunaannya dapat diperdebatkan. Psikolog harus bersedia untuk secara objektif mengidentifikasi efek kecil dan menafsirkannya.
2018	Paul Monsarrat ^{1,2} and Jean-Noel Vergnes ^{3,4} . <i>Data mining of effect sizes from PubMed abstracts: a cross-study conceptual replication</i> . Bioinformatics. 2018	1. Menghitung tingkat kesalahan dalam Effect size (ESs) yang dilaporkan dari database Studi Effect size 2. Menilai evolusi temporal ES	Replikasi langsung berguna karena dapat memeriksa kurangnya validitas internal tetapi dibatasi oleh fakta bahwa setiap kelemahan metodologis dalam penelitian asli akan diabadikan dalam penelitian berikutnya. Di sisi lain, replikasi konseptual hanya mensyaratkan bahwa kondisi esensial sangat cocok dengan aslinya, sementara memungkinkan fleksibilitas untuk memvariasikan kondisi non-esensial lainnya	Untuk mempersiapkan replikasi konseptual, dapat melakukan replikasi langsung dari Studi Kesalahan, menggunakan data mentah yang ada. Kemudian dilanjutkan menerapkan metode yang sama ke database Studi effect size. Sehingga menemukan hasil serupa mengenai tingkat perbedaan, dengan kisaran kesalahan yang diperkirakan antara 5,3 dan 8,1% dari ES yang dilaporkan dalam abstrak PubMed, dan juga menemukan hasil serupa mengenai penurunan tingkat kesalahan dari waktu ke

			waktu
2018	<p>Julie Lorah Correspondence: jlorah@iu.edu Universitas Indiana, Bloomington, AS 2018 Lorah Large-scale Assess Educ (2018) 6:8 https://doi.org/10.1186/s40536-018-0061-2</p>	<p>Efek size adalah bagian penting dari interpretasi untuk studi pemodelan bertingkat yang diterapkan. Peneliti dapat menggunakan ICC untuk mewakili besarnya efek acak yang dapat mewakili efek negara dan/atau sekolah dan koefisien regresi standar atau untuk mewakili besarnya efek tetap yang dapat mewakili hubungan kepentingan saat memeriksa pertanyaan substantif menggunakan kumpulan data internasional. Kompleksitas yang terkait dengan model tiga tingkat, pelaporan, dan lereng acak telah dieksplorasi. Topik dalam penelitian ini telah didemonstrasikan dengan menggunakan data TIMSS, tetapi saran yang diberikan dapat diterapkan pada analisis multilevel data primer atau sekunder</p>	<p>Efek size adalah bagian penting dari interpretasi untuk studi pemodelan bertingkat yang diterapkan</p>
	<p>Cristiano Ialongo Laboratory Medicine Department, "Tor Vergata" University Hospital, Rome, Italy Department of Human Physiology and Pharmacology, University of Rome Sapienza, Rome, Italy cristiano.ialongo@gmail.com http://dx.doi.org/10.11613/BM.2016.015</p>	<p>Menghitung tingkat kesalahan dalam Effect size (ESs) yang dilaporkan dari database Studi efek size</p>	<p>Paradigma mendikte nilai-P membuat kita lupa apa itu bukti ilmiah dan apa maknanya dalam penilaian statistiknya. Meskipun demikian, melalui ES dapat mulai mengajari diri kami sendiri tentang pembobotan temuan terhadap peluang dan besarnya, dan itu akan sangat membantu dalam apresiasi kami atas pencapaian ilmiah apa pun. N</p>
	<p>Gail M. Sullivan, MD, MPH Richard Feinn, PhD University of Connecticut, 253 Farmington Avenue, Farmington, CT 06030-5215, gsullivan@nso1.uconn.edu DOI: http://dx.doi.org/10.4300/JGME-D-12-00156.1</p>	<p>Effect size membantu pembaca memahami besarnya perbedaan yang ditemukan, sedangkan signifikansi statistik memeriksa apakah temuan tersebut kemungkinan besar terjadi secara kebetulan. Intervensi pendidikan menghasilkan peningkatan nilai ujian mata pelajaran dengan total rata-rata 15 dari 50 pertanyaan dibandingkan dengan intervensi lain, efek size absolut adalah 15 pertanyaan atau 3 tingkat kelas (30%) lebih baik pada ujian. Effect size absolut tidak memperhitungkan variabilitas skor, karena tidak setiap subjek mencapai hasil rata-rata</p>	<p>Efek size membantu pembaca memahami besarnya perbedaan yang ditemukan</p>
2017	<p>Jolynn Pek and David B. Flora Pek, J., & Flora,</p>	<p>Menginterpretasikan efek size dan CI studi psikologis empiris merupakan dasar untuk melakukan</p>	<p>Efect size dalam konteks sejarah dan penelitian berpotensi membandingkan</p>

