PENGUATAN KAPASITAS PENULISAN ILMIAH GURU GEOGRAFI MELALUI *DRAFT-BASED WRITING CLINIC* DI DKI JAKARTA

Muhammad Zid[1], Della Ayu Lestari[2], Sony Nugratama Hijrawadi[1], Lona Syafana Pasya[1], Tri Wandi Januar[2], Hikari Dwi Saputro[1], Fauzi Ramadhoan A'R[1], Wini Mustikarani[1], Denny Mawardi[3], Waldi Rohmatullah[3], Mahfuzh Anshori[3], Nono[3], Singgih Denny Santosa[3], Afiifah Huwaidatun Nisaa Khoiriyyah[1], Imam Nasuha[1]

[1]Program Studi Pendidikan Geografi, Fakultas Ilmu Sosial, Universitas Negeri Jakarta, Indonesia

[2]Program Studi Geografi, Fakultas Ilmu Sosial, Universitas Negeri Jakarta, Indonesia

[3]MGMP Geografi DKI Jakarta, Indonesia

zidmuhammad@unj.ac.id; dellaayulestari@unj.ac.id; sonynugratama@unj.ac.id; lonasyafana@unj.ac.id; triwandijanuar@unj.ac.id; hikari.dwi@unj.ac.id; fauzi.ramadhoan@unj.ac.id; wini.mustikarani@unj.ac.id

Abstract

Scientific writing is a core professional demand for geography teachers, yet training programs rarely report measured evidence of changes in writing capacity. This article evaluates a one-day draft-based writing clinic *for the Geography Teachers Association (MGMP) of DKI Jakarta held on 8 May 2026. The clinic positioned teachers' article drafts as the main object of learning and combined scientific writing with three supporting topics: Generative AI and StoryMaps, Google Earth Engine, and NVivo. A one-group pre-test and post-test design was applied to 79 valid pre-test and 63 valid post-test responses, of which 61 could be matched individually. Because the instruments were not fully identical, change analysis was restricted to 13 equivalent items and tested using the Wilcoxon Signed-Rank test. Overall mastery rose from 79.7% to 82.7% ($p = 0.062$), a positive but statistically non-significant trend attributable to a ceiling effect in the initial scores. Knowledge of scientific writing concepts remained high and stable (83.6% to 82.3%; $p = 0.591$), whereas Google Earth Engine improved significantly (85.2% to 92.1%; $p = 0.006$) and readiness to develop StoryMaps rose from 3.36 to 3.57 on a four-point scale ($p = 0.014$). Final mastery was lowest for NVivo (63.8%). A multiple-choice test captures conceptual knowledge but not drafting skill; strengthening teachers' scientific writing capacity therefore requires sustained draft-based mentoring and performance-based assessment.*

Keywords: *scientific writing; writing clinic; geography teachers; teacher professional development; pre-test post-test*

Abstrak

Guru geografi dituntut mampu menulis karya ilmiah, tetapi program pelatihan jarang melaporkan bukti terukur mengenai perubahan kapasitas menulis. Artikel ini mengevaluasi draft-based writing clinic *bagi MGMP Geografi DKI Jakarta pada 8 Mei 2026 yang menempatkan draf tulisan guru sebagai objek belajar utama dan memadukan penulisan ilmiah dengan tiga materi penunjang, yaitu Generative AI dan StoryMaps, Google Earth Engine, serta NVivo. Desain yang digunakan adalah satu kelompok dengan pre-test dan post-test terhadap 79 respons pre-test valid dan 63 respons post-test valid, dengan 61 responden yang dapat dipasangkan. Karena susunan soal tidak sepenuhnya sama, analisis perubahan dibatasi pada 13 butir yang setara dan diuji dengan Wilcoxon Signed-Rank. Penguasaan keseluruhan meningkat dari 79,7% menjadi 82,7% ($p = 0,062$), yaitu positif tetapi belum signifikan karena skor awal peserta sudah tinggi. Pengetahuan konseptual penulisan ilmiah relatif stabil pada tingkat tinggi (83,6% menjadi 82,3%; $p = 0,591$), sedangkan Google Earth Engine meningkat signifikan dari 85,2% menjadi 92,1% ($p = 0,006$) dan kesiapan membuat StoryMaps naik dari 3,36 menjadi 3,57 pada skala empat ($p = 0,014$). Penguasaan akhir terendah terdapat pada NVivo (63,8%). Tes pilihan ganda hanya menjangkau pengetahuan konseptual dan belum menangkap keterampilan menyusun draf, sehingga penguatan kapasitas menulis ilmiah guru memerlukan pendampingan draf yang berkelanjutan serta penilaian berbasis kinerja.*

