

Efektivitas Penggunaan Duolingo dalam Pembelajaran Bahasa Arab: Overall Effect Size dan Heterogenitas

Julham Hukom

Universitas Negeri Makassar, Makassar, Indonesia

julham.hukom@unm.ac.id

Irfan

Universitas Negeri Makassar, Makassar, Indonesia

irfan90@unm.ac.id

Akmal

Universitas Islam Ahmad Dahlan, Sinjai, Indonesia

latoqalu@gmail.com

Abstrak

Berbagai penelitian telah mengkaji keefektifan penggunaan *Duolingo* pada pembelajaran bahasa arab jika dibandingkan dengan media tradisional. Meskipun demikian, hasil penelitian masih menunjukkan inkonsistensi. Penelitian ini bertujuan untuk menganalisis *overall effect size* dan heterogenitas terkait penggunaan *Duolingo* dalam pembelajaran bahasa arab dibandingkan dengan media tradisional. Metode penelitian ini menggunakan desain meta-analisis dengan pendekatan *group contrast*, dengan melibatkan delapan sampel penelitian utama ($k = 8$) yang memenuhi kriteria inklusi. Hasil analisis menggunakan *software OpenMee* menunjukkan bahwa penggunaan *Duolingo* secara signifikan lebih efektif dalam meningkatkan kemampuan bahasa Arab dibandingkan media tradisional, dengan nilai *overall effect size* sebesar ($g = 0,91$; $p < 0,05$). Hasil analisis heterogenitas menunjukkan bahwa variabel tipe kemampuan yang diukur mempengaruhi efektivitas penggunaan *duolingo* dalam pembelajaran bahasa arab ($Qb = 9.56$; $p < 0.05$), penggunaan *duolingo* lebih efektif untuk meningkatkan kemampuan kosakata dibandingkan kemampuan mendengar dan lainnya. Namun pada variabel jenjang pendidikan dan kapasitas kelas tidak mempengaruhi dampak penggunaan *Duolingo* dalam pembelajaran bahasa arab. Penelitian ini memiliki keterbatasan, khususnya pada jumlah studi yang dianalisis dan ketidakterlibatan analisis variabel moderator. Penelitian selanjutnya dapat melibatkan sampel yang lebih banyak dan mengeksplorasi faktor-faktor lain seperti jenis durasi penggunaan aplikasi atau

integrasi *Duolingo* dengan strategi pembelajaran lain untuk memperoleh pemahaman secara komprehensif.

Kata kunci: Bahasa arab, *Duolingo*, Meta-Analysis

Abstract

Various studies have examined the effectiveness of using *Duolingo* in Arabic language learning compared to traditional media. However, the results of the study still show inconsistencies. This study aims to analyze the overall effect size and heterogeneity related to the use of *Duolingo* in Arabic language learning compared to traditional media. This research method uses a meta-analysis design with a group contrast approach, involving eight main research samples ($k = 8$) that meet the inclusion criteria. The results of the analysis using OpenMee software show that the use of *Duolingo* is significantly more effective in improving Arabic language skills than traditional media, with an overall effect size value of ($g = 0.91$; $p < 0.05$). The results of the heterogeneity analysis show that the type of ability variable measured affects the effectiveness of using *Duolingo* in Arabic language learning ($Q_b = 9.56$; $p < 0.05$), the use of *Duolingo* is more effective in improving vocabulary skills than listening skills and others. However, the variables of education level and class capacity do not affect the impact of using *Duolingo* in Arabic language learning. This study has limitations, especially in the number of studies analyzed and the lack of involvement in the analysis of moderator variables. Future research could involve a larger sample and explore other factors such as the type of app usage duration or *Duolingo* integration with other learning strategies to gain a more comprehensive understanding.

Keywords: Arabic Language, *Duolingo*, Meta-Analysis

Pendahuluan

Perkembangan teknologi digital dalam beberapa dekade terakhir membawa perubahan signifikan dalam dunia pendidikan, termasuk dalam bidang pembelajaran bahasa asing. Salah satu inovasi berbasis teknologi yang semakin banyak digunakan adalah platform pembelajaran berbasis aplikasi seperti *Duolingo*. Dengan pendekatan gamifikasi, personalisasi pembelajaran, dan integrasi kecerdasan buatan, *Duolingo* menjadi salah satu aplikasi yang populer di kalangan pembelajar bahasa di seluruh dunia, termasuk dalam konteks pembelajaran bahasa Arab. Tren ini mencerminkan pergeseran paradigma dari metode pembelajaran tradisional menuju model pembelajaran yang lebih mandiri, adaptif, dan berbasis teknologi. Kehadiran *Duolingo* sebagai media pembelajaran mencerminkan upaya dunia

pendidikan dalam merespons kebutuhan peserta didik abad ke-21 untuk menguasai bahasa secara aktif dan relevan dalam lingkungan digital yang terus berkembang.

Duolingo memiliki potensi besar dalam meningkatkan kemampuan berbahasa arab, karena menggabungkan prinsip personalisasi pembelajaran yang efektif dengan teknologi interaktif. Beberapa fitur seperti latihan berbasis konteks, umpan balik instan, sistem pengulangan bertingkat (*spaced repetition*), serta penggunaan gamifikasi untuk mempertahankan motivasi belajar, terbukti efektif dalam mendukung pengembangan keterampilan bahasa (Ajisoko, 2020; Fakhrurriana et al., 2024; Wagner, 2020). Aplikasi ini tidak hanya meningkatkan kosakata siswa (Irawan et al., 2020; Jaelani & Sutari, 2021) dan mendengar (Jiang et al., 2021; Octaviani & others, 2022), tetapi jika diterapkan dengan metode pembelajaran yang tepat juga dapat meningkatkan kemampuan berbicara dan menulis (Kazu & Kuvvetli, 2025). Dengan fleksibilitas waktu dan tempat belajar, serta adaptasi tingkat kesulitan sesuai dengan progres pengguna, *Duolingo* juga mampu mengakomodasi keragaman gaya belajar peserta didik, yang menjadi salah satu faktor penting dalam meningkatkan efektivitas pembelajaran.

Berbagai penelitian telah dilakukan untuk mengukur efektivitas penggunaan *Duolingo* dalam pembelajaran bahasa arab dibandingkan dengan media tradisional. Meskipun demikian, berbagai penelitian tersebut menghasilkan temuan yang inkonsisten. Misalnya, studi yang dilakukan oleh (Adira et al., 2024; Al-Gifari et al., 2025; Idris & Kasih, 2023; Muttaqin et al., 2025; Renita, 2023; Yazid et al., 2023; Zakiyah et al., 2024), melaporkan bahwa penggunaan *Duolingo* berpengaruh signifikan pada pembelajaran bahasa arab jika dibandingkan dengan media pembelajaran tradisional. Namun, temuan mereka inkonsisten dengan temuan Baroroh et al. (2022). Temuan mereka menunjukkan bahwa meskipun ada peningkatan prestasi belajar bahasa arab menggunakan *Duolingo*, namun pengaruhnya tidak berbeda signifikan dengan penggunaan media tradisional. Inkonsistensi hasil ini dapat menyebabkan keraguan di kalangan pendidik dan

pembuat kebijakan dalam mengintegrasikan aplikasi *Duolingo* secara luas karena tidak berbasis bukti empiris yang kuat.