	<p>DB (2017, 9 Maret). Melaporkan Effect size dalam Penelitian Psikologis Asli: Diskusi dan Tutorial. Metode Psikologis. Memajukan publikasi online. http://dx.doi.org/10.1037/met0000126</p>		<p>penelitian dan akumulasi pengetahuan. Berfokus pada effect size dan variabilitas pengambilan sampelnya menghambat interpretasi hasil secara dikotomis yaitu, signifikan secara statistik atau tidak (Waldman & Lilienfeld, 2015). Konsisten dengan revolusi metodologis modern untuk pendekatan berbasis aturan (cf., Gonzalez, 1994; Rodgers, 2010; Tukey, 1969), penelitian ini juga harus dipertimbangkan dari sudut pandang di mana effect size adalah landasan untuk membangun argumen yang masuk akal (Abelson, 1995). Maka dari itu untuk meningkatkan effect size, maka akan dilakukan petinjauan prinsip-prinsip umum untuk penulisan ilmiah yang baik dan menerapkan prinsip-prinsip ini pada konteks effect size. Karena mengembangkan teori psikologi dan melakukan penelitian menggunakan ukuran psikologis adalah cair di alam, effect size dan interpretasinya harus cukup bervariasi untuk mencocokkan pertanyaan penelitian yang sedang berkembang, desain penelitian, dan estimasi statistik. Dalam konteks pelaporan penelitian utama, effect size harus dioptimalkan untuk mengomunikasikan temuan penting dengan cara yang dapat diakses, intuitif, dan berharga yang berkontribusi pada literatur yang ada.</p>	<p>efek penting dengan efek yang dipublikasikan, dan idealnya memotivasi pertanyaan penelitian tambahan yang berpotensi menambah pengetahuan baru ke paradigma penelitian. Interpretasi yang bermakna dari effect size yang dilaporkan adalah inti dari apa yang berkontribusi pada sains kita. Interpretasi mereka harus mencerminkan kematangan teori dan program penelitian kami, ketepatan estimasi mereka, dan pengetahuan tentang signifikansi praktisnya. Contoh empiris yang diberikan, yang menggambarkan rekomendasi kami untuk praktik dengan menekankan interpretasi effect size dalam kaitannya dengan pertanyaan penelitian mereka dan kematangan paradigma penelitian mereka, berfungsi sebagai tutorial bagi peneliti substantif untuk membantu meningkatkan praktik dalam pelaporan penelitian tentang ukuran efek.</p>
2018	<p>Farid Chakhssi1, Jannis T. Kraiss, Marion Sommers-Spijkerman and Ernst T. Bohlmeijer Chakhssi1dkk. Psikiatri BMC (2018) 18:211 https://doi.org/10.1186/s12888-018-1739-2</p>	<p>Tinjauan sistematis dan meta-analisis dilakukan mengikuti pedoman PRISMA. PsycINFO, PubMed, dan Scopus</p>	<p>Tiga puluh studi dimasukkan, mewakili 1.864 pasien dengan gangguan klinis. Pada pasca-intervensi, PPI menunjukkan effect size yang signifikan dan kecil untuk kesejahteraan (Hedges'g =0,24) dan depresi (g =0,23) dibandingkan dengan kondisi kontrol saat menghilangkan outlier. Perbaikan moderat yang signifikan diamati untuk kecemasan (g =0,36). Effect size untuk stres tidak signifikan. Efek tindak lanjut (8-12 minggu), jika tersedia, menghasilkan effect size yang serupa. Kualitas studi rendah sampai sedang.</p>	<p>Temuan ini menunjukkan bahwa PPI, di mana fokusnya adalah memunculkan perasaan, kognisi, atau perilaku positif, tidak hanya berpotensi meningkatkan kesejahteraan, tetapi juga dapat mengurangi tekanan pada populasi dengan gangguan klinis. Mengingat meningkatnya minat PPI dalam pengaturan klinis, penelitian berkualitas lebih tinggi diperlukan untuk menentukan efektivitas PPI dalam sampel klinis</p>
2019	<p>Victor Orozco-Olvera1D , Fuyuan Shen, Lucie Cluver https://doi.org/10.1371/journal.pone.0209969 February 12, 2019</p>	<p>Pencarian elektronik dan manual dilakukan untuk mengidentifikasi studi eksperimental dan kuasi-eksperimental dengan desain kontrafaktual kuat yang diterbitkan antara tahun 1985 dan kuartal pertama 2017. Effect size dianalisis secara meta menggunakan model efek campuran.</p>	<p>Delapan penelitian eksperimental dan dua penelitian kuasi-eksperimental memenuhi kriteria inklusi kami. Ukuran sampel agregat adalah 23.476 peserta, dengan median 902 peserta per studi. Narasi pendidikan hiburan memiliki efek kecil namun signifikan untuk tiga perilaku seksual. Ini mengurangi jumlah pasangan seksual [perbedaan rata-rata standar, (SMD) = 0,17, interval kepercayaan 95% (CI) = 0,02–0,33, tiga ukuran efek], mengurangi hubungan seks tanpa kondom (SMD = 0,08, 95% CI =</p>	<p>Meskipun hiburan media massa memiliki efek kecil dalam mempromosikan praktik seksual yang lebih aman, skala ekonominya dibandingkan intervensi tatap muka menunjukkan potensinya untuk menjadi alat yang hemat biaya di atas ambang batas penonton. Penggunaan peserta studi dari populasi pemuda umum dan penggunaan sebagian besar uji coba efektivitas mengurangi kekhawatiran</p>