Kata Kunci: *penulisan ilmiah;* writing clinic*; guru geografi; pengembangan keprofesian guru; pre-test post-test*

1. PENDAHULUAN (Introduction)

Penulisan ilmiah merupakan kompetensi keprofesian yang semakin dituntut dari guru geografi. Publikasi ilmiah diposisikan sebagai salah satu unsur pengembangan keprofesian berkelanjutan di samping pengembangan diri dan karya inovatif (Peraturan Menteri Negara Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 16 Tahun 2009), dan menulis menjadi sarana bagi guru untuk mendokumentasikan, merefleksikan, serta membagikan praktik baik pembelajaran kepada komunitas yang lebih luas. Dalam pembelajaran geografi, tuntutan tersebut beririsan dengan meluasnya ketersediaan data spasial terbuka sehingga guru tidak hanya menjelaskan konsep keruangan, tetapi juga memiliki bahan empiris yang dapat diolah menjadi tulisan ilmiah. Musyawarah Guru Mata Pelajaran (MGMP) menjadi wadah strategis untuk memenuhi kebutuhan tersebut karena bersifat kolegial, berkelanjutan, dan dekat dengan persoalan nyata di sekolah.

Penulisan ilmiah, bagaimanapun, bukan sekadar keterampilan teknis yang dapat dialihkan melalui paparan materi. Model kognitif menggambarkan menulis sebagai proses rekursif yang melibatkan perencanaan, penuangan gagasan, dan peninjauan ulang (Flower & Hayes, 1981), sedangkan kajian wacana akademik menempatkannya sebagai praktik sosial yang terikat pada konvensi komunitas disiplin (Hyland, 2016). Implikasinya, penguatan kapasitas menulis lebih tepat diarahkan pada pengerjaan tulisan yang sesungguhnya, yakni draf yang dikomentari, direvisi, dan didiskusikan secara berulang, daripada pada penyampaian kaidah penulisan semata (Lee & Boud, 2003; Aitchison & Lee, 2006). Gagasan inilah yang melandasi *draft-based writing clinic*, yaitu format pendampingan yang menjadikan draf tulisan peserta sebagai objek belajar utama.

Kajian tentang pengembangan keprofesian guru menunjukkan bahwa dampak lebih besar dicapai apabila kegiatan berfokus pada konten, melibatkan pembelajaran aktif, berlangsung dalam durasi memadai, dan terhubung dengan praktik (Desimone, 2009; Darling-Hammond, Hyler, & Gardner, 2017). Penulisan artikel geografi yang memanfaatkan data geospasial dan data kualitatif juga menuntut kompetensi teknologi yang menopang proses menulis, misalnya Generative AI untuk membantu penyusunan dan penyuntingan, StoryMaps untuk menyajikan narasi keruangan, Google Earth Engine untuk menghasilkan data empiris wilayah, serta NVivo untuk mengolah data kualitatif (Gorelick et al., 2017; Cope, Mikhailova, Post, Schlautman, & Carbajales-Dale, 2018; Woods, Paulus, Atkins, & Macklin, 2016). Kerangka Technological Pedagogical Content Knowledge menegaskan bahwa penguasaan perangkat tidak otomatis berubah menjadi praktik yang bermakna apabila tidak dipadukan dengan pengetahuan konten dan pedagogi (Mishra & Koehler, 2006).

Meskipun demikian, banyak laporan kegiatan pengabdian kepada masyarakat bertema penulisan ilmiah guru di Indonesia masih berhenti pada deskripsi keterlaksanaan dan kepuasan peserta. Bukti perubahan yang diuji secara statistik, terutama yang mempertimbangkan kesetaraan butir soal pre-test dan post-test, masih jarang disajikan. Kebaruan artikel ini terletak pada evaluasi terukur atas pelatihan penulisan ilmiah berbasis draf yang menggabungkan analisis butir setara, uji nonparametrik terhadap data berpasangan, serta pemisahan antara perubahan pengetahuan dan perubahan kesiapan menggunakan teknologi penunjang penulisan.