Beberapa faktor yang dapat mempengaruhi inkonsistensi hasil ini diantaranya adalah karakteristik metode penelitian yang digunakan, maupun karakteristik dari penggunaan teknologi *Duolingo* itu sendiri. Misalnya penelitian hukom (2025), Purnomo et al. (2022), dan Sulistyowati et al. (2023) melaporkan bahwa ukuran sampel, jenjang pendidikan, dan tipe teknologi adalah beberapa faktor yang mempengaruhi efektivitas penggunaan pembelajaran berbasis teknologi terhadap pencapaian akademik siswa.

Berdasarkan kasus ini, perlu dilakukan penelitian lebih lanjut untuk mengevaluasi secara komprehensif faktor-faktor yang menyebabkan variasi hasil penelitian mengenai dampak *Duolingo* dalam pembelajaran bahasa arab. Salah satu pendekatan yang tepat untuk kasus ini adalah metode meta-analisis (Cahyani et al., 2024; Hukom et al., 2023; Kamsurya et al., 2022; Mawardi et al., 2024; Muhtadi et al., 2022; Purnomo et al., 2022; Samritin et al., 2023; Setiawan et al., 2022; Sulistyowati et al., 2023; Ulum & Hukom, 2025; Zuliana et al., 2025). Meta-analisis bertujuan untuk mensintesis atau menggabungkan berbagai penelitian dengan topik yang sama (Borenstein et al., 2021). Meta-analisis berfokus pada penghitungan effect size dengan menganalisis informasi statistik dari berbagai penelitian yang mengkaji topik yang sama (Lipsey & Wilson, 2001). Meta-analisis menggabungkan sampel dari berbagai penelitian sehingga standar error yang dihasilkan lebih kecil (Borenstein et al., 2009).

Meta-analisis terdahulu yang membandingkan pembelajaran bahasa asing menggunakan teknologi mobile dan media tradisional telah dilakukan oleh Burston, and Giannakou (2022), Chen et al. (2020), Li (2023), Santaliestra et al. (2024), Xu et al. (2022), dan Yu and Trainin (2022). Akan tetapi, meta-analisis yang dilakukan tersebut tidak berfokus untuk mengevaluasi dampak tipe teknologi mobile dengan *Duolingo* secara spesifik dalam pembelajaran bahasa asing. Sementara temuan

Sulistyowati et al. (2023) menemukan bahwa tipe teknologi yang digunakan dapat mempengaruhi efektivitas pembelajaran terhadap prestasi belajar siswa. Selain itu, faktor-faktor lain seperti jenjang pendidikan, kapasitas kelas, dan tipe kemampuan juga perlu diidentifikasi untuk melihat apakah keefektifan penggunaan *Duolingo* dalam pembelajaran bahasa arab dibandingkan media tradisional berbeda berdasarkan ketiga faktor ini.

Berangkat dari keterbatasan tersebut, penelitian ini bertujuan untuk mengisi kesenjangan dalam penelitian yang ada dengan melakukan analisis yang lebih komprehensif. Secara spesifik, ada dua tujuan utama yang ingin dicapai: (1) mengukur efektivitas penggunaan *Duolingo* dalam pembelajaran bahasa arab jika dibandingkan dengan media tradisional; dan (2) mengidentifikasi faktor-faktor yang dapat memengaruhi efektivitas penggunaan *Duolingo* dalam pembelajaran bahasa arab. Untuk menjawab tujuan tersebut, penelitian ini merumuskan beberapa pertanyaan penelitian (RQ) sebagai berikut.

- RQ 1: Bagaimana efektivitas penggunaan *Duolingo* pada pembelajaran bahasa arab dibandingkan media tradisional?
- RQ 2: Apakah variabel jenjang pendidikan mempengaruhi efektivitas penggunaan *Duolingo* pada pembelajaran bahasa arab?
- RQ 3: Apakah variabel kapasitas kelas mempengaruhi efektivitas penggunaan *Duolingo* pada pembelajaran bahasa arab?
- RQ 4: Apakah variabel tipe kemampuan mempengaruhi efektivitas penggunaan *Duolingo* pada pembelajaran bahasa arab?

Kajian Teori

Duolingo

Duolingo menawarkan sekitar 95 bahasa yang dapat dipelajari, dan pengguna juga dapat belajar menggunakan bahasa selain bahasa Inggris (*Duolingo*, 2021). Pengguna memulai dengan memilih bahasa target yang ingin dipelajari dan

dapat mengikuti tes penempatan jika sudah memiliki pengetahuan dasar sebelumnya. Mereka menetapkan target harian berupa jumlah poin pengalaman (*experience points*) dan akan mendapatkan bonus ketika mencapai target tersebut. Menyelesaikan satu pelajaran per hari akan menambah satu hari pada “*Streak*”, namun jika pada suatu hari tidak ada pelajaran yang diselesaikan, *Streak* tersebut akan direset ke nol (sebagai bentuk penguatan). Aplikasi ini juga sesekali menawarkan tantangan opsional kepada pengguna, seperti mempertahankan *Streak* selama beberapa hari berturut-turut, membandingkan poin pengalaman dengan pengguna lain di berbagai liga, atau memberikan hadiah setelah tantangan berhasil diselesaikan (berorientasi pada kesenangan dan kompetisi) (Nah et al., 2013).

Sistem pelajaran dalam *Duolingo* dibangun berdasarkan topik-topik tertentu seperti keluarga, makanan, dan perjalanan. Setiap topik memperkenalkan konsep tata bahasa dan budaya dengan penjelasan yang sangat terbatas, sementara fokus utama pelajaran adalah memperkenalkan kosakata baru dan latihan-latihan berbentuk drill. Jenis latihan yang ditawarkan meliputi penerjemahan, soal pilihan ganda untuk pengenalan kata, dan ejaan. Jawaban yang salah ditangani dengan dua cara. Pada beberapa akun, terdapat sistem “hati” (*heart*), di mana setiap kesalahan akan mengurangi satu dari lima hati yang tersedia. Jika seluruh hati habis, pengguna tidak dapat melanjutkan latihan hingga mereka memulihkan sebagian hati. Di beberapa perangkat, kesalahan akan memicu pengulangan materi lebih sering serta mengurangi jumlah poin pengalaman atau pencapaian yang diperoleh setelah menyelesaikan pelajaran. Tampaknya fitur “hati” hanya diterapkan pada beberapa akun tertentu dan tidak tersedia di versi web aplikasi. Biasanya, kesalahan disertai dengan komentar singkat, sedangkan jawaban yang benar akan mendapatkan bonus tambahan di akhir pelajaran serta umpan balik positif dalam bentuk singkat. Dalam kedua situasi tersebut, pengguna juga memiliki opsi untuk mengakses forum

diskusi yang tersedia untuk setiap pertanyaan tertentu dan berinteraksi dengan pelajar lain.

