

BAB III

METODE PENELITIAN

3.1 Tinjauan Meta-Analisis

Istilah "meta-analisis" pada awalnya diciptakan oleh Gene V. Glass pada tahun 1976 (Glass, 1976). Dalam kaitannya dengan analisis yang ada, Glass mengklasifikasikan analisis penelitian menjadi analisis primer (studi primer), analisis sekunder (studi sekunder), dan meta-analisis. Analisis primer adalah analisis asli dari data mentah. Analisis sekunder menggunakan teknik yang sama untuk menjawab pertanyaan yang berbeda dari data yang sama.

Referensi Glass untuk meta-analisis sebagai "analisis analisis" (Glass, 1976) telah dijelaskan oleh (Kulik & Kulik, 1988) sebagai "definisi paling ringkas yang belum diajukan untuk metodologi ini". Kulik lebih lanjut menyoroti beberapa karakteristik Glass dari meta-analisis, sebagai berikut:

1. Sebuah meta-analisis meliputi menganalisis hasil studi yang sudah dilakukan.
Glass tidak menggunakan istilah tersebut untuk merujuk pada analisis serangkaian investigasi yang direncanakan.
2. Sebuah meta-analisis diterapkan pada ringkasan statistik, bukan data mentah.
Ini termasuk cara, deviasi standar, dan hasil dari uji statistik.
3. Sebuah meta-analisis berfokus pada ukuran efek perlakuan, bukan hanya signifikansi statistik.
4. Sebuah meta-analisis berfokus pada fitur dan hasil studi. Tujuan dari meta-analisis tidak hanya untuk merangkum seluruh literatur dengan ukuran efek

tunggal atau tingkat signifikansi keseluruhan tetapi juga untuk menentukan bagaimana fitur studi mempengaruhi ukuran efek.

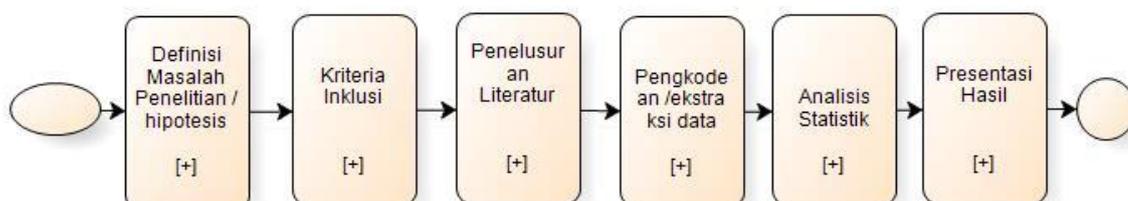
Metode khas sebelum pengembangan meta-analisis adalah tinjauan pustaka dimana studi ditinjau dan hasil dibandingkan untuk sampai pada kesimpulan umum mengenai pertanyaan penelitian. Karena studi yang berbeda terkadang memiliki hasil yang berbeda bahkan kontradiktif, maka penarikan kesimpulan tentang pertanyaan penelitian umum dapat bersifat subjektif (Franzen, 2018). Untuk itu, meta-analisis dikembangkan sebagai alat untuk mengevaluasi dan membandingkan hasil di antara berbagai penelitian. Ini dikembangkan untuk memberikan metode yang lebih obyektif untuk menarik kesimpulan ketika ada banyak penelitian dengan hasil yang bervariasi (Cooper et al. 2009). Meta-analisis memberikan evaluasi keseluruhan dengan analisis statistik data kuantitatif yang diperoleh dalam studi independen pada subjek tertentu (Glass, 1976; Shelby & Vaske, 2008; Schwarzer, Carpenter, & Rücker, 2015; Cleophas & Zwinderman, 2017).

Unit analisis dalam penelitian meta-analisis adalah ukuran efek (Borenstein et al., 2009). Dalam istilah dasar, ukuran efek dapat digambarkan sebagai angka yang mencerminkan besarnya hubungan antara dua variabel (Cooper & Hedges, 2009). Ukuran efek merupakan cara sederhana untuk mengukur perbedaan antara dua kelompok yang memiliki banyak keunggulan dibandingkan dengan penggunaan uji signifikansi statistik saja (Coe, 2002; Thalheimer & Cook, 2002; Ellis, 2010). Ukuran efek adalah nilai yang mencerminkan besarnya efek perlakuan (Borenstein et al., 2009). Ini dianggap

sebagai praktik terbaik untuk menggambarkan besarnya pengaruh dari suatu perlakuan dalam semua penelitian kuantitatif (Brendel, 2011).

3.2 Desain Penelitian

Desain penelitian ini adalah meta-analisis yang menggabungkan dua penelitian utama atau lebih tentang pengaruh penggunaan CAME terhadap kemampuan matematis siswa, yang telah dipublikasikan untuk mengakumulasikan temuan. Secara umum ada 5 tahapan dalam meta-analisis yaitu perumusan masalah, pengumpulan data, evaluasi data, analisis dan interpretasi data, dan presentasi hasil (Cooper, 1982). Proses meta-analisis yang lebih rinci dapat ditemukan dalam literatur dan mencakup 7 langkah (Cooper, Patall, Lindsay, Bickman, & Rog, 2009). Namun beberapa langkah dapat digabungkan menjadi satu langkah proses komprehensif dan dipandang secara terperinci sebagai bagian dalam subproses. Model proses sederhana berikut menunjukkan ikhtisar metodologi meta-analisis dan dibuat dengan model NVIVO 10:



Gambar 3.1: Model Proses meta-analisis: Gambaran Umum

Gambar 3.1 menunjukkan semua tahap dalam penelitian meta-analisis, dengan setiap kegiatan yang bernotasi [+], berarti setiap langkah proses dapat diperluas ke subproses lebih lanjut.

3.3 Kriteria Inklusi

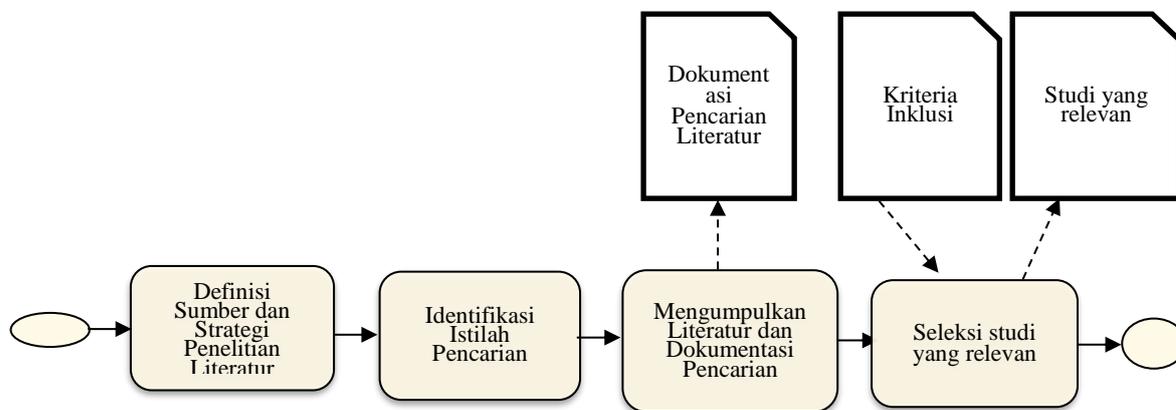
Kriteria berikut digunakan untuk menentukan apakah penelitian akan dimasukkan dalam meta-analisis atau dikeluarkan dari analisis. Sesuai dengan tujuan penelitian maka kriteria inklusi dalam penelitian yaitu:

- 1) Jenis Studi: Studi yang dianalisis mengevaluasi pembelajaran matematika tentang pengaruh CAME pada kemampuan matematis siswa.
- 2) Jenis peserta: Penelitian ini menganalisis studi utama tentang pengaruh CAME yang dilakukan di jenjang pendidikan menengah atas (SMA).
- 3) Jenis pengaturan: Studi ini mencakup semua studi utama yang telah dilakukan di seluruh Indonesia dan sudah dipublikasikan pada tahun 2010 atau lebih baru.
- 4) Jenis desain: Studi yang dimasukkan dalam analisis primer yang menggunakan kelompok kontrol sebagai pembanding. Studi tanpa kelompok kontrol, dikeluarkan. Studi yang tidak memuat informasi statistik untuk transformasi ukuran efek, dikeluarkan.