Permasalahan yang dijawab adalah sejauh mana *draft-based writing clinic* memperkuat pengetahuan dan kesiapan guru geografi, serta aspek mana yang paling dan paling kurang berkembang. Tujuan artikel ini adalah (1) mendeskripsikan perubahan penguasaan peserta pada butir yang setara, (2) menguji signifikansi perubahan tersebut, (3) memetakan tingkat penguasaan pada akhir kegiatan, dan (4) merumuskan rekomendasi pengembangan klinik penulisan lanjutan bagi MGMP Geografi DKI Jakarta.

2. TINJAUAN LITERATUR (Literature Review)

Menulis ilmiah dipahami sebagai proses rekursif, bukan urutan linear dari perencanaan hingga penyuntingan. Flower dan Hayes (1981) memperlihatkan bahwa penulis berpindah-pindah antara merencanakan, menuangkan, dan meninjau teks sambil terus menyesuaikan tujuan retoris. Dari sisi sosial, Hyland (2016) menegaskan bahwa tulisan akademik dibentuk oleh genre dan konvensi komunitas pembacanya, sehingga penulis pemula perlu menginternalisasi konvensi tersebut melalui keterlibatan dengan teks yang nyata. Panduan praktis seperti Cargill dan O'Connor (2013) serta Murray (2013) menguraikan langkah penyusunan artikel riset dan strategi menulis secara terjadwal, tetapi penerapannya tetap memerlukan latihan dan umpan balik yang kontekstual. Pada guru, keyakinan atas kemampuan menulis (*self-efficacy*) turut memengaruhi motivasi dan capaian menulis (Pajares, 2003), sehingga penguatan kapasitas perlu memperhatikan kepercayaan diri di samping pengetahuan.

Klinik penulisan berbasis draf berpijak pada tiga gagasan. Pertama, umpan balik efektif ketika menjawab pertanyaan tentang tujuan, capaian saat ini, dan langkah berikutnya, serta diarahkan pada tugas, proses, dan regulasi diri (Hattie & Timperley, 2007). Kedua, umpan balik yang baik mendorong pembelajar menilai tulisannya sendiri dan memberi kesempatan menutup kesenjangan melalui revisi (Nicol & Macfarlane-Dick, 2006); dalam penulisan, efektivitasnya juga bergantung pada konteks dan relasi antara pemberi dan penerima umpan balik (Hyland & Hyland, 2006). Ketiga, dukungan terstruktur yang dikurangi secara bertahap sejalan dengan konsep *scaffolding* (Wood, Bruner, & Ross, 1976). Pada tataran kelembagaan, kelompok menulis dipandang sebagai strategi pengembangan riset yang menempatkan menulis sebagai titik awal, bukan titik akhir, proses penelitian (Lee & Boud, 2003), dan pedagogi pembimbingan penulisan doktoral memperlakukan teks serta umpan balik sebagai inti pembelajaran (Kamler & Thomson, 2014). Komunitas praktik seperti MGMP menyediakan ruang sosial bagi proses tersebut (Wenger, 1998).

Efektivitas pengembangan keprofesian guru dipengaruhi oleh fokus konten, pembelajaran aktif, koherensi, durasi, dan partisipasi kolektif (Desimone, 2009), serta oleh pemodelan praktik, dukungan ahli, umpan balik, dan refleksi (Darling-Hammond et al., 2017). Kennedy (2016) menambahkan bahwa efektivitas ditentukan oleh cara program memfasilitasi pembelajaran guru, tidak sekadar oleh daftar fitur program, sedangkan Guskey (2002) menunjukkan bahwa perubahan keyakinan dan sikap guru umumnya muncul setelah mereka melihat bukti keberhasilan praktik baru. Rujukan tersebut menjadi dasar untuk membaca pelatihan satu hari sebagai titik awal siklus pendampingan, bukan sebagai intervensi tunggal. Evaluasi dalam artikel ini berada pada tingkat pembelajaran dalam kerangka Kirkpatrick dan Kirkpatrick (2006), yakni perubahan pengetahuan dan kesiapan, sedangkan perubahan perilaku menulis dan hasil berupa publikasi memerlukan penelusuran jangka yang lebih panjang.