Memfasilitasi Kemandirian Belajar dengan Duolingo

Duolingo telah menjadi salah satu platform pembelajaran bahasa yang mendukung pembelajaran mandiri secara luas di berbagai belahan dunia. Pembelajaran mandiri (*self-directed learning*) menuntut siswa untuk mengambil inisiatif, mengatur ritme belajar, serta mengevaluasi kemajuan belajar mereka secara otonom (Meece, 2023). Dalam konteks ini, *Duolingo* menyediakan fitur-fitur yang memperkuat karakteristik pembelajaran mandiri, seperti pelacakan kemajuan otomatis, pengaturan target harian, pengulangan adaptif berdasarkan performa pengguna, dan motivasi berbasis penghargaan gamifikasi (tingkatan, hadiah, dan tantangan harian).

Penelitian oleh Payne (2020) mengungkapkan bahwa penggunaan *Duolingo* secara mandiri mampu meningkatkan akuisisi kosa kata dan struktur tata bahasa dasar secara signifikan, meskipun hasil yang optimal tetap memerlukan konsistensi dan keterlibatan pengguna dalam jangka waktu tertentu. Selain itu, Li dan Bonk (2025) menyoroti bahwa *Duolingo*, dengan pendekatan mikro-pembelajarannya (*microlearning*), mendorong pembelajar untuk mengembangkan disiplin diri, karena latihan singkat tetapi berulang setiap hari terbukti lebih efektif dalam mempertahankan pengetahuan jangka panjang dibandingkan sesi belajar panjang tetapi jarang.

Meta-Analysis

Meta-analisis adalah metode kuantitatif untuk mensintesis hasil-hasil penelitian sebelumnya, dengan tujuan memperoleh estimasi efek yang lebih akurat dan dapat digeneralisasi. Dalam bidang pendidikan, khususnya dalam evaluasi efektivitas model pembelajaran, meta-analisis berfungsi untuk mengatasi keterbatasan studi individual yang mungkin memiliki sampel kecil atau hasil yang inkonsisten. Menurut Borenstein et al. (Borenstein et al., 2009), meta-analisis

menggabungkan ukuran efek dari berbagai studi menggunakan metode statistik tertentu, seperti *fixed-effects model* atau *random-effects model*, tergantung pada tingkat heterogenitas antar studi.

Dengan mengkaji ukuran efek gabungan, meta-analisis memberikan bukti yang lebih kuat mengenai efektivitas suatu intervensi, termasuk efektivitas penggunaan *Duolingo* dalam pembelajaran bahasa Arab. Salah satu keunggulan utamanya adalah kemampuannya dalam mengurangi bias antar studi individual, seperti perbedaan ukuran sampel, metode penelitian, atau kualitas pelaporan hasil. Dengan menggabungkan ukuran efek melalui model statistik seperti *fixed-effects* atau *random-effects*, meta-analisis dapat memperhitungkan heterogenitas antar studi dan menghasilkan ukuran efek gabungan yang lebih stabil (Borenstein et al., 2009).

Penggunaan meta-analisis dalam studi pembelajaran berbasis teknologi menjadi penting untuk menyaring bukti ilmiah yang tersebar dan memberikan dasar pengambilan keputusan yang berbasis data. Meta-analisis berfokus pada penghitungan effect size dengan menganalisis informasi statistik dari berbagai penelitian yang mengkaji topik yang sama (Lipsey & Wilson, 2001). Meta-analisis menggabungkan sampel dari berbagai penelitian sehingga standar error yang dihasilkan lebih kecil (Borenstein et al., 2009).

Metode

Desain Penelitian

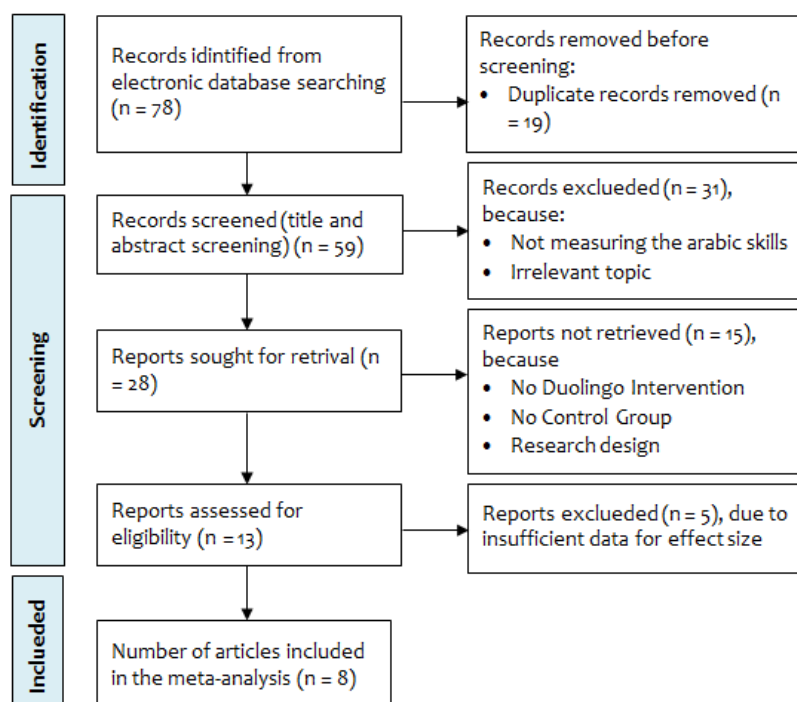
Penelitian ini menggunakan pendekatan meta-analisis untuk menilai seberapa efektif penggunaan teknologi *Duolingo* dalam pembelajaran bahasa Arab siswa jika dibandingkan dengan media tradisional. Selain itu, penelitian ini juga bertujuan untuk melakukan analisis terhadap variabel moderator untuk mengidentifikasi faktor-faktor yang dapat mempengaruhi keefektifan penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab. Meta-analisis dipilih karena

metode ini memungkinkan peneliti menggabungkan hasil dari berbagai penelitian sebelumnya, sehingga menghasilkan gambaran yang lebih lengkap dan akurat tentang pengaruh *Duolingo* terhadap hasil belajar. Dengan mengumpulkan data dari banyak studi, pendekatan ini juga membantu mengurangi kemungkinan bias dari satu studi tertentu.