		0,03–0,12 , sembilan ukuran efek), dan peningkatan pengujian dan manajemen untuk IMS (SMD = 0,29, 95% CI = 0,11-0,46, dua ukuran efek). Intervensi tidak efektif dalam mengurangi seks antar generasi, diukur melalui perbedaan usia dengan pasangan seksual (SMD = 0,06, 95% CI = -0,06–0,19, empat ukuran efek). Pendidikan hiburan memiliki efek ukuran sedang pada hasil pengetahuan (SMD = 0,67, 95% CI = 0,32-1,02, tujuh ukuran efek), di mana hubungan waktu-peluruhan diamati. Tidak ada efek yang ditemukan pada sikap.	mengenai skalabilitasnya. Kurangnya keseluruhan studi berkualitas tinggi menegaskan perlunya memperkuat basis bukti pendidikan hiburan. Penelitian masa depan harus dilakukan untuk memahami efek moderator untuk subkelompok yang berbeda dan karakteristik intervensi.
2017	Tom Meyvis Stijn M. J. Van Osselaer Oxford University Journal of Consumer Researc Vol. 0 November 2017 DOI: 10.1093/jcr/ucx110	Perhitungan ES adalah bagian dari kerangka analisis daya, sehingga perhitungan ukurannya biasanya disediakan tertanam dalam paket perangkat lunak statistik atau dicapai melalui standar. Pentingnya memberikan efek size di samping nilai-P harus ditekankan, karena ini adalah nilai tambah untuk setiap penelitian yang mewakili langkah menuju kebenaran ilmiah. Untuk alasan ini, peneliti harus didorong untuk menunjukkan ES dalam pekerjaan mereka, terutama melaporkannya setiap kali nilai-P disebutkan. Sebaiknya juga menyediakan CI bersama dengan ES	Krisis replikasi dalam psikologi dan riset konsumen telah menyebabkan fokus pada peningkatan ukuran sampel dan ketidakpercayaan umum terhadap pendekatan alternatif untuk meningkatkan kekuatan: peningkatan ukuran efek. Namun, mempelajari efek yang menarik secara psikologis tanpa mengambil langkah yang tepat untuk meningkatkan effect size yang diamati seringkali membutuhkan ukuran sampel yang besar yang akan membebani sebagian besar sumber daya peneliti yang terbatas (dalam hal pendanaan, tenaga kerja, dan peserta yang tersedia). Oleh karena itu kami telah mengusulkan bahwa peningkatan yang wajar dalam ukuran sampel harus dilengkapi dengan desain dan keputusan analisis yang meningkatkan effect size dengan meningkatkan dampak manipulasi atau dengan mengurangi varians kesalahan. Secara khusus, kami telah menyediakan pedoman untuk menghapus peserta, menggunakan kovariat, menggunakan desain dalam subjek, memilih manipulasi dan rangsangan yang lebih efektif, menggunakan tindakan yang lebih sensitif dan andal, merekrut peserta yang tepat, mengurangi variasi latar belakang, melakukan uji statistik yang lebih kuat, dan mengejar fenomena yang tidak terlalu halus dan berlawanan dengan intuisi. Untuk mengambil keuntungan dari strategi ini untuk meningkatkan effect size yang diamati, peneliti perlu