Materi penunjang yang dilatihkan memiliki karakter yang berbeda. Generative AI dapat membantu pengembangan gagasan dan penyuntingan bahasa, tetapi disertai peringatan mengenai halusinasi informasi, bias, serta isu kepenulisan dan integritas akademik (Kasneci et al., 2023; Dwivedi et al., 2023). Karena itu, literasi AI mencakup kemampuan mengevaluasi keluaran secara kritis, bukan sekadar mengoperasikan aplikasi (Ng, Leung, Chu, & Qiao, 2021; Zawacki-Richter, Marín, Bond, & Gouverneur, 2019; Chiu, Xia, Zhou, Chai, & Cheng, 2023). Google Earth Engine menyediakan pemrosesan citra berbasis komputasi awan untuk analisis tutupan lahan, vegetasi, dan suhu permukaan perkotaan (Gorelick et al., 2017; Tamiminia et al., 2020; Amani et al., 2020), sedangkan StoryMaps berperan sebagai media naratif geospasial yang dipersepsikan positif oleh guru (Strachan & Mitchell, 2014; Cope et al., 2018). NVivo menuntut pemahaman prosedural bertahap, mulai dari penyiapan data, pembuatan Cases, pengodean, hingga query lanjutan (Woods et al., 2016; Dhakal, 2022). Adopsi teknologi oleh guru dipengaruhi ekspektasi kinerja, ekspektasi usaha, dan kondisi pendukung (Venkatesh, Thong, & Xu, 2012), sehingga kesiapan menggunakan teknologi relevan diukur berdampingan dengan pengetahuan. Perbedaan karakter materi inilah yang diduga menghasilkan pola perubahan yang tidak seragam antardomain.

3. METODE PELAKSANAAN (Materials and Method)

Kegiatan dilaksanakan pada 8 Mei 2026 bersama MGMP Geografi DKI Jakarta dalam bentuk pelatihan satu hari dengan empat materi, yaitu penulisan artikel ilmiah sebagai materi inti, serta pemanfaatan Generative AI dan StoryMaps, analisis geospasial dengan Google Earth Engine, dan pengolahan data kualitatif menggunakan NVivo sebagai materi penunjang penulisan. Setiap sesi disampaikan melalui paparan konsep, demonstrasi, dan praktik terbatas.

3.1 Rancangan *draft-based writing clinic*

Rancangan kegiatan mengacu pada logika klinik penulisan berbasis draf, yaitu menjadikan draf tulisan guru sebagai objek belajar utama dan memosisikan umpan balik serta revisi sebagai inti proses belajar. Logika tersebut dioperasionalkan dalam lima tahap berurutan, yakni (1) diagnosis bagian artikel yang perlu diperkuat, (2) pemodelan melalui contoh artikel, (3) praktik terbimbing pada bagian draf, (4) umpan balik dari fasilitator dan sejawat, serta (5) revisi sebagai tindak lanjut (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). Materi penunjang diposisikan sebagai alat untuk menghasilkan dan menyajikan isi artikel, yaitu data spasial, narasi keruangan, dan analisis data kualitatif. Karena durasi satu hari hanya memungkinkan tahap awal siklus tersebut, pelatihan diposisikan sebagai pemicu bagi siklus umpan balik dan revisi berikutnya di lingkungan MGMP.

3.2 Desain dan partisipan

Evaluasi menggunakan desain satu kelompok dengan pre-test dan post-test. Instrumen berbentuk pilihan ganda yang disebarkan melalui Google Forms sebelum dan sesudah kegiatan, dilengkapi dua butir persepsi kesiapan menggunakan skala satu sampai empat. Data awal berjumlah 80 respons pre-test dan 64 respons post-test; masing-masing memuat satu respons uji coba yang dikeluarkan sehingga diperoleh 79 dan 63 respons valid. Pencocokan berdasarkan alamat surat elektronik menghasilkan 61 pasangan data.

3.3 Rekonstruksi skor dan dasar perbandingan

Skor otomatis Google Forms tidak dapat digunakan karena seluruhnya tercatat 0/0, sehingga skor direkonstruksi dari pilihan jawaban peserta terhadap kunci jawaban, dengan nilai satu untuk jawaban benar dan nol untuk jawaban salah. Karena susunan soal pre-test dan post-test tidak sepenuhnya sama, analisis perubahan pengetahuan dibatasi pada 13 butir yang identik secara substansi dan pilihan jawaban, terdiri atas lima butir penulisan ilmiah, tiga butir AI dan StoryMaps, serta lima butir Google Earth Engine. Butir NVivo pada kedua tes berbeda cakupan sehingga hanya dilaporkan sebagai capaian akhir dan tidak ditafsirkan sebagai gain.