3.4 Penelusuran Studi

Tujuan dari penelusuran studi adalah pengambilan lengkap dari studi yang relevan mengenai topik penelitian meta-analisis tanpa menciptakan bias oleh kebiasaan penelitian literatur peneliti (White, 2009). Pada akhir langkah

ini, studi yang memenuhi kriteria relevansi dipilih untuk langkah pengkodean selanjutnya. Gambar 3.2 menunjukkan proses yang lebih rinci dari “Pencarian dan Pengumpulan Literatur”:



Gambar 3.2: Proses Pencarian dan Pengumpulan Literatur
(Diadaptasi dari Cooper, 2010 and Schewe et al., 2014)

Berdasarkan Gambar 3.2, penelusuran studi dimulai dengan sub tugas “Definisi Sumber dan Strategi Penelitian”. Untuk tiga strategi terakhir dapat diidentifikasi: 1. penelitian dengan istilah pencarian dalam database elektronik, 2. menggunakan referensi dari laporan penelitian yang sudah diperoleh (juga disebut "pengejaran catatan kaki"), dan 3. penyelidikan sistematis dari penelitian yang relevan. Strategi-strategi ini berurutan, ini berarti setelah pengambilan pertama sumber dari strategi (1), daftar referensi dapat digunakan sebagai dasar untuk mencapai studi (2), sampai mencapai kriteria berhenti yang ditentukan (misalnya studi menjadi terlalu tua) dan akhirnya kontak langsung dengan peneliti lain untuk secara langsung meminta studi yang sesuai atau set data individu yang disebutkan dalam studi (3).

Sumber tambahan adalah, misalnya ulasan penelitian (kuantitatif atau kualitatif), referensi dalam literatur yang diambil, database elektronik dan *world wide web*, jurnal terkait topik, program konferensi, dan lain-lain (Lipsey & Wilson, 2003). Prosedur pengumpulan literatur dilakukan dengan menggunakan sumber yang ditentukan dan strategi pencarian yang direncanakan. Untuk menutup langkah penelitian literatur, semua literatur yang dikumpulkan dievaluasi oleh kriteria inklusi yang ada di akhir langkah satu, sehingga peneliti meta-analisis memutuskan studimana yang memenuhi syarat. Hasilnya adalah kumpulan studi yang akan dikodekan pada langkah proses besar berikutnya (Schewe et al., 2014).

Menurut Strohmeier, (2015) tantangan pencarian literatur adalah bagaimana menghindari bias. Sumber *peer review*, khususnya, jurnal ilmiah, cenderung menderita bias publikasi (Rothstein & Hopewell, 2009). Bias terhadap temuan nol mencirikan fakta bahwa artikel yang signifikan secara statistik memiliki peluang lebih tinggi untuk dipublikasikan dan bahwa para peneliti juga jarang (6%) mencoba untuk mempublikasikan penelitian yang tidak signifikan. Bias konfirmasi menggambarkan bahwa hasil yang tidak sesuai dengan keyakinan penelitian yang berlaku juga lebih kecil kemungkinannya untuk dipublikasikan, bahkan jika hasilnya signifikan secara statistik (Cooper, 2010). Hal ini mengarah pada representasi yang berlebihan dari dan pada gilirannya sama dengan hilangnya studi yang sebenarnya ada (Rothstein & Hopewell, 2009). Namun, literatur *peer review* dapat dilihat

sebagai sumber berkualitas tinggi dan harus menjadi bagian dari pencarian literatur, tetapi bukan satu-satunya sumber (Cooper, 2010).

Dalam penelitian ini pencarian literatur yang menyeluruh dilakukan, dengan menggunakan beberapa sumber sehingga sumber yang berbeda dapat saling melengkapi. Selain itu strategi pencarian memastikan pencarian literatur yang lengkap dengan: pencarian basis data referensi; peninjauan cermat jurnal yang relevan; inspeksi referensi dalam penelitian primer dan sintesis penelitian yang ada; dan kontak langsung dengan "peneliti aktif dan terkemuka" lainnya.

Penting untuk diperhatikan bahwa dalam penelitian meta-analisis pencarian literatur didekati secara sistematis dalam upaya untuk mencapai semaksimal mungkin studi terkait baik yang telah diterbitkan maupun yang belum. Menurut Lipsey dan Wilson (2001), pengecualian pencarian dan termasuk kemungkinan akan mengarah pada bias ke atas dalam ukuran efek. Pencarian mendalam untuk studi dan penelitian dicari menggunakan kombinasi kata kunci "Kata kunci yang dimasukkan: {Pengaruh (memasukkan salah satu dari jenis software berikut secara bergantian: GeoGebra, Cabri, Autograph, Maple, Wingeom, Matlab, Macromedia Flash, Math, Mathematica) terhadap kemampuan matematis}.

3.4.1 Basis Data Elektronik

Basis data elektronik yang dicari termasuk ERIC (*Education Resources Information Center*), Sage, dan Elsevier (pencarian terakhir dilakukan Desember 2020).

3.4.2 Kontak Pribadi

Tidak dapat dipungkiri bahwa ada informasi penting yang terkait dengan penelitian ini tetapi tidak ditemukan dalam studi primer. Untuk mengantisipasi hal tersebut, penulis harus mengupayakan komunikasi dengan para penulis asli. Dalam penelitian ini para penulis dihubungi dengan mengirim email atau kontak langsung melalui telepon untuk meminta atau melengkapi data yang diperlukan.

3.4.3 Pencarian Tangan

Pencarian tangan dilakukan untuk memverifikasi informasi yang tidak lengkap dari artikel jurnal yang sebelumnya telah diperoleh. Sebagian besar studi yang tidak memuat informasi lengkap adalah dari artikel jurnal yang merupakan produk tesis dan disertasi Universitas Pendidikan Indonesia. Jadi pencarian tangan dalam penelitian ini dilakukan dengan mengunjungi perpustakaan Universitas Pendidikan Indonesia.

3.4.4 Pencarian Internet

Pencarian kata kunci (sebagaimana disebutkan di atas) dilakukan menggunakan [googlescholar.com](https://scholar.google.com), dan portal Garuda.