Kriteria Inklusi dan Pengumpulan Data

Kriteria inklusi dalam penelitian meta-analisis ini meliputi: (1) artikel penelitian yang diterbitkan dalam rentang waktu 5 tahun terakhir; (2) menggunakan desain penelitian eksperimental atau kuasi-eksperimental yang melibatkan kelompok eksperimen (*duolingo*) dan kelompok kontrol secara jelas; (3) penelitian mengevaluasi efektivitas *duolingo* dalam pembelajaran bahasa arab di berbagai jenjang pendidikan; dan (4) studi harus melaporkan informasi statistik yang cukup untuk perhitungan effect size, seperti nilai rata-rata (mean), standar deviasi (SD), dan jumlah sampel dari masing-masing kelompok.

Pengumpulan data dilakukan secara sistematis melalui penelusuran di database Google Scholar dan mesin pencari Google dengan menggunakan kata kunci gabungan, seperti ("duolingo") AND ("arabic" OR "arab") AND ("experimental" OR "quasi-experimental"). Artikel yang diperoleh kemudian diseleksi menggunakan pedoman *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) yang meliputi empat tahap, yaitu: identifikasi (*identification*), penyaringan (*screening*), dan inklusi (*inclusion*). Dari proses seleksi yang ketat, diperoleh delapan studi primer yang memenuhi kriteria untuk dianalisis lebih lanjut. Gambar 1 menyajikan proses pengumpulan data menggunakan protokol PRISMA.



Gambar 1. Proses Screening Literatur Menggunakan PRISMA

Ekstraksi Data

Informasi statistik dari kelompok eksperimen (pembelajaran terbalik) dan kontrol seperti nilai rata-rata, simpangan baku, dan ukuran sampel dikumpulkan untuk memudahkan analisis data. Tabel 1 menyajikan hasil ringkasan ekstraksi data.

Tabel 1. Hasil Ekstraksi Data Numerik

| Studi | Duolingo | | | Not Duolingo | | | t-hitung | Sig. |
|--------------------------|----------|-------|----|--------------|-------|----|----------|-------|
| | Mean | SD | N | Mean | SD | N | | |
| (Al-Gifari et al., 2025) | 94.1 | 5.22 | 20 | 73.1 | 4.91 | 20 | | |
| (Muttaqin et al., 2025) | 8.16 | 0.94 | 25 | 7.22 | 0.99 | 23 | | |
| (Baroroh et al., 2022) | 73.88 | 9.24 | 18 | 70.24 | 8.33 | 25 | | |
| (Zakiah et al., 2024) | 74.84 | 26.77 | 31 | 54.67 | 26.41 | 30 | | |
| (Idris & Kasih, 2023) | | | 20 | | | 20 | 2.88 | |
| (Yazid et al., 2023) | | | 26 | | | 26 | 6.97 | |
| (Adira et al., 2024) | | | 21 | | | 21 | | 0.003 |
| (Renita, 2023) | | | 31 | | | 30 | | 0.001 |

Analisis Data

Dalam penelitian ini, perangkat lunak *OpenMEE* digunakan sebagai alat utama untuk mengolah dan menganalisis data dari berbagai studi yang telah dikumpulkan. Analisis data dilakukan melalui beberapa tahapan sistematis, yaitu. (1) menghitung *effect size* dari masing-masing studi, (2) dilakukan uji heterogenitas untuk melihat seberapa besar variasi antar studi. Hasil uji heterogenitas ini menjadi acuan dalam memilih model analisis yang tepat, apakah menggunakan model *fixed-effect* atau *random-effect* (3) hitung *overall effect size* yang menggambarkan dampak rata-rata penggunaan *Duolingo* terhadap prestasi akademik siswa, dan (4) analisis variabel moderator. Adapun Interpretasi terhadap nilai *effect size* mengacu pada pedoman yang dikembangkan oleh Cohen (1988), sebagaimana ditampilkan pada Tabel 2 berikut.

Table 2. Klasifikasi Effect Size

| Effect Size (<i>d</i>) | Kategori |
|--------------------------|------------|
| $0.00 \leq d < 0.20$ | Ignored |
| $0.20 \leq d < 0.50$ | Small |
| $0.50 \leq d < 0.80$ | Moderate |
| $0.80 \leq d < 1.30$ | Large |
| $1.30 \leq d$ | Very Large |

Hasil

Karakteristik Umum Penelitian Primer

Berdasarkan kriteria inklusi yang ditetapkan, kami memperoleh 8 sampel independen yang bersumber dari penelitian primer. Tabel 3 menyajikan ringkasan penelitian yang menjadi objek meta-analisis yang terdiri dari nama penelitian dan tahun publikasi, jenjang pendidikan, kapasitas kelas eksperimen, dan jenis kemampuan yang diukur. Distribusi jenjang pendidikan menunjukkan bahwa tujuh studi ($n = 7$) dilakukan pada jenjang SMP/MTs dan hanya satu studi ($n = 1$) dilakukan pada jenjang SMA/MA. Ditinjau dari kapasitas kelas, menunjukkan bahwa enam studi ($n = 6$) dilakukan pada kapasitas kelas kecil (≤ 30), sedangkan dua studi ($n = 2$) dilakukan pada kapasitas kelas besar (> 30). Sementara jika ditinjau dari tipe

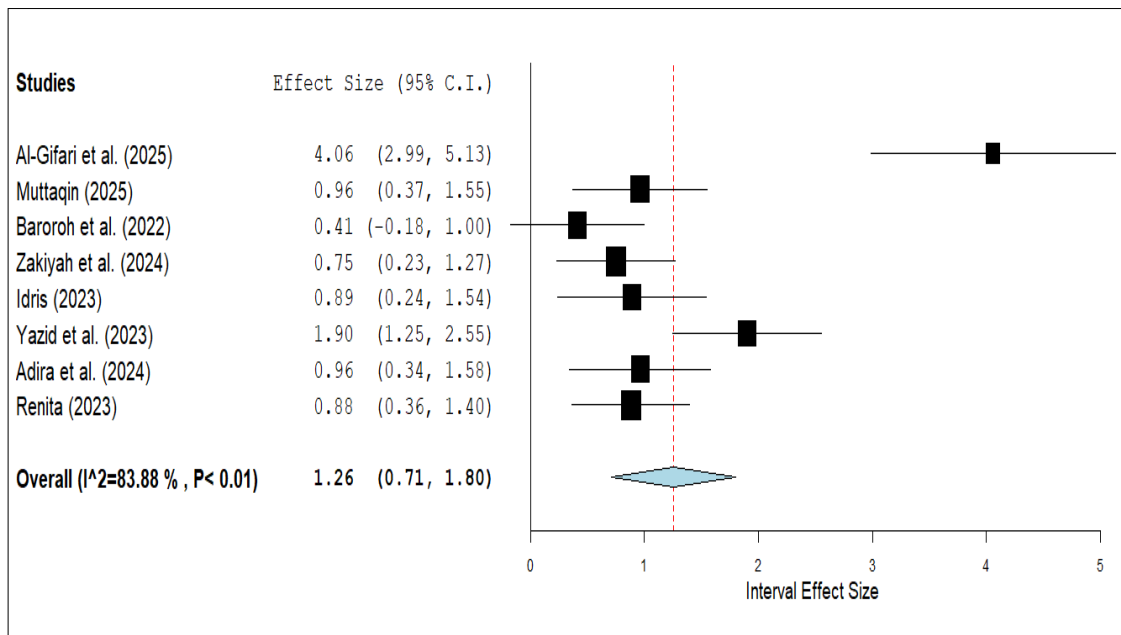
kemampuan yang diukur, menunjukkan bahwa enam studi ($n = 6$) mengukur kemampuan kosakata, satu studi ($n = 1$) mengukur kemampuan mendengar, dan satu studi ($n = 1$) mengukur minat belajar.