merencanakan studi dan prosedur mereka dengan hati-hati sambil menjaga dari penyalahgunaan derajat kebebasan peneliti dan p-hacking. Namun, jika dilakukan dengan benar, pendekatan ini akan memungkinkan peneliti untuk secara memadai memperkuat studi mereka tanpa harus merekrut (dan memberi kompensasi) sejumlah besar peserta. Seringkali jauh lebih efisien dan diagnostik untuk menjalankan beberapa studi yang direncanakan dengan hati-hati daripada banyak studi yang kurang bertenaga. merekrut peserta yang tepat, mengurangi variasi latar belakang, melakukan uji statistik yang lebih kuat, dan mengejar fenomena yang kurang halus dan berlawanan dengan intuisi. Untuk mengambil keuntungan dari strategi ini untuk meningkatkan effect size yang diamati, peneliti perlu merencanakan studi dan prosedur mereka dengan hati-hati sambil menjaga dari penyalahgunaan derajat kebebasan peneliti dan p-hacking. Namun, jika dilakukan dengan benar, pendekatan ini akan memungkinkan peneliti untuk secara memadai memperkuat studi mereka tanpa harus merekrut (dan memberi kompensasi) sejumlah besar peserta. Seringkali jauh lebih efisien dan diagnostik untuk menjalankan beberapa studi yang direncanakan dengan hati-hati daripada banyak studi yang kurang bertenaga. merekrut peserta yang tepat, mengurangi variasi latar belakang, melakukan uji statistik yang lebih kuat, dan mengejar fenomena yang kurang halus dan berlawanan dengan intuisi. Untuk mengambil keuntungan dari strategi ini untuk meningkatkan effect size yang diamati, peneliti perlu merencanakan studi dan prosedur mereka dengan hati-hati sambil menjaga dari penyalahgunaan derajat kebebasan peneliti dan p-hacking. Namun, jika dilakukan dengan benar, pendekatan ini akan