3.4.5 Daftar Referensi

Daftar referensi studi yang ditemukan relevan untuk tinjauan ini serta studi terkait dari meta-analisis sebelumnya diperiksa untuk sumber data yang relevan digunakan lebih lanjut.

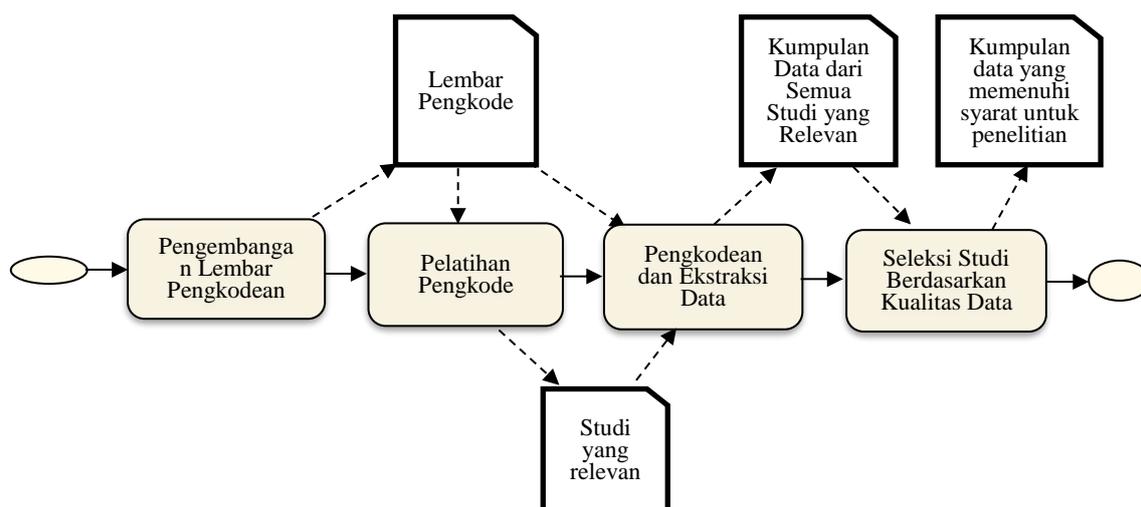
3.5 Koding dan Ekstraksi Data

3.5.1 Instrumen Penelitian

Instrumen dalam meta-analisis ini dilakukan dengan lembaran pemberian kode (*coding category*). Formulir pengkodean dibuat mengikuti karakteristik yang sudah ditetapkan sesuai karakteristik studi. Selain itu formulir pengkodean juga dilengkapi dengan informasi statistik sesuai dengan rumus yang dipilih.

Proses Pengkodean atau ekstraksi data adalah komponen penting dari meta-analisis dan "sering kali bisa rumit, membosankan dan memakan waktu" (Schmidt & Hunter, 2015). Namun langkah ini sangat penting untuk kualitas meta-analisis, sehingga harus dijalankan dengan hati-hati untuk mengesampingkan potensi kesalahan atau bias (Orwin & Vevea, 2009).

Untuk melakukan tugas ini, lembar pengkodean dikembangkan dan divalidasi oleh ahli. Selanjutnya pengkode dapat mengekstraksi dan mengkodekan data yang ditemukan dalam penelitian berdasarkan kualitas data yaitu studi yang memenuhi syarat yang telah ditetapkan. Gambar 3.3 berikut menunjukkan proses ini:



Gambar 3.3: Proses Pengkodean dan Ekstraksi Literatur Terpilih
(Diadaptasi dari Cooper, 2010)

Dalam penelitian ini lembar pengkodean terdiri dari "formulir pengkodean" dan "manual pengkodean". Formulir pengkodean adalah instrumen penelitian dimana studi diberi kode (Wilson, 2009). Manual pengkodean memberikan panduan "tentang cara menerapkan item formulir pengkodean untuk studi" (Wilson, 2009). Draf manual pengkodean pertama mencakup daftar karakteristik studi yang harus diisi dari setiap studi dan ditambah dengan informasi statistik yang diperlukan dalam perhitungan ukuran efek. Draf manual pengkodean ini harus ditinjau oleh kolega atau pembimbing yang berpengalaman luas untuk memberikan masukan tentang pertanyaan pengkodean lebih lanjut, kemungkinan nilai jawaban dan juga untuk menyoroiti ketidakjelasan yang akan diklarifikasi oleh meta-analisis.

Dalam penelitian ini lembar pengkodean mencakup kategori informasi berikut:

1. Informasi pengkode yaitu identitas berupa nama, pendidikan, pekerjaan dan email pengkode.
2. Petunjuk pengkodean yaitu pernyataan yang mengarahkan pengkode untuk mengekstraksi hasil studi kedalam kormat koding data.
3. Koding data berupa tabel yang isinya nama penulis, judul penelitian, informasi statistik yang diperlukan dalam menentukan ukuran efek dan karakteristik studi yaitu jenis CAME, rasio penggunaan komputer, variabel dependen, ukuran sampel, peran instruksi komputer, durasi

perlakuan, jenjang sekolah, kelas penelitian, tahun penelitian/studi, materi pembelajaran, dan jurusan yang dipilih.

Berdasarkan kategori informasi tersebut Tabel 3.1 berikut menyajikan informasi tentang karakteristik studi dan kelompok studi yang selanjutnya akan dianalisis dalam penelitian ini:

Tabel 3.1 Informasi Studi

Variabel Moderator	Kategori
Jenis CAME	CAS
	DGS
	CAS dan DGS
Ukuran Sampel	Ne=30 atau kurang
	Ne = 31 atau lebih
Rasio Penggunaan Komputer (RPK)	Individu
	Rombongan
	Tidak ditentukan
Peran Instruktur Komputer	Pengganti
	Pelengkap
	Tidak ditentukan
Durasi Perlakuan	$t \leq 4p$
	$4p < t \leq 7 p$
	$t \geq 7 p$
	t tidak ditentukan
Jurusan	IPA
	IPS
	Belum ditentukan
Kemampuan yang diukur	Analogi Matematis
	Berpikir Matematis
	Komunikasi matematis
	Koneksi matematis
	Pemahaman Matematis
	Pemecahan Masalah
	Penalaran Matematis
	Representasi Matematis
Spasial matematis	
Materi	Bangun Datar
	Bagun Ruang
	Fungsi
	Integral
	Sistem Pertidaksamaan
	Trigonometri
	Statistik

	Grafik Fungsi
	Limit Fungsi
	Program Linear
	Lingkaran
	Persamaan Linear
	Transformasi geometri
Kelas Penelitian	X
	XI
	XII
Tahun Publikasi	2010-2012
	2013-2015
	2016-2018
	2019-2020

Berdasarkan Tabel 3.1 tampak bahwa jumlah variabel moderator yang dianalisis dalam penelitian adalah 10 dengan 37 kategori. Perbandingan ES antar kategori dilakukan untuk memperjelas peran moderator terhadap ES studi.