Tabel 3. Ringkasan penelitian yang menjadi objek meta-analisis

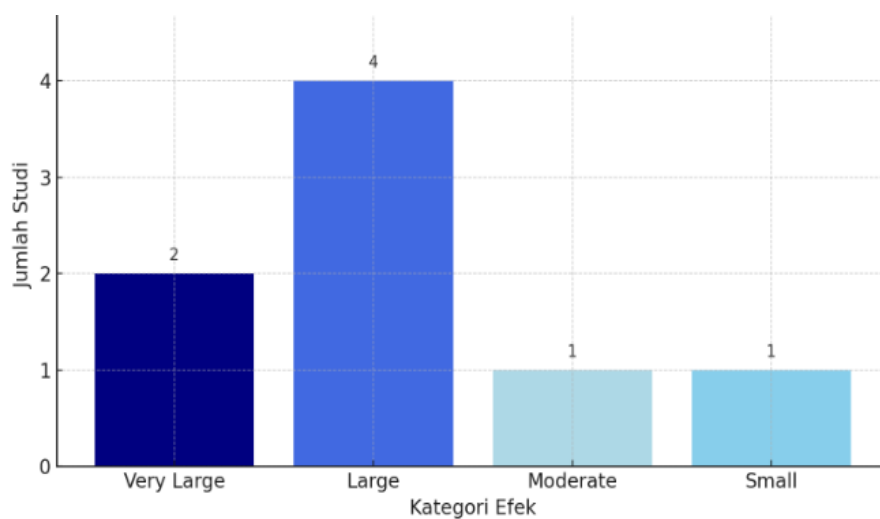
| Studi | Jenjang Pendidikan | Kapasitas Kelas Eksperimen | Tipe Kemampuan |
|--------------------------|--------------------|----------------------------|----------------|
| (Al-Gifari et al., 2025) | SMP/MTs | Kecil | Kosakata |
| (Muttaqin et al., 2025) | SMP/MTs | Kecil | Kosakata |
| (Baroroh et al., 2022) | SMP/MTs | Kecil | Mendengar |
| (Zakiah et al., 2024) | SMP/MTs | Besar | Kosakata |
| (Idris & Kasih, 2023) | SMP/MTs | Kecil | Kosakata |
| (Yazid et al., 2023) | SMP/MTs | Kecil | Kosakata |
| (Adira et al., 2024) | SMA/MA | Kecil | Minat Belajar |
| (Renita, 2023) | SMP/MTs | Besar | Kosakata |

Analisis Overall Effect Size

Tujuan pertama meta-analisi ini adalah untuk menghitung effect size gabungan. Berdasarkan hasil analisis meta yang divisualisasikan dalam forest plot (Gambar 2), sebanyak delapan dianalisis untuk mengevaluasi efektivitas penggunaan teknologi *duolingo* terhadap prestasi belajar bahasa arab siswa. Rentang nilai *effect size* berkisar antara 0.41 hingga 4.06, yang menunjukkan adanya variasi tingkat efektivitas antar studi. Berdasarkan klasifikasi Cohen (1988), dari delapan studi yang dianalisis, dua studi ($n = 2$) menunjukkan efek *very large*, yaitu studi Al-Gifari et al. (2025) dan Yazid et al. (2023), empat studi ($n = 4$) menunjukkan efek *large*, yaitu studi Adira et al. (2024), Idris (2023), Muttaqin (2025), dan Renita (2023), satu studi ($n = 1$) menunjukkan efek *moderate*, yaitu studi Zakiah et al. (2024), dan satu studi menunjukkan efek *small*, yaitu studi (2022)Baroroh et al. (2022). Gambar 3 menyajikan ringkasan rangkuman *effect size* dari delapan sebaran *effect size* pada *forest plot*.



Gambar 2. Forest Plot



Gambar 3. Distribusi kategori *effect size* terkait duolingo

Hasil analisis heterogenitas (Lihat Tabel 4) menunjukkan bahwa terdapat sebaran effect size bersifat heterogen ($I^2 = 83.88\%$; $p < 0.01$). Oleh karena itu, metode estimasi yang digunakan untuk menghitung *overall effect size* adalah model *random-effects*. Berdasarkan model estimasi *random-effect*, diperoleh nilai *overall effect size* sebesar ($d = 1.26$; $p < 0.01$). Nilai ini diklasifikasikan dalam kategori *large*

effect (Cohen, 1988), yang menunjukkan bahwa secara keseluruhan, penerapan teknologi *Duolingo* memiliki pengaruh besar dan signifikan pada pembelajaran bahasa arab jika dibandingkan dengan media tradisional.

Tabel 4. Hasil Perhitungan Overall Effect Size

| Metode Estimasi | k | d | 95% Confidence Interval | p | df | Heterogeneity | | |
|-----------------|---|------|-------------------------|------|----|---------------|------|----------------|
| | | | | | | Q | p | i ² |
| Random-Effect | 8 | 1.26 | [0.71 - 1.80] | 0.00 | 7 | 43.90 | 0.00 | 83.88 |
| Fixed-Effect | 8 | 1.06 | [0.84 - 1.13] | 0.00 | 7 | | | |

Analisis Variabel Moderator

Tujuan kedua dari penelitian ini adalah untuk mengevaluasi heterogenitas dari studi-studi yang dianalisis, dengan melakukan analisis variabel moderator. Analisis ini bertujuan untuk mengidentifikasi faktor-faktor yang dapat mempengaruhi efektivitas penggunaan teknologi *duoligo* dalam pembelajaran bahasa arab. Variabel moderator dalam penelitian ini terdiri dari jenjang pendidikan, kapasitas kelas eksperimen, dan tipe kemampuan yang diukur. Ringkasan analisis variabel moderator dapat dilihat pada Tabel 3 berikut.