			memungkinkan peneliti untuk secara memadai memperkuat studi mereka tanpa harus merekrut (dan memberi kompensasi) sejumlah besar peserta. Seringkali jauh lebih efisien dan diagnostik untuk menjalankan beberapa studi yang direncanakan dengan hati-hati daripada banyak studi yang kurang	
2019	Ting Shen & Spyros Konstantopoulos Ting Shen & Spyros Konstantopoulos (2019): Estimating causal effects of class size in secondary education: evidence from TIMSS, Research Papers in Education, DOI: 10.1080/02671522.2019.1697733	Penelitian ini menerapkan metode variabel instrumental (IV) dan menggunakan desain diskontinuitas regresi (RDD) untuk melakukan analisis data TIMSS pada tahun 2003, 2007 dan 2011	menunjukkan bahwa di Rumania pada tahun 2003 kelas yang lebih kecil memiliki efek yang signifikan dan positif pada nilai akademik dalam matematika, fisika, kimia dan ilmu bumi dan tahun 2007 tentang kenikmatan belajar matematika. Di Lituania, pada tahun 2011 kelas kecil berpengaruh signifikan dan positif terhadap kenikmatan belajar biologi dan kimia serta belajar kimia dengan baik. Namun pada tahun 2007, efeknya terbalik pada beberapa hasil non-kognitif terkait biologi. Secara keseluruhan, efek ukuran kelas yang signifikan juga besar dan biasanya jauh lebih besar daripada efek yang dilaporkan dalam Project STAR	bahwa untuk negara-negara tertentu (yaitu Rumania dan Lithuania), selama bertahun-tahun dan untuk hasil kognitif dan nonkognitif tertentu, efeknya cukup besar dan biasanya lebih besar daripada efek yang dilaporkan di Project STAR. Namun, sebagian besar efek ukuran kelas tidak berbeda dari nol, yang menunjukkan bahwa pengurangan ukuran kelas tidak secara otomatis menjamin peningkatan kinerja siswa. Banyak faktor proses dan dinamika kelas lainnya dan harus bekerja sama dengan baik untuk mencapai hasil yang sukses dalam pembelajaran siswa.
	Michael Sanders King's College London https://ssrn.com/abstract=3532325	Uji coba acak	menemukan bahwa effect size rata-rata kira-kira sepertiga ukuran yang digambarkan sebagai 'kecil' dalam aturan praktis yang banyak digunakan, bahwa kurang dari 5% dari semua effect size memenuhi ambang batas. untuk 'kecil' dan effect size itu cukup stabil dari waktu ke waktu dan usia anak-anak yang terlibat. Kami menemukan bahwa intervensi yang dikelompokkan untuk pengacakan biasanya menghasilkan effect size yang lebih kecil, dan intuisi peneliti tentang peringkat ordinal temuan cukup akurat	pertama bahwa perhitungan ukuran sampel berdasarkan aturan praktis yang diusulkan oleh Cohen (1988), atau oleh Hattie (2008) biasanya terlalu ambisius. Kami juga menemukan bahwa ini tidak bervariasi secara signifikan dari waktu ke waktu (sejauh ini), atau berdasarkan usia anak muda, meskipun kami menemukan bahwa intervensi yang diacak pada tingkat individu rata-rata lebih efektif daripada intervensi yang diacak di tingkat kelompok kedua bahwa meskipun peneliti terlalu optimis dalam memperkirakan ukuran efek, intuisi mereka tentang urutan peringkat kemungkinan efek cukup akurat, menyarankan tempat untuk aturan praktis secara umum, tetapi hanya jika mereka dipusatkan dengan tepat. Oleh karena itu, kami mengusulkan bahwa aturan praktis baru diadopsi untuk uji coba terkontrol secara acak pendidikan. Mengikuti formulasi Cohen "kecil, sedang, besar", kami menyarankan pembaruan

				dengan nilai numerik baru, daripada keberangkatan grosir. Aturan yang kami sarankan dapat ditemukan pada tabel di bawah ini, di samping studi Cohen dan contoh dari EEF dengan perkiraan efek tersebut.
2018	Mark J. Lachowicz and Kristopher J. Preacher Psychological Methods 2018, Vol. 23, No. 2, 244–261	Demonstrasikan aplikasi dan interpretasi effect size menggunakan data nyata	Analisis mediasi digunakan untuk menguji proses melalui mana prediktor memiliki efek pada hasil melalui variabel intervensi yang disebut mediator. Komponen efek yang ditransmisikan melalui mediator dikenal sebagai efek tidak langsung. Meskipun efek tidak langsung umumnya dilaporkan, effect size untuk mereka belum ditetapkan secara pasti. Kami menunjukkan bahwa ukuran yang kami usulkan adalah pilihan yang menarik karena beberapa alasan, tetapi yang paling penting bahwa (a) memiliki interpretasi intuitif, (b) mengukur efek tidak langsung independen dari pilihan desain yang sewenang-wenang, dan (c) dapat digunakan untuk menarik kesimpulan yang valid untuk ukuran sampel dan besaran efek yang umum dalam penelitian terapan.	bahwa memajukan penelitian tentang model mediasi dengan menurunkan effect size yang bermakna secara teoritis dan sangat berguna. Selanjutnya, bahwa telah terbukti menjadi penaksir kualitas dan dengan demikian menawarkan kemajuan tidak hanya untuk literatur metode tetapi juga untuk pengguna model mediasi
2018	Miočević, M., O'Rourke, H. P., MacKinnon, D. P., & Brown, H. C. (2018). Statistical properties of four effect-size measures for mediation models. <i>Behavior Research Methods</i> , 50(1), 285–301. https://doi.org/10.3758/s13428-017-0870-1	Dalam studi yang berisi 320 kombinasi parameter, perangkat lunak SAS (Versi 9.2 dari Sistem SAS untuk Windows, Cary, NC, USA) digunakan untuk melakukan simulasi yang menghitung bias, bias relatif, dan standar deviasi effect size lebih dari 1.000 ulangan. Kekuatan empiris didefinisikan sebagai persentase interval kepercayaan untuk ukuran efek-ukuran yang tidak mengandung nol ketika efek yang benar ada dalam populasi; nilai 0,8 dan lebih tinggi dianggap diinginkan. Lebar interval didefinisikan sebagai perbedaan antara batas kepercayaan atas dan batas kepercayaan bawah; lebar interval yang lebih kecil menunjukkan perkiraan yang lebih presisi, namun, karena keempat effect size tidak pada metrik yang sama, lebar intervalnya tidak dapat dibandingkan secara langsung, sehingga mengurangi lebar interval dengan peningkatan ukuran	Effect size standar merupakan tindakan mediasi terbaik. Solusi analitik untuk kesalahan standar dari setiap effect size untuk efek mediasi dapat mengkonfirmasi dan menjelaskan temuan dari studi simulasi dan memperkuat rekomendasi tentang kegunaan effect size tertentu dalam meta-analisis model mediasi. Namun, interval kepercayaan berdasarkan distribusi normal yang diasumsikan untuk ukuran effect size dan rumus analitik yang sesuai untuk kesalahan standar ukuran effect size mungkin tidak akurat, sehingga solusinya adalah menggunakan metode bootstrap ideal. Situasi ini menyoroti pentingnya interpretasi untuk pemilihan ukuran efek. Ketika memilih effect size untuk mewakili efek tidak langsung, seseorang harus memilih effect size yang menjawab pertanyaan penelitian dengan paling akurat.	