Tidak dapat dipungkiri bahwa selama proses pengkodean, banyak masalah yang harus ditangani. Terutama ketika jumlah studi banyak, membaca ulang semua studi untuk mengambil data tambahan bisa menjadi "sangat memakan waktu" (Cooper, 2010). Secara umum keandalan data yang dikodekan harus dinilai. Cooper (2010) mengidentifikasi beberapa masalah seperti kesalahan perekaman, laporan tidak jelas - pertanyaan pengkodean tidak dapat dijawab, definisi yang ambigu atau bahkan kecenderungan pembuat kode mengenai topik penelitian. Untuk melucuti kelemahan tersebut dalam penelitian ini dapat dilakukan beberapa langkah:

a. Penggunaan manual pengkodean

Peneliti menyiapkan manual pengkodean yang selanjutnya digunakan oleh pengkode. Penggunaan manual pengkodean ini dilakukan untuk menjaga konsistensi pengkodean dengan memberikan penjelasan yang lebih rinci

tentang pertanyaan pengkodean, jawaban dan contohnya. Dalam penelitian ini manual pengkodean divalidasi oleh pembimbing dan pihak lain yang dianggap ahli dalam penelitian meta-analisis.

b. Pengodean dalam tim

Studi harus dikodekan ganda (oleh setidaknya 2 pengkode) untuk menilai tingkat kesepakatan (Cooper, 2010; Wilson, 2009). Kemudian pengkodean dibandingkan sehingga tidak ditemukan tumpang tindih atau salah pengisian data.

c. Seleksi dan pelatihan Para Pengkode

Pemilihan pengkode diperlukan untuk memastikan mereka memiliki pengetahuan yang cukup dalam domain penelitian. Selanjutnya pelatihan diperlukan untuk memastikan bahwa hasil koding akurat dan konsisten dengan para pengkode lain (Wilson, 2009). Dalam melatih para pembuat kode Syed & Nelson (2015) menyarankan untuk mengikuti prosedur tiga langkah. Pertama, peneliti harus menyediakan manual pengkodean untuk semua pembuat kode. Manual ini harus dibahas secara rinci dan semua pertanyaan awal ditangani. Kedua, peneliti harus menyediakan data sampel yang diambil secara acak dari kumpulan data, yang dengannya pembuat kode dapat mempraktikkan skema pengkodean. Kode awal ini harus didiskusikan secara menyeluruh dengan tim pengkodean, dan proses pengecualian dan penyertaan harus dirinci. Pada tahap ini, manual pengkodean sering direvisi untuk mencerminkan pertanyaan umum yang diajukan kepada peneliti oleh pembuat kode untuk memberikan dokumentasi dari setiap keputusan yang dibuat pada tahap pelatihan awal.

Ketiga, setelah mencapai pemahaman tentang manual pengkodean, peneliti dapat mulai menetapkan sebagian kecil dari kumpulan data (dipilih secara acak) untuk pengkodean. Setelah penilai cukup terlatih, mereka dapat diberikan kasus untuk dikodekan secara independen. Akhirnya, dan mungkin yang paling penting, peneliti harus membuat keputusan tentang proses pengkodean, persentase yang akan dikodekan untuk mencapai reliabilitas, koefisien reliabilitas yang dibutuhkan, dan setiap keputusan tentang siapa yang akan membantu proses pengkodean.

3.5.2 Uji Reliabilitas

Semua studi yang memenuhi syarat diberi kode menggunakan skema pengkodean terperinci berdasarkan manual pengkodean dan protokol koding. Manual dan protokol pengkodean divalidasi oleh pembimbing dan ahli (lihat Lampiran A). Setelah manual pengkodean dan protokol koding dikembangkan penulis meminta penilaian dari tim ahli di luar pembimbing. Mereka adalah Prof. Ying Zhou dari Guangxi Normal University, China; Prof. Sule Bayraktar dari *Faculty of Education*, Giresun University, Giresun, Turkey; Assoc Prof. Sedat Turgut dari Bartın University, Turkey; dan Prof. Terri Deocampo Pigott dari Georgia State University, USA. Penulis mengirimkan instrumen tersebut melalui email dari masing-masing pakar.

Dari keempat pakar tersebut hanya Prof. Ying Zhou dari Guangxi Normal University China yang bersedia. Kenyataan ini tidak sesuai dengan harapan bahwa penilai minimal dua orang yang dianggap pakar dalam bidang

meta-analisis diluar pembimbing. Peneliti kemudian meminta bantuan Prof. Ying Zhou untuk merekomendasikan pakar lain yang dianggap dapat menilai instrumen penelitian ini. Prof. Ying Zhou merekomendasikan asistennya yaitu Tommy Tanu Wijaya M.Ed dari Guangxi Normal University China. Keduanya mengirimkan hasil penilaian secara bersamaan pada bulan awal Januari 2021 (lihat lampiran A2 halaman 169).

Skala penilaian dalam instrumen yang dinilai terdiri dari 1-5. Nilai 1 berarti sangat tidak baik (tidak dapat digunakan), nilai 2 berarti tidak baik (juga dapat digunakan dengan revisi banyak), nilai 3 berarti cukup baik (dapat digunakan dengan revisi sedang), nilai 4 berarti baik (dapat digunakan dengan revisi sedikit), nilai 5 berarti sangat baik (dapat digunakan tanpa revisi).

Dari hasil pemeriksaan, penilai pertama memberikan rata-rata skor 4,5 dan penilaian kedua 4,25. Rata-rata penilaian dari kedua pakar adalah 4,4 yang berarti instrumen penelitian ini hampir sempurna. Meskipun demikian peneliti merevisinya sesuai masukan dari kedua penilai.

Manual pengkodean disempurnakan melalui proses berulang dimana masalah-masalah yang diangkat dalam tinjauan, dan informasi baru dari studi utama yang diperoleh dalam proses pengkodean awal digunakan untuk menambah, menghapus, atau mengklarifikasi variabel yang menarik. Misalnya, pada tahap awal pengkodean, tidak jelas apakah penting untuk memberi kode tentang perbedaan jurusan siswa saat penelitian dalam studi primer (misalnya, jurusan IPA, IPS, dan kelompok yang belum dibagikan). Karakteristik tentang

jurusan dengan tiga kategori dijadikan sebagai salah satu moderator meskipun rancangan awalnya itu belum dipikirkan.

Instrumen penelitian adalah bentuk pengkodean yang telah dikembangkan untuk mengekstrak informasi dari studi individu menjadi data numerik. Informasi ini mencakup: nama studi, tahun penelitian, informasi statistik, jenis CAME, rasio penggunaan komputer, ukuran sampel, peran instruksi komputer, durasi perlakuan, kelas penelitian, jurusan, materi pembelajaran, dan variabel dependen yang diukur.

Dengan menggunakan lembar ini, informasi ukuran variabel dan efek dikodekan secara terpisah untuk setiap studi oleh dua orang yang sebelumnya dilatih secara khusus. Keduanya adalah mahasiswa program magister pendidikan matematika universitas pendidikan Indonesia. Untuk menentukan reliabilitas, sampel acak 5 dari banyaknya studi yang memenuhi syarat diduplikasi dan didistribusikan di antara para pembuat kode. Setiap pembuat kode diberi salinan artikel, formulir pengkodean, dan protokol pengkodean. Untuk menilai keandalan antar pengkode, digunakan tingkat kesepakatan persen persetujuan (*percent agreement*) atau yang disingkat PA. Metode PA adalah pendekatan yang paling mudah dan intuitif untuk membangun keandalan (Syed & Nelson, 2015). PA hanyalah rasio item dimana dua pembuat kode setuju dengan jumlah total item yang dinilai, dan dihitung menggunakan rumus berikut:

$$PA = \frac{N_A}{N_A + N_P} \times 100$$

dimana N_A adalah jumlah total kesepakatan dan N_P adalah jumlah total ketidaksepakatan. Tingkat kesepakatan 85% atau lebih besar ditentukan sebelumnya untuk dianggap tinggi.