Tabel 3. Ringkasan analisis variabel moderator

| Variabel Moderator | k | d | p-value | Heterogeneity | | | |
|----------------------------|---|------|---------|---------------|--------|------|---------|
| | | | | Q-value | df (Q) | Qb | p-value |
| Jenjang Pendidikan | | | | | | | |
| SMP/MTs | 7 | 1.07 | 0.00 | 0.00 | 6 | 0.11 | 0.74 |
| SMA/MA | 1 | 0.96 | 0.00 | 43.76 | 0 | | |
| Kapasitas Kelas Eksperimen | | | | | | | |
| Besar (> 30) | 2 | 0.81 | 0.00 | 0.12 | 1 | 2.71 | 0.10 |
| Kecil (≤ 30) | 6 | 1.19 | 0.00 | 40.37 | 5 | | |
| Tipe Kemampuan | | | | | | | |
| Kosakata | 6 | 1.18 | 0.00 | 37.61 | 5 | 9.56 | 0.04 |
| Mendengar | 1 | 0.41 | 0.18 | 0.00 | 0 | | |
| Minat Belajar | 1 | 0.96 | 0.00 | 0.00 | 0 | | |

Jenjang Pendidikan

Variabel moderator jenjang pendidikan terdiri dari dua kelompok, yaitu SMP/MTs dan SMA/MA. Hasil analisis (lihat Tabel 3) mengungkapkan bahwa rata-

rata *effect size* kedua jenjang pendidikan tidak berbeda signifikan ($Q_b = 0.11$; $p = 0.74$). Hal ini mengindikasikan bahwa variabel jenjang pendidikan tidak berpengaruh terhadap efektivitas penggunaan teknologi *duolingo* pada pembelajaran bahasa arab jika dibandingkan dengan media tradisional. Namun demikian, penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab pada semua jenjang terbukti efektif jika dibandingkan dengan media tradisional. Penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab paling efektif jika diterapkan pada jenjang SMP/MTs ($d = 1,07$; $p < 0,05$) dibandingkan jenjang SMA/MA ($d = 0,96$; $p < 0,05$).

Kapasitas Kelas Eksperimen

Variabel moderator kapasitas kelas eksperimen terdiri dari dua kelompok, yaitu besar (>30) dan kecil (≤ 30). Hasil analisis (lihat Tabel 3) mengungkapkan bahwa rata-rata *effect size* kedua kelompok kapasitas kelas tidak berbeda signifikan ($Q_b = 2.71$; $p = 0.10$). Hal ini mengindikasikan bahwa variabel kapasitas kelas eksperimen tidak berpengaruh terhadap efektivitas penggunaan teknologi *duolingo* pada pembelajaran bahasa arab jika dibandingkan dengan media tradisional. Namun demikian, penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab yang diukur pada kapasitas kelas kecil maupun besar terbukti efektif. Penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab paling efektif jika diterapkan pada kapasitas kelas kecil ($d = 1,19$; $p < 0,05$) dibandingkan kapasitas kelas besar ($d = 0,81$; $p < 0,05$).

Tipe Kemampuan

Variabel moderator tipe kemampuan terdiri dari tiga kelompok, yaitu kosakata, mendengar, dan minat belajar. Hasil analisis (lihat Tabel 3) mengungkapkan bahwa rata-rata *effect size* ketiga kelompok berbeda signifikan ($Q_b = 9.56$; $p = 0.04$). Hal ini mengindikasikan bahwa variabel tipe kemampuan berpengaruh terhadap efektivitas penggunaan teknologi *duolingo* pada pembelajaran bahasa arab jika dibandingkan dengan media tradisional. Dari ketiga kelompok tersebut, Penggunaan teknologi *duolingo* dalam pembelajaran bahasa

arab paling efektif jika diterapkan untuk meningkatkan kemampuan kosakata bahasa arab ($d = 1.18$; $p < 0,05$), dibandingkan dengan kelompok minat ($d = 0.96$; $p < 0,05$). Sementara Penggunaan teknologi *duolingo* dalam pembelajaran bahasa arab tidak efektif untuk meningkatkan kemampuan mendengar bahasa arab jika dibandingkan dengan media tradisional ($d = 0.41$; $p > 0.05$).

Pembahasan

Hasil meta-analisis ini menunjukkan bahwa penggunaan teknologi *Duolingo* secara keseluruhan memberikan pengaruh besar terhadap prestasi belajar bahasa Arab siswa dibandingkan dengan media tradisional. Hal ini tercermin dari nilai *effect size* gabungan sebesar 1,26 (CI 95%: 0,71–1,80), yang dikategorikan sebagai efek besar (*large effect*) berdasarkan interpretasi Cohen (1988). Temuan ini sejalan dengan hasil *systematic review* yang dilakukan oleh Shortt et al. (2023), yang mereview 35 studi terbitan tahun 2012–2020 dan menemukan bahwa *Duolingo* memberikan kontribusi positif dalam pembelajaran bahasa asing. Lebih jauh, hasil ini juga mengonfirmasi pandangan Finardi et al. (2016) yang menekankan bahwa penggunaan aplikasi *mobile-assisted language learning* (MALL) seperti *Duolingo* dapat memperkaya pengalaman belajar melalui pendekatan berbasis personalisasi dan gamifikasi. Temuan Hukom (2024), Samal (2024), dan Sulistyowati et al. (2024) juga menunjukan bahwa pembelajaran yang dipersonalisasi dapat meningkatkan prestasi akademik siswa.

Dari sebelas studi yang dianalisis, variasi ukuran efek menunjukkan adanya keberagaman dalam tingkat keberhasilan penggunaan teknologi *Duolingo* terhadap kemampuan bahasa Arab siswa. Dua studi, yaitu Al-Gifari et al. (2025) dan Yazid et al. (2023), menunjukkan efek yang sangat besar (*very large*). Empat studi, yakni Adira et al. (2024), Idris (2023), Muttaqin (2025), dan Renita (2023), menunjukkan efek besar (*large*), sedangkan satu studi oleh Zakiyah et al. (2024) menunjukkan efek sedang (*moderate*), dan satu studi oleh Baroroh et al. (2022) menunjukkan efek

kecil (*small*). Analisis lebih lanjut dengan menggunakan variabel moderator menunjukkan bahwa efektivitas *Duolingo* berbeda secara signifikan berdasarkan tipe kemampuan yang diukur, namun tidak menunjukkan perbedaan signifikan berdasarkan jenjang pendidikan dan kapasitas kelas eksperimen. Temuan ini mendukung laporan Chen et al. (2020) yang menyatakan bahwa variasi konteks pendidikan tidak selalu menghasilkan perbedaan signifikan dalam efektivitas penggunaan aplikasi MALL.