		<p>sampel digunakan sebagai kriteria. Bias standar dihitung dengan membagi bias effect size dengan standar deviasinya pada setiap kombinasi nilai parameter dan ukuran sampel. Empat effect size memiliki rentang bias standar yang sebanding untuk semua ukuran sampel yang berarti bahwa bias dan efisiensi proporsional untuk keempat ukuran efek. Dengan kata lain, semakin bias ukuran efek, semakin kurang efisiennya. Sebaliknya, effect size yang memiliki bias rendah juga memiliki standar deviasi yang kecil.</p>	
2013	<p>Kirby, K. N., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. <i>Behavior Research Methods</i>, 45(4), 905–927. https://doi.org/10.3758/s13428-013-0330-5</p>	<p>Metode bootstrap membutuhkan komputasi yang intensif karena membutuhkan banyak sampel ulang dan perhitungan. Bootstrap CI tidak dapat menangani masalah dg ukuran sampel yang kecil. Meskipun CI BCa mungkin mengungguli CI tradisional untuk sampel kecil dari populasi tidak normal, cakupannya untuk ukuran sampel kecil masih dapat berbeda secara substansial dari nominal 1. Ini karena CI bootstrap sangat bergantung pada nilai sampel di bagian akhir distribusi sampel. Beberapa statistik sampel terkecil dan terbesar akan dihitung dari sampel ulang yang terjadi untuk menarik pengamatan terutama dari bagian akhir distribusi nilai sampel. Metode bootstrap membutuhkan komputasi yang intensif karena membutuhkan banyak sampel ulang dan perhitungan. Karena subjek biasanya merupakan faktor acak, metode pengambilan sampel ulang harus secara acak mengambil ulang subjek dalam kelompok, dengan tetap memperhatikan jumlah subjek dalam setiap kelompok dalam sampel</p>	<p>BootES adalah paket gratis yang menawarkan kombinasi fitur yang tidak peneliti temukan di perangkat lunak lainnya. BootES ini dapat menghitung CI bootstrap untuk effect size standar dan tidak standar untuk kontras dan melakukan penskalaan effect size tidak standar untuk mempertahankan unit pengukuran asli. Dengan memanfaatkan struktur data untuk membantu memilih fungsi effect size bawaan yang sesuai, bootES dapat melakukan sebagian besar hal ini dengan satu perintah dan sejumlah kecil spesifikasi opsi.</p>