Instrumen yang telah divalidasi selanjutnya digunakan oleh dua pengkode yang telah diseleksi dan dilatih yaitu Suparman, dan Yunita. Keduanya adalah mahasiswa program magister pendidikan matematika Universitas Pendidikan Indonesia. Gambar 4.2 proses pelatihan menggunakan aplikasi zoom.

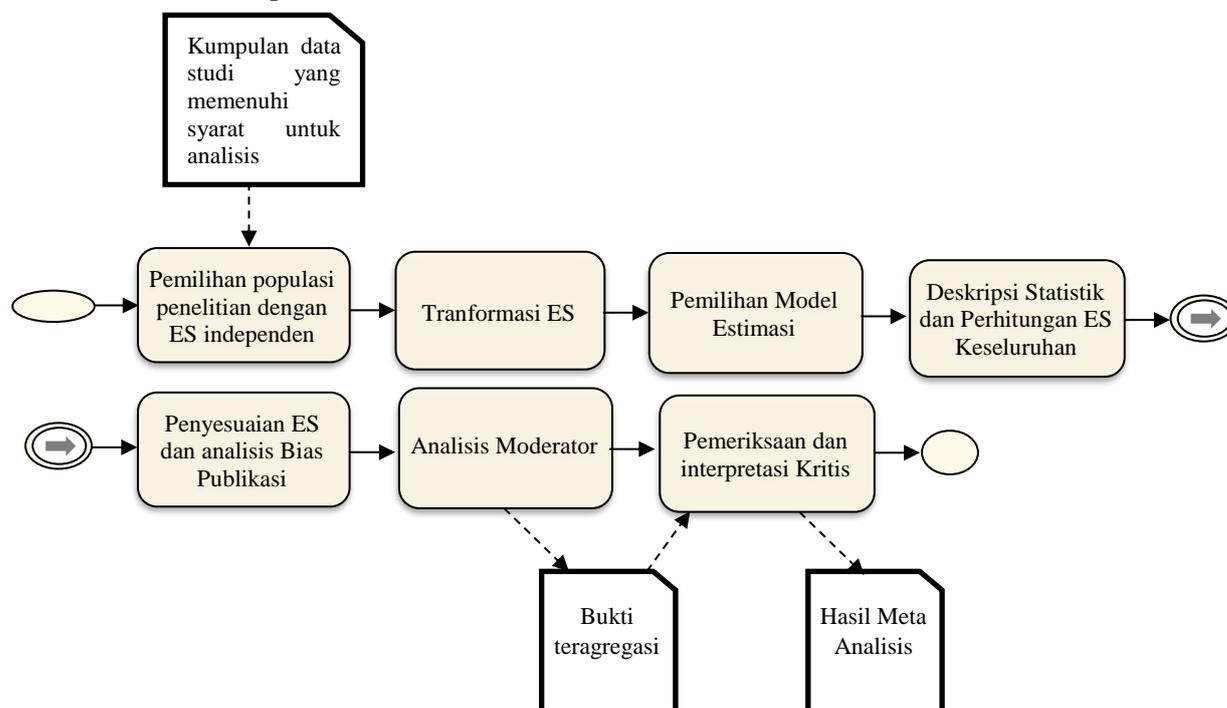


Gambar 3.4 Proses pelatihan pengkode

Selanjutnya uji coba instrumen untuk menentukan reliabilitasnya dilakukan dengan memilih sampel acak 5 dari 20 studi individu yang sebelumnya telah ditemukan melalui pencarian awal menggunakan basis data elektronik. Ke-lima studi tersebut diduplikasi dan didistribusikan di antara para pembuat kode. Dari hasil perhitungan diperoleh tingkat PA keduanya adalah 87,27% (perhitungan lengkapnya lihat di lampiran A3 halaman 172). Angka tersebut menunjukkan ada kecocokan substansial hingga hampir sempurna antara pembuat kode. Dengan demikian instrumen yang akan digunakan dalam penelitian ini dapat diandalkan.

3.6 Analisis Statistik

Analisis statistik merupakan identifikasi dan penerapan metode untuk secara statistik mengintegrasikan hasil studi dan menafsirkannya. Gambar 3.4 berikut merupakan alur analisis statistik.



Gambar 3.5 Proses Analisis Statistik Metode Meta-analisis (Berdasarkan Borenstein et al., 2009)

3.6.1 Populasi Penelitian

Populasi dalam penelitian ini adalah seluruh studi yang dilakukan di Indonesia yang menyelidiki pengaruh CAME dalam pembelajaran matematika jika dibandingkan dengan pendekatan konvensional selama 10 tahun terakhir. Sampel dari penelitian ini diambil dari populasi yang dipilih dengan teknik *Purposive Sampling* yaitu teknik penentuan sampel dengan pertimbangan tertentu. Pemilihan studi dalam purposive sampling, didasarkan atas kriteria

inklusi yang diterapkan berdasarkan tujuan penelitian atau permasalahan penelitian.

3.6.2 Transformasi *Effect Size* (ES)

Sebagaimana halnya studi primer yang menganalisis data dari hasil observasi, meta-analisis juga mengolah data. Data tersebut dapat dinyatakan dengan berbagai ukuran yang dihitung dengan formula yang dinyatakan dengan persamaan matematika. Ukuran tersebut dinamakan sebagai ES.

Dalam meta-analisis, perhitungan ES dianggap sebagai inti dari studi karena menentukan temuan inti dari studi yang diminati (Borenstein, 2009). Penting untuk memeriksa ES, karena mereka menggambarkan tingkat efektivitas studi yang bersangkutan. Model proses sebagaimana yang ditunjukkan pada gambar 3.4 menunjukkan analisis statistik metode meta-analisis dengan fokus pada ES).

Perhitungan ES menggunakan indikator perbedaan rata-rata terstandarisasi. Rumus berikut tersedia untuk perhitungan skor terstandarisasi ini, yaitu, Cohen's *d*, Glass's '*s*', dan Hedges' *g* (Fern & Monroe, 1996)

$$\text{Cohen's } d = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{pooled}}; \quad \text{Glass's } \Delta = \frac{\bar{X}_1 - \bar{X}_2}{S_c}; \quad \text{Hedges's } g = \frac{\bar{X}_1 - \bar{X}_2}{S_{pooled}};$$

Keterangan:

\bar{X}_1 : Rata-rata kemampuan matematis pada kelompok eksperimen

\bar{X}_2 : Rata-rata kemampuan matematis pada kelompok kontrol

σ_{pooled} : Deviasi standar gabungan dari populasi

S_c : Deviasi standar kontrol

S_{pooled} : Deviasi standar gabungan

Penggunaan rumus di atas disesuaikan dengan ketersediaan informasi statistik dari studi primer. Penggunaan Cohen's d jarang dimungkinkan karena σ_{pooled} jarang dilaporkan dalam studi primer. Pilihannya adalah antara rumus Glass atau Hedges' g . Dapat dilihat bahwa satu-satunya perbedaan antara keduanya terletak pada penyebutnya.