Penggunaan *Duolingo* terbukti lebih efektif secara signifikan dalam meningkatkan penguasaan kosakata bahasa Arab dibandingkan kemampuan *listening*. Salah satu faktor utama yang mendukung hasil ini adalah karakteristik fitur dalam *Duolingo*, yang lebih banyak berfokus pada aktivitas *drilling* kosakata melalui latihan pencocokan kata, permainan interaktif, dan kuis berbasis teknologi. Pendekatan berbasis pengulangan adaptif ini memfasilitasi proses akuisisi leksikal yang lebih cepat dan meningkatkan daya ingat jangka panjang (Dağdeler, 2023). Meskipun demikian, hasil temuan kami juga menunjukkan bahwa penggunaan *Duolingo* juga secara signifikan meningkatkan kemampuan *listening* siswa dibandingkan dengan media tradisional. Artinya, baik penguasaan kosakata maupun keterampilan menyimak tetap memperoleh efek yang signifikan melalui penggunaan *Duolingo*.

Temuan ini mempertegas bahwa integrasi *Duolingo* dalam pembelajaran bahasa Arab dapat menjadi pendekatan yang transformatif, dengan menghadirkan lingkungan belajar yang lebih interaktif, berbasis data, mendukung motivasi intrinsik, serta memperkuat pembelajaran mandiri (*self-regulated learning*) (Zimmerman, 2002). Di tengah tuntutan pembelajaran abad ke-21, penggunaan *Duolingo* dapat mengatasi keterbatasan metode konvensional yang cenderung satu arah dan tidak adaptif terhadap kebutuhan individual siswa. Oleh karena itu, integrasi *Duolingo* perlu dioptimalkan tidak hanya dalam praktik pembelajaran di kelas, tetapi juga melalui dukungan kebijakan pendidikan, pelatihan guru, serta

penguatan literasi digital dalam rangka menciptakan inovasi instruksional yang berkelanjutan (Selwyn, 2021).

Simpulan

Berdasarkan meta-analisis ini, dapat disimpulkan bahwa penggunaan teknologi *Duolingo* secara signifikan memberikan dampak yang besar pada pembelajaran bahasa Arab jika dibandingkan dengan media tradisional. Selain itu, analisis variabel moderator menunjukkan bahwa variabel tipe kemampuan mempengaruhi efektivitas penggunaan *Duolingo* pada pembelajaran bahasa arab. Artinya bahwa besar kecilnya pengaruh *Duolingo*, tergantung pada tipe kemampuan yang diukur. Namun, hasil analisis heterogenitas menunjukkan bahwa variabel jenjang pendidikan dan kapasitas kelas tidak mempengaruhi efektivitas penggunaan *Duolingo* pada pembelajaran bahasa Arab. Artinya bahwa penggunaan *Duolingo* tetap memberikan efek yang sama baik diterapkan pada berbagai jenjang pendidikan maupun berbagai kapasitas kelas siswa belajar.

Penelitian ini juga memiliki beberap keterbatasan, diantaranya adalah meta-analisis ini hanya melibatkan delapan sampel penelitian. Selain itu, variabel moderator yang dianalisis terbatas pada variabel jenjang, kapasitas kelas, dan tipe kemampuan yang diukur. Penelitian selanjutnya disarankan untuk melibatkan sampel yang lebih banyak bahkan terfokus pada studi-studi yang hanya dipublikasikan pada jurnal bereputasi. Selain itu disarankan juga kajian mengenai efektivitas penggunaan *Duolingo* diperluas dengan mempertimbangkan variasi faktor moderasi lainnya seperti durasi eksperimen maupun integrasi *Duolingo* dengan strategi pembelajaran lain, sehingga memperoleh pemahaman secara komprehensif.

Referensi

- Adira, W., Hamdun, D., Ramadhan, A. R., & Hasna, T. S. (2024). Is Duolingo Media Effective? An Experimental Study on Increasing Student Interest in Arabic Learning. *Arabiyat: Journal of Arabic Education & Arabic Studies/Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 11(2).
- Ajisoko, P. (2020). The use of Duolingo apps to improve English vocabulary learning. *International Journal of Emerging Technologies in Learning (IJET)*, 15(7), 149–155.
- Al-Gifari, R. A., Rahmawati, S. U., & Susiawati, W. (2025). Efektivitas Aplikasi Duolingo dalam Pembelajaran Mufradat di MTs Islamiyah Ciputat. *Jurnal Intelek Insan Cendikia*, 2(3), 5989–5998.
- Baroroh, R. U., Pohan, N. K., & Hania, I. (2022). Pengembangan Aplikasi Duolingo untuk Meningkatkan Kemahiran Mendengar Siswa Kelas VIII di Ma'had al-Ansor dengan Standar CEFR. *PROSIDING KONIPBSA: Konferensi Nasional Inovasi Pembelajaran Bahasa Dan Sastra Arab*, 2(1), 67–74.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons. <https://doi.org/10.1002/9780470743386>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John wiley & sons.
- Burston, J., & Giannakou, K. (2022). MALL language learning outcomes: A comprehensive meta-analysis 1994–2019. *ReCALL*, 34(2), 147–168.
- Cahyani, I. A., Sujarwo, S., Imaroh, Y. R., Hukom, J., Yanuar, F. S., Martaputri, N. A., & Nisrina, N. (2024). Effectiveness of geogebra integration into flipped classroom (gfc) on students mathematics skills: a meta-analysis study. *Jurnal Pendidikan Progresif*, 14(3), 1754–1769.
- Chen, Z., Chen, W., Jia, J., & An, H. (2020). The effects of using mobile devices on language learning: A meta-analysis. *Educational Technology Research and Development*, 68(4), 1769–1789.
- Cohen, J. (1988). The effect size. *Statistical power analysis for the behavioral sciences*. Abingdon: Routledge, 77–83.
- Dağdeler, K. O. (2023). A systematic review of mobile-assisted vocabulary learning research. *Smart Learning Environments*, 10(1), 19.
- Fakhrurriana, R., Nisa, A., & Noni, N. (2024). The Perceptions of Using Duolingo Application in Learning English for Student's Vocabulary Mastery. *BATARA DIDI: English Language Journal*, 3(1), 45–54.
- Finardi, K. R., Leao, R. G., & Amorim, G. B. (2016). Mobile assisted language learning: Affordances and limitations of duolingo. *Education and Linguistics Research*, 2 (2), 48-65.
- Hukom, J. (2024). Penerapan Prinsip Redundansi dalam Media Pembelajaran untuk