2019	<p>(Schäfer & Schwarz, 2019) Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. <i>Frontiers in Psychology</i>, 10, 1–13. https://doi.org/10.3389/fpsyg.2019.00813</p>	<p>asli.</p> <p>Ada tiga elemen metodologis kunci dalam penelitian ini. Pertama, untuk mendapatkan gambaran representatif dari efek yang dipublikasikan dalam psikologi, peneliti menganalisis pilihan acak dari studi empiris yang dipublikasikan. Keacakan memastikan bahwa setiap studi memiliki probabilitas yang sama untuk ditarik, yang merupakan jalur paling andal menuju kesimpulan yang dapat digeneralisasikan.</p> <p>Kedua, untuk memperkirakan seberapa kuat efek yang dipublikasikan mungkin bias, peneliti membedakan antara studi dengan dan tanpa pra-pendaftaran. Ketiga, untuk membandingkan sub-disiplin yang berbeda, peneliti mengkategorikan berbagai cabang psikologi menjadi sembilan kelompok dan secara acak menggambar dan menganalisis efek dalam setiap kelompok.</p>	<p>Pendekatan perbandingan maupun pendekatan konvensi tidak dapat diterapkan untuk menafsirkan kebermaknaan suatu efek tanpa mengalami masalah yang besar. Perbandingan sulit dilakukan bila tidak ada dasar empiris yang dapat diandalkan dari efek populasi nyata; dan konvensi global tidak berguna ketika perbedaan antara sub-disiplin dan antara desain studi begitu dramatis. Temuan peneliti menunjukkan bahwa efek dalam penelitian psikologis mungkin jauh lebih kecil daripada yang muncul pada publikasi terdahulu yang memiliki implikasi menguntungkan dan merugikan. Ketika menafsirkan efek dari satu penelitian, tentu saja baik untuk mengetahui bahwa banyak efek agak kecil dan karenanya efeknya sendiri tidak menonjol. Pada sisi negatifnya, effect size yang lebih kecil memperlihatkan bahwa studi psikologi menjadi kurang optimal bahkan lebih dramatis daripada yang baru-baru ini. Karena efek populasi yang lebih kecil akan membutuhkan sampel yang lebih besar untuk menghasilkan signifikansi statistik. Dengan demikian, temuan peneliti sekali lagi menggarisbawahi perlunya perhitungan kekuatan dalam penelitian psikologis untuk menghasilkan pengetahuan yang andal.</p>
	<p>Carl, E., Stein A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers M. B. (2018). Virtual reality exposure therapy for anxiety and related disorders: A metaanalysis of randomized controlled trials. <i>Journal of Anxiety Disorders</i>. https://doi.org/10.1016/j.janxdis.2018.08.003</p>	<p>Efect size antar-kelompok untuk setiap studi dihitung menggunakan g Hedge. Hedge's g secara unik cocok untuk memberikan perkiraan effect size yang tidak bias ketika ukuran sampel kecil. Ketika data yang diperlukan tersedia, effect size dihitung secara langsung menggunakan rumus ini: $g = \bar{x}_T - \bar{x}_C = sp$, di mana \bar{x}_T adalah rerata kelompok perlakuan, \bar{x}_C adalah rerata kelompok pembanding, dan sp adalah deviasi standar gabungan. Ketika data ini tidak tersedia, g diperkirakan menggunakan persamaan konversi untuk uji signifikansi. Sebuah effect size gabungan dihitung untuk studi menggunakan beberapa ukuran hasil. Hedge's g dapat diinterpretasikan sama dengan Cohen's d,</p>	<p>Meta-analisis dari 30 percobaan VRET (peserta $N = 1,057$) mendukung hipotesis penelitian. Konsisten dengan hipotesis 1, VRET menunjukkan effect size yang besar dibandingkan dengan kondisi daftar tunggu dan effect size sedang hingga besar dibandingkan dengan kontrol psikologis. Konsisten dengan hipotesis 2, VRET tidak secara signifikan lebih atau kurang efektif daripada paparan in vivo. Analisis di setiap gangguan menunjukkan bahwa effect size ini relatif konsisten. Sebagai kesimpulan, meta-analisis ini menunjukkan bahwa, untuk sejumlah gangguan terkait kecemasan, VRET memiliki effect size yang besar dibandingkan dengan kondisi kontrol dan effect size yang sama dengan paparan in vivo. Hasil ini konsisten untuk gangguan yang berbeda, dengan effect size sedang atau besar untuk VRET dibandingkan dengan kontrol untuk fobia spesifik, SAD dan kecemasan kinerja, PTSD, dan PD. Dalam kasus di mana VRET lebih disukai atau lebih mudah diakses, VRET dapat dianggap sebagai alternatif yang dapat diterima dan berkhasiat untuk</p>