Dalam penelitian ini metrik utama untuk penghitungan ukuran efek dalam ulasan ini adalah Hedges' g karena memiliki koreksi bawaan untuk bias untuk ukuran sampel kecil (Borenstein et al., 2009; Cooper & Hedges, 2009). Metrik Hedges' g dikategorikan sebagai perbedaan rata-rata standar (*standardized mean difference*) sehingga memungkinkan untuk memasukan skor dengan skala yang berbeda. Ini relevan dengan kenyataan bahwa penelitian primer tentang pengaruh CAME dalam pembelajaran matematika sering melaporkan skor kemampuan matematis dalam berbagai skala.

Untuk studi yang melaporkan statistik nilai t , F , atau p , rumus konversi seperti Hedges' g akan digunakan untuk menghitung perbedaan rata-rata standar untuk perkiraan ukuran efek. Dalam penelitian ini perhitungan ukuran efek rata-rata dan pengujian hipotesis menggunakan perangkat lunak CMA. ES yang digunakan adalah Hedges' g sesuai dengan informasi statistik yang umumnya ditemukan. Interpretasi ukuran efek, menggunakan klasifikasi Cohen et al. (2018) yaitu:

Tabel 3.2 Interpretasi Ukuran Efek

Range of Effect Size (ES)	Interpretasi
$ES \leq 0,20$	Efek lemah (<i>weak effect</i>)

$0,20 < ES \leq 0,50$	Efek sederhana (<i>modest effect</i>)
$0,50 < ES \leq 1,00$	Efek sedang (<i>moderate effect</i>)
$ES > 1,00$	Efek kuat (<i>strong effect</i>)

ES menginformasikan seberapa kuat hubungan antara dua variabel (Cooper, 2010; Lipsey & Wilson, 2003). Lebih lanjut menggunakan ES memungkinkan perbandingan numerik dan dengan demikian analisis satu set studi dengan konsep standarisasi (Lipsey & Wilson, 2003).

Beberapa ketentuan yang harus diperhatikan untuk mencapai hasil statistik yang valid adalah kemandirian statistik dari studi yang diintegrasikan. Pertama, harus dipastikan bahwa semua studi dalam populasi yang dipilihnya independen (Lipsey & Wilson, 2003). Kedua, ES diubah atau dihitung dari data yang dikodekan yang sudah dinyatakan memenuhi kriteria. Langkah transformasi diperlukan untuk membakukan jenis ES yang mungkin berbeda dari studi yang berbeda. Kemudian analisis selanjutnya tergantung pada model estimasi yaitu model efek-tetap atau efek-acak yang akan dipilih (Borenstein et al., 2009).

3.6.3 Pemilihan Model Estimasi

Model efek tetap didasarkan pada asumsi bahwa populasi penelitian memiliki ES sejati yang sama tetapi tidak diketahui. Hal ini berarti, semua karakteristik penelitian yang mempengaruhi ES adalah sama di seluruh populasi penelitian. Jadi semua penelitian menggambarkan ES sejati yang identik. Variasi sampel penelitian hanya dapat dijelaskan oleh kesalahan pengambilan sampel.

Model efek-acak mengasumsikan bahwa penelitian dapat bervariasi mengenai ES yang sebenarnya. Biasanya studi berbeda satu sama lain karena dalam kebanyakan kasus, studi yang akan diintegrasikan dilakukan oleh peneliti yang berbeda, dengan peserta dengan karakteristik berbeda yang berbeda atau desain penelitian yang berbeda-beda. Dengan demikian, ES sejati yang mendasarinya diharapkan serupa, tetapi tidak identik. Jadi setiap studi memiliki ES sejati lain yang mendasarinya, ES yang diamati dari penelitian ini adalah sampel acak dari masing-masing studi ES sejati.

3.6.4 Perhitungan ES Keseluruhan

Apapun model yang digunakan, efek keseluruhan adalah rata-rata tertimbang dari populasi ES. Bobot diberikan sesuai dengan masing-masing studi presisi, sehingga studi yang lebih tepat (dapat diandalkan) menerima bobot yang lebih berat daripada yang kurang tepat. Selanjutnya distribusi set ES dijelaskan dan rata-rata tertimbang dan interval kepercayaan untuk rata-rata tertimbang diestimasi menggunakan model yang dipilih (Lipsey & Wilson, 2003).

Setelah itu distribusi ES diuji untuk heterogenitas. Pendekatan yang sering digunakan untuk mengidentifikasi heterogenitas pada meta-analisis adalah perhitungan probabilitas dari jumlah total varian studi-studi yang diamati hanya disebabkan oleh kesalahan pengambilan sampel (asumsi model efek-tetap). Oleh karena itu, dalam pendekatan ini probabilitas rendah menunjukkan heterogenitas. Jika model efek tetap digunakan, uji heterogenitas dapat

membuktikan secara positif pilihan ini, karena heterogenitas bertentangan dengan asumsi untuk model ini. Apapun model kesalahan yang dipilih, jika heterogenitas ditunjukkan, peneliti perlu melakukan analisis lebih lanjut untuk mengidentifikasi pengaruh moderator, seperti yang diusulkan secara apriori oleh teori-teori yang mendasarinya. Metode untuk analisis moderator adalah “analisis subkelompok” (Borenstein et al., 2009).

Model estimasi digunakan untuk memperkirakan efek keseluruhan. Kedua model memberikan bobot yang berbeda. Dalam model efek-tetap, dimana ES hanya bervariasi berdasarkan kesalahan pengambilan sampel, studi yang lebih besar diberi bobot lebih banyak karena ketepatan informasi tentang ES umum lebih baik daripada dalam studi kecil. Ini mencerminkan fakta, bahwa semakin besar ukuran sampel maka interval kepercayaan semakin sempit (Borenstein et al., 2009). Dalam model ini efek keseluruhan M dihitung sebagai berikut:

$$M = \frac{\sum_{i=1}^k W_i Y_i}{\sum_{i=1}^k W_i} \quad (1)$$

W_i melambangkan bobot studi, Y_i mewakili ES studi. Bobot studi merupakan kebalikan dari varians dalam studi V_{Y_i} :

$$W_i^{fixed} = \frac{1}{V_{Y_i}} \quad (2)$$

Perbedaan antara model efek-tetap dan efek-acak merupakan tambahan varians antar-studi acak tambahan dalam model efek-acak. Model efek-acak memperkirakan rata-rata distribusi efek, sehingga semua efek sebenarnya harus terkandung dalam efek keseluruhan. Oleh karena itu, tidak ada studi yang lebih kecil yang harus diabaikan, meskipun mereka kurang tepat,

juga studi besar dengan presisi yang lebih tinggi tidak akan menerima bobot yang lebih besar secara proporsional. Dalam model efek-acak varians disesuaikan dengan varians antar studi tambahan T^2 . T^2 adalah estimasi absolut varians antara studi. Dalam hal ini varian ES menyertakan varian studi dalam V_{Y_i} seperti dalam model efek tetap dan tambahan varian antar-studi T^2 .