- Mengurangi Beban Kognitif. *VARIABLE RESEARCH JOURNAL*, 1(03), 966–970.
- Hukom, J. (2025). Meta-Analysis of the Effectiveness of Computer-Assisted Language Learning (CALL) on Students' Arabic Language Ability. *EL-FUSHA: Jurnal Bahasa Arab Dan Pendidikan*, 6(1), 30–41.
- Hukom, J., Prihatmojo, A., Manaf, A., Suciati, I., Ratau, A., & others. (2023). Integration of Blended Learning and Project-Based Learning (BPjBL) on Achievement of Students' Learning Goals: A Meta-Analysis Study. *Pegem Journal of Education and Instruction*, 13(4), 274–281.
- Idris, M. S., & Kasih, B. W. N. (2023). The Impact of Using Duolingo in Enhancing Arabic Vocabulary Retention for Junior High School Students in Indonesia: في تأثير استخدام وسائل "Duolingo" ترقية حفظ المفردات لطلاب الثانوية بإندونيسيا. *Journal of Arabic Language Learning and Teaching*, 1(1), 47–60.
- Irawan, A., Wilson, A., & Sutrisno, S. (2020). The implementation of duolingo mobile application in English vocabulary learning. *Scope: Journal of English Language Teaching*, 5(1), 8–14.
- Jaelani, A., & Sutari, D. R. (2021). Students' perception Of Using Duolingo Application As A Media In Learning Vocabulary. *Bogor English Student and Teacher (BEST) Conference*, 2, 40–47.
- Jiang, X., Rollinson, J., Plonsky, L., Gustafson, E., & Pajak, B. (2021). Evaluating the reading and listening outcomes of beginning-level Duolingo courses. *Foreign Language Annals*, 54(4), 974–1002.
- Kamsurya, M. A., Wijaya, A., Ramadhani, R., & Hukom, J. (2022). The effect of self-efficacy on students' mathematical abilities: a meta-analysis study. *Jurnal Pendidikan Progresif*, 12(2), 451–463.
- Kazu, İ. Y., & Kuvvetli, M. (2025). Digital language learning with Duolingo: assessing its impact on listening, speaking, reading, and writing skills. *Journal of Computers in Education*, 1–27.
- Li, Z., & Bonk, C. J. (2025). Self-directed language learning with Duolingo in an out-of-class context. *Computer Assisted Language Learning*, 38(3), 569–591.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications, Inc.
- Mawardi, D. N., Sulistyowati, E., & Hukom, J. (2024). Meta-Analysis investigasi model kelas terbalik pada keterampilan berpikir tingkat tinggi (HOTS) siswa matematika: Analisis efek gabungan dan heterogenitas. *Jurnal Math Educator Nusantara: Wahana Publikasi Karya Tulis Ilmiah Di Bidang Pendidikan Matematika*, 10(1), 154–166.
- Meece, J. L. (2023). The role of motivation in self-regulated learning. In *Self-regulation of learning and performance* (pp. 25–44). Routledge.

- Muhtadi, A., Assagaf, G., & Hukom, J. (2022). Self-efficacy and students' mathematics learning ability in Indonesia: A meta analysis study. *International Journal of Instruction*, 15(3), 1131–1146.
- Muttaqin, M., Zuhdi, H., & Ridwan, R. (2025). The Use of Gamification-Based Duolingo Application in Increasing Student Motivation is Reviewed From the Theory of Self-Determination. *AL-WIJDĀN Journal of Islamic Education Studies*, 10(1), 81–95.
- Nah, F. F.-H., Telaprolu, V. R., Rallapalli, S., & Venkata, P. R. (2013). Gamification of education using computer games. *Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business, 15th International Conference, HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part III* 15, 99–107.
- Octaviani, S. K., & others. (2022). Duolingo mobile application for English listening skill improvement of vocational school students. *Journal of English Language and Education*, 7(2), 84–95.
- Payne, J. S. (2020). Developing L2 productive language skills online and the strategic use of instructional tools. *Foreign Language Annals*, 53(2), 243–249.
- Purnomo, B., Muhtadi, A., Ramadhani, R., Manaf, A., & Hukom, J. (2022). The Effect of Flipped Classroom Model on Mathematical Ability: A Meta Analysis Study. *Jurnal Pendidikan Progresif*, 12(3), 1201–1217. <https://doi.org/10.23960/jpp.v12.i3.202216>
- Renita, H. (2023). *Pengaruh Penggunaan Duolingo Terhadap Penguasaan Kosakata Dalam Pembelajaran Bahasa Arab Siswa Mts Muhammadiyah 2 Jenangan Ponorogo*. Universitas Muhammadiyah Yogyakarta.
- Samal, D. (2024). Evaluation of blended learning-based utilization using CSE-UCLA Model. *Pegem Journal of Education and Instruction*, 14(2), 131–139.
- Samritin, S., Susanto, A., Manaf, A., & Hukom, J. (2023). A meta-analysis study of the effect of the blended learning model on students' mathematics learning achievement. *Jurnal Elemen*, 9(1), 15–30.
- Selwyn, N. (2021). *Education and technology: Key issues and debates*. Bloomsbury Publishing.
- Setiawan, A. A., Muhtadi, A., & Hukom, J. (2022). Blended Learning and Student Mathematics Ability in Indonesia: A Meta-Analysis Study. *International Journal of Instruction*, 15(2), 905–916. <https://doi.org/10.29333/iji.2022.15249a>
- Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, 36(3), 517–554.
- Sulistiyowati, E., Hukom, J., & Muhtadi, A. (2023). Meta-analysis of flipped classroom

- on students' mathematics abilities: effectiveness and heterogeneity analysis. *JTP-Jurnal Teknologi Pendidikan*, 25(2), 140–159.
- Sulistiyowati, E., Rohman, A., & Hukom, J. (2024). Flipped classroom model: Minimizing gaps in understanding mathematical concepts for students with different academic abilities. *European Journal of Mathematics and Science Education*, 5(1), 27–37.
- Ulum, F., & Hukom, J. (2025). Flipped Learning in Foreign Language Learning in Higher Education: Analysis of Effectiveness and Moderator Variables. *Jurnal Pendidikan Progresif*, 15(2), 1025–1040.
- Wagner, E. (2020). Duolingo English test, revised version July 2019. *Language Assessment Quarterly*, 17(3), 300–315.
- Yazid, A., Ma'arif, M. S., & Firdaus, M. A. (2023). The effect of duolingo media application in improving arabic vocabulary mastery of class VIII Students at MTsN 6 Nganjuk. *Al-Wasil*, 1(1), 26–49.
- Yu, A., & Trainin, G. (2022). A meta-analysis examining technology-assisted L2 vocabulary learning. *ReCALL*, 34(2), 235–252.
- Zakiah, F., Renita, H., & Mirma, A. F. (2024). The Effect of Using Duolingo on Vocabulary Mastery in Arabic Learning for MTs Muhammadiyah 2 Jenangan Ponorogo Students. *Al Mahara: Jurnal Pendidikan Bahasa Arab*, 10(2), 420–434.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70.
- Zuliana, E., Dwiningrum, S. I. A., Wijaya, A., & Hukom, J. (2025). The effect of culture-based mathematics learning instruction on mathematical skills: a meta-analytic study. *Journal of Education and Learning (EduLearn)*, 19(1), 191–201.