dengan cutoff untuk efek kecil (0,2), sedang (0,5) dan besar (0,8). Effect size rata-rata keseluruhan untuk studi gabungan dihitung menggunakan rumus tertimbang, $= w_j g_j / \sum w_j g_j$, di mana w_j adalah bobot untuk setiap studi dan g_j adalah effect size untuk setiap studi. Analisis efek acak digunakan, yang mengasumsikan bahwa studi yang disertakan hanya sampel dari seluruh populasi studi.

paparan in vivo untuk pengobatan gangguan terkait kecemasan.

PEMBAHASAN

Effect size merupakan salah satu uji statistik yang mengukur sejauh mana suatu variabel mempengaruhi variabel lain dalam suatu penelitian atau menunjukkan seberapa efektif suatu variabel mempengaruhi variabel lainnya. *Effect size* merupakan ukuran mengenai signifikansi praktis hasil penelitian yang berupa ukuran besarnya korelasi atau perbedaan, atau efek dari suatu variabel pada variabel lain (Santoso, 2010). Selain itu,

Effect size juga dapat menjadi salah satu tindakan mediasi statistik terbaik. Solusi analitik untuk kesalahan standar dari setiap effect size untuk efek mediasi dapat mengkonfirmasi dan menjelaskan temuan dari studi simulasi dan memperkuat rekomendasi tentang kegunaan effect size tertentu dalam meta-analisis model mediasi. Namun, interval kepercayaan berdasarkan distribusi normal yang diasumsikan untuk ukuran effect size dan rumus analitik yang sesuai untuk kesalahan standar ukuran effect size mungkin tidak akurat, sehingga solusinya adalah menggunakan metode bootstrap ideal. Situasi ini menyoroti pentingnya interpretasi untuk pemilihan ukuran efek. Ketika memilih effect size untuk mewakili efek tidak langsung, seseorang harus memilih effect size yang menjawab pertanyaan penelitian dengan paling akurat (Miočević et al., 2018).

Berdasarkan hasil studi literatur review, beberapa metode perhitungan yang dapat digunakan dalam mengukur *effect size*, diantaranya yaitu adalah g Hedge, karena penggunaannya yang cocok untuk memberikan perkiraan ukuran efek yang tidak bias ketika ukuran sampel kecil. (Carl et al., 2019) dan penggunaan Bootstrap CI untuk melakukan penskalaan ukuran efek yang tidak memenuhi standar untuk mempertahankan unit pengukuran asli (Kirby & Gerlanc, 2013).

SIMPULAN

Effect size adalah bagian dari uji statistik yang mudah yang berperan dalam membantu pengguna memahami besarnya perbedaan yang ditemukan dalam suatu percobaan penelitian.

DAFTAR PUSTAKA

- Anam, C. (2020). Jenis uji statistik untuk analisis hasil penelitian. *Study*, 23(4), 115–117. https://ejournal.undip.ac.id/index.php/berkala_fisika/article/view/34495
- Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials.

Journal of Anxiety Disorders, 61, 27–36.

<https://doi.org/10.1016/j.janxdis.2018.08.003>

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Science*. US: Lawrence, Erlbaum.

Kirby, K. N., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods*, 45(4), 905–927.

<https://doi.org/10.3758/s13428-013-0330-5>

Miočević, M., O'Rourke, H. P., MacKinnon, D. P., & Brown, H. C. (2018). Statistical properties of four effect-size measures for mediation models. *Behavior Research Methods*, 50(1), 285–301.

<https://doi.org/10.3758/s13428-017-0870-1>

Santoso, A. (2010). Studi Deskriptif Effect Size Penelitian-Penelitian Di Fakultas Psikologi Universitas Sanata Dharma. *Jurnal Penelitian, Fakultas Psikologi Universitas Sanata Dharma, Yogyakarta*, 14(1), 17.

<http://repository.usd.ac.id/id/eprint/9419>

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10(APR), 1–13.

<https://doi.org/10.3389/fpsyg.2019.00813>