Diperoleh hubungan sebagai berikut:

$$V_{Y_i}^{random} = V_{Y_i} + T^2 \quad (3)$$

$$W_i^{random} = \frac{1}{V_{Y_i}^{random}} \quad (4)$$

Karena T^2 adalah ukuran mutlak, maka T^2 distandarisasi oleh C menurut metode momen. Q yang diharapkan sama dengan varians antar-studi standar ditambah varians dalam studi:

$$Q = C.T^2 + df \leftrightarrow T^2 = \frac{Q-df}{C} \quad (5)$$

Perhitungan efek keseluruhan tetap secara formal sama kecuali untuk bobot yang diubah;

$$M^{random} = \frac{\sum_{i=1}^k W_i^{random} Y_i}{\sum_{i=1}^k W_i^{random}} \quad (6)$$

Pilihan model kesalahan tergantung pada harapan peneliti tentang set studinya dan tujuan meta-analisis. Model efek tetap kurang cocok untuk meta-analisis, karena semua studi yang dimasukkan harus identik secara fungsional (semua studi memiliki efek yang sama) dan tujuannya bukan untuk menggeneralisasi hasil bagi populasi lain (Cooper, 2010). Sebagian besar model efek acak akan sesuai karena studi yang dikumpulkan cenderung dilakukan oleh peneliti yang beroperasi secara independen dan biasanya studi berbeda dalam beberapa hal. Tujuan meta-analisis adalah untuk membuat generalisasi tertentu, sehingga model efek tetap dengan batasan yang disebutkan di atas akan dibatasi.

Model efek tetap hanya boleh diterapkan jika benar-benar sesuai. Dengan cara apa pun, jika dimulai dengan model efek-acak dan tidak ada heterogenitas yang terdeteksi, model tersebut akan direduksi secara otomatis menjadi model efek-tetap, yang merupakan kasus khusus dari model efek-acak dengan varian antar studi dari nol (Borenstein et al., 2009). Oleh karena itu Borenstein et al. (2009) dan Cumming (2012) merekomendasikan untuk "secara rutin menggunakan model efek acak.

3.6.5 Penyesuaian ES

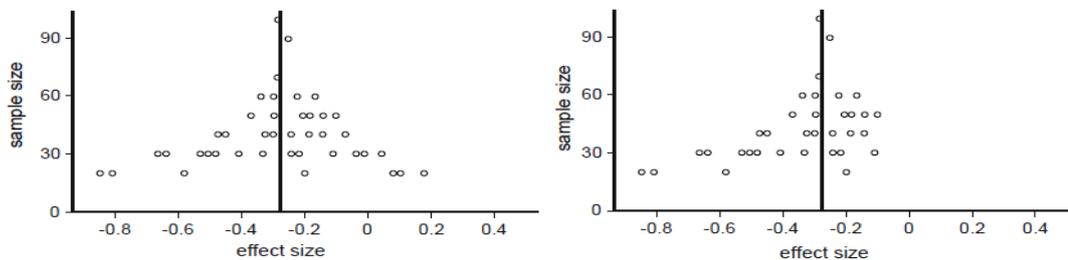
Saat menyesuaikan ES, setiap pencilan statistik harus diatasi dengan beberapa cara. Kesalahan yang mengakibatkan pencilan dapat berasal dari pengkodean, transfer data (oleh para pengkode atau peneliti utama) atau alasan yang tidak diketahui lainnya (Cooper, 2010; Wood & Eagly, 2009). Jika terlalu ekstrim, pencilan statistik dapat dikeluarkan dari meta-analisis, karena pencilan tidak mungkin menjadi bagian dari populasi yang diselidiki (Cooper, 2010).

Penyesuaian masing-masing ES, yaitu koreksi yang dijelaskan oleh Schmidt dan Hunter (2015), memungkinkan meta-analisis untuk mendekati besaran ES dalam kondisi ideal, yang meningkatkan keandalan. Namun informasi yang dibutuhkan tidak tersedia dalam kebanyakan kasus atau hanya untuk mayoritas, sehingga meta-analisis perlu memutuskan antara akurasi atau komparabilitas: Koreksi ES, dimana informasi tersedia, atau membiarkan semua ES tidak disesuaikan.

Kelemahan dalam studi meta-analisis adalah adanya bias publikasi. Jika beberapa dari studi yang dimasukkan dalam analisis merupakan sampel bias, maka ukuran efek keseluruhan yang dilaporkan akan mencerminkan bias ini (Borenstein et al., 2009). Akibatnya ada kekhawatiran bahwa hasil meta-analisis ini mungkin melebih-lebihkan ukuran efek sebenarnya. Karena itu analisis bias publikasi dilakukan untuk menafsirkan dan menarik kesimpulan tentang generalisasi dan keterbatasan temuan kumulatif (Cooper, 2010). Umumnya dalam perangkat lunak meta-analisis, analisis sensitivitas dilakukan dengan menghapus sebuah studi atau lebih kemudian mengamati perubahan yang terjadi (Zedeck, 2014).

Untuk menafsirkan dan menarik kesimpulan tentang generalisasi temuan kumulatif, kemungkinan bias publikasi harus diatasi. Salah satu metode umum apakah ada bias publikasi (Borenstein et al., 2009) adalah dengan menggunakan plot corong untuk secara grafis mengidentifikasi studi yang hilang karena "masalah *file laci*" ini. Masalah *file laci* merupakan fakta bahwa sebagian besar studi tidak dipublikasikan (hanya ada di laci arsip seseorang) karena berbagai alasan. Plot corong adalah plot pencar, dimana kesalahan standar individu atau ukuran sampel diplot terhadap ES individu studi dalam sistem koordinat kartesius. Varians penelitian semakin menurun dengan meningkatnya ukuran sampel, sehingga studi dengan ukuran sampel kecil dan ES mendekati nol lebih sering secara statistik tidak signifikan. Penelitian harus menyebar secara simetris ke kiri dan kanan garis vertikal menandai ringkasan ES (bentuk corong), jika tidak ada temuan yang hilang. Oleh karena itu plot

corong asimetris menandakan bias publikasi (Cumming, 2012). Gambar 3.4 berikut menunjukkan contoh plot corong simetris dan asimetris:



Gambar 3.5: Plot Corong Simetri (kiri) and Asimetri (kanan)
(Sumber Gambar dari Bortz & Döring, 2006)

Satu pendekatan lebih lanjut, apakah bias publikasi sepenuhnya bertanggung jawab untuk efek yang diamati, adalah uji Trim dan Fill digunakan. Tidak ada bias jika tidak ada perbedaan antara ukuran efek yang diamati dan efek virtual yang dibuat menurut model efek-acak.

Jadi dalam penelitian untuk menentukan bias publikasi, plot corong dan uji Trim dan Fill diselidiki. Tidak ada bias jika ukuran efek dari penelitian menunjukkan distribusi simetris di sekitar ukuran efek keseluruhan dalam plot corong (Borenstein et al., 2009).

3.6.6 Analisis Heterogenitas

Selanjutnya, himpunan ES diperiksa untuk menganalisis heterogenitas. Heterogenitas adalah sejauh mana ukuran efek berbeda satu sama lain (Petricrew & Roberts, 2006). Dalam meta-analisis, perlu untuk menggunakan uji statistik untuk menilai apakah inkonsistensi dalam ukuran efek yang diamati lebih besar

daripada yang diharapkan secara kebetulan. Jika demikian, maka efek yang diamati dikatakan heterogen.

Jika hasilnya menunjukkan heterogenitas, tidak semua titik data adalah bagian dari populasi yang menggambarkan ES sejati yang sama dan bervariasi lebih dari tingkat kesalahan pengambilan sampel pada tingkat dalam studi. Dengan demikian ada variasi tambahan yang ada di tingkat antar-studi. Pendekatan umum adalah membandingkan jumlah variasi keseluruhan yang diamati dan variasi perkiraan karena kesalahan pengambilan sampel dari efek yang diamati. Perbedaan variasi-variasi ini adalah variabilitas antar-studi ES yang sebenarnya.

Sebaliknya, homogenitas adalah ketika variabilitas dalam ukuran efek yang diamati tidak lebih besar dari yang diharapkan jika diberikan kesempatan atau kesalahan pengambilan sampel. Untuk menentukan apakah heterogenitas statistik lebih besar daripada kebetulan, dilakukan dengan uji chi-kuadrat hipotesis homogenitas efek menggunakan statistik Cochran's Q yang selanjutnya disebut dengan statistik untuk menilai apakah efeknya sama (Kulinskaya et al., 2008).

Statistik Cochran's Q menguji perbedaan efek antara dua atau lebih perlakuan yang diterapkan pada set komponen eksperimental yang sama (Borenstein, 2009). Jika hipotesis nol gagal ditolak, maka estimasi nilai Q akan memiliki kira-kira distribusi chi-kuadrat χ^2 dengan derajat kebebasan sama dengan jumlah studi dikurangi satu, $k-1$. Jika statistik Q signifikan, kita dapat menganggap bahwa heterogenitas ada. Jika Q ditemukan secara statistik tidak

signifikan, aman untuk memperkirakan bahwa ukuran efek homogen, menganggap analisis moderator tidak perlu (Brendel, 2011).

Meskipun tes signifikansi statistik Q yang sering digunakan di meta-analisis, mereka tidak boleh digunakan sebagai sumber informasi tunggal untuk penilaian heterogenitas, karena mereka rentan terhadap kesalahan. Untuk itu I^2 berguna untuk mengukur tingkat inkonsistensi (varians relatif) untuk menghindari salah tafsir (Borenstein et al., 2009). Statistik I^2 , yang merupakan pelengkap dari statistik Q, berguna karena menentukan tingkat heterogenitas (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). Statistik I^2 mewakili rasio persentase heterogenitas variabel penelitian dalam kaitannya dengan total variabilitas dalam ukuran efek (Carter, 2012). Dikatakan I^2 besar, bahwa varian yang diamati sebenarnya ada dan perlu dijelaskan, mis. oleh moderator-analisis. Higgins et al. (2003) juga menetapkan skala untuk inkonsistensi: rendah ($I^2 = 25\%$), sedang (50%) dan tinggi (75%). Kelebihan dari ukuran ini adalah, bahwa hal itu tidak secara langsung dipengaruhi oleh jumlah studi, juga bukan merupakan perkiraan dari efek sebenarnya yang mendasarinya, sehingga ini adalah statistik deskriptif dan hanya heterogenitas dari langkah-langkah yang diamati yang ditangani.

Dalam penelitian saat ini, uji heterogenitas diimplementasikan menggunakan perangkat lunak *Comprehensive Meta-Analysis* (CMA) [Versi 3] melalui metode statistik Q. CMA digunakan untuk menghitung ukuran efek g Hedges serta menghitung informasi statistik seperti nilai- p , nilai- t , statistik Q, I^2 , dan interval kepercayaan. Plot hutan, plot corong, dan analisis bias publikasi

juga dibuat menggunakan perangkat lunak ini. Sebagai hasil dari perhitungan, ketika ukuran efek secara statistik heterogen ($Q_b > \chi^2_{,95}$; $p < 0,05$), hipotesis pada homogenitas ukuran efek ditolak (Borenstein et al., 2009).

3.6.7 Analisis Moderator

Analisis moderator dilakukan setelah diperoleh hasil pengujian bahwa hipotesis pada homogenitas ukuran efek ditolak. Analisis moderator dalam kebanyakan kasus, analisis heterogenitas akan menemukan heterogenitas ES, yang merupakan varian yang tidak dapat dijelaskan dan menyiratkan analisis lebih lanjut dari calon moderator. Moderator adalah variabel yang mempengaruhi atau kovary dengan efek yang dipelajari dan mampu menjelaskan variabilitas sistematis antara varians studi (Wood & Eagly, 2009). Analisis moderator mencari moderator dan memperkirakan sejauh mana mereka dapat menjelaskan variabilitas dalam ES dibandingkan studi (Cumming, 2012). Oleh karena itu, analisis moderator dapat digunakan untuk menjelaskan heterogenitas sebelumnya yang tidak dapat dijelaskan, jika moderator telah dikodekan dalam langkah ekstraksi data (Cumming, 2012). Analisis moderator hanya dapat menunjukkan korelasi antara moderator dan perbedaan ES, tetapi tidak ada hubungan sebab akibat. Namun, kesimpulan mungkin menawarkan dukungan untuk hipotesis mengenai efektivitas intervensi yang dapat diteliti lebih lanjut dalam studi masa depan (Littell et al., 2008).

3.6.8 Analisis Subgroup

Metode ini memisahkan himpunan studi (ES) menjadi subkelompok sesuai dengan moderator kategoris atau karakteristik studi yang diidentifikasi, menghitung efek rata-rata dan varians untuk setiap subkelompok, kemudian efek rata-rata dari subkelompok dibandingkan (misalnya dengan uji-Z, uji-Q berdasarkan analisis dari varians atau heterogenitas) dan sejauh mana perbedaan dijelaskan dengan memberikan interval kepercayaan 95%.

Sebagian besar analisis statistik meta-analisis sudah diimplementasikan dalam perangkat lunak (Cumming, 2012; Schmidt & Hunter, 2015), sehingga bagian penting dari pekerjaan manusia adalah mengumpulkan data yang diperlukan untuk mengeksekusi metode meta-analisis ini dan menerapkan alat meta-analisis dengan benar. Dalam penelitian ini alat bantu analisis adalah perangkat lunak CMA (*comprehensive meta-analysis*). CMA menghitung nilai Z untuk menentukan uji signifikansi dan memberikan ukuran efek rata-rata dengan interval kepercayaan untuk setiap kelas variabel serta homogenitas antar kelompok yaitu nilai Q_b .

Sebagai hasil perhitungan, jika $Z_{hitung} > Z_{tabel}$ dengan $p < 0,05$ maka hipotesis nol ditolak (Borenstein & Hedges, 2009). Menolak hipotesis nol artinya menolak pernyataan bahwa perbedaan risiko sebenarnya adalah 0,0. Dengan perkataan lain penerapan CAME menghasilkan ukuran efek yang lebih besar pada kemampuan matematis siswa daripada pendekatan konvensional.

Ketika ukuran efek secara statistik heterogen ($Q_b > \chi^2_{2,95}$; $p < 0,05$), hipotesis pada homogenitas ukuran efek ditolak (Demir & Başol, 2014). Model

efek acak digunakan ketika Q_b ditolak. Menolak Q_b menyiratkan bahwa ukuran efek dari kelompok karakteristik penelitian mungkin tidak mengukur parameter populasi yang sama (Borenstein & Hedges, 2009). Dengan kata lain, ada perbedaan yang signifikan secara statistik dalam ukuran efek rata-rata untuk setiap kelompok karakteristik penelitian (Bayir & Bozkurt, 2018).