3.4 Analisis data

Analisis mencakup rerata, persentase penguasaan, selisih skor, distribusi peserta yang meningkat, tetap, dan menurun, serta uji Wilcoxon Signed-Rank untuk data berpasangan (Wilcoxon, 1945) pada taraf signifikansi 5%. Uji nonparametrik dipilih karena skor bersifat diskret, rentangnya terbatas, dan terdapat indikasi efek plafon pada beberapa domain, suatu kondisi yang dapat membiaskan estimasi besaran perubahan (Wang, Zhang, McArdle, & Salthouse, 2008). Evaluasi dibatasi pada tingkat pembelajaran dalam kerangka Kirkpatrick dan Kirkpatrick (2006). Ringkasan data yang dianalisis disajikan pada Tabel 1.

Tabel 1. Ringkasan data yang dianalisis

Komponen data	Jumlah	Keterangan
Respons pre-test valid	79	Setelah satu respons uji coba dikeluarkan
Respons post-test valid	63	Setelah satu respons uji coba dikeluarkan
Responden berpasangan	61	Dasar analisis perubahan
Butir identik	13	5 menulis ilmiah, 3 AI/StoryMaps, 5 GEE
Butir post-test	23	Dasar pemetaan penguasaan akhir

4. HASIL DAN PEMBAHASAN (Results and Discussion)

4.1 Perubahan penguasaan secara keseluruhan

Pada 61 responden berpasangan, rerata skor 13 butir identik meningkat 0,39 poin, dari 10,36 menjadi 10,75 dari skor maksimum 13. Penguasaan bertambah dari 79,7% menjadi 82,7% atau naik 3,0 poin persentase. Sebanyak 25 peserta mengalami peningkatan, 21 peserta memperoleh skor tetap, dan 15 peserta menurun. Uji Wilcoxon menghasilkan $p = 0,062$ sehingga perubahan keseluruhan menunjukkan kecenderungan positif, tetapi belum signifikan pada taraf 5%. Rangkuman hasil disajikan pada Tabel 2.

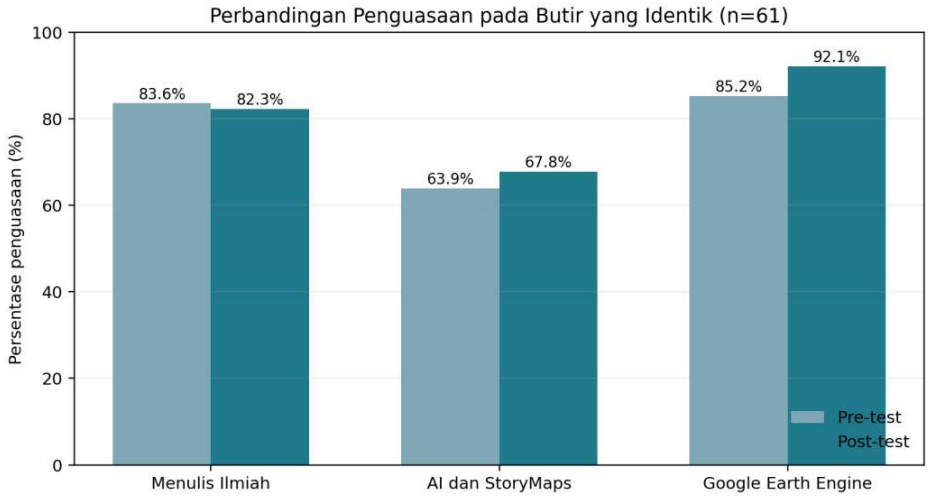
Tabel 2. Perubahan hasil pre-test dan post-test secara keseluruhan (n = 61)

Indikator	Pre-test	Post-test	Perubahan	Interpretasi
Rerata skor (maks. 13)	10,36	10,75	+0,39	Meningkat
Persentase penguasaan	79,7%	82,7%	+3,0 p.p.	Meningkat
Uji Wilcoxon	–	–	$p = 0,062$	Belum signifikan
Distribusi perubahan	15 menurun	25 meningkat	21 tetap	Cenderung positif

Hasil tersebut perlu dibaca bersama kondisi skor awal yang relatif tinggi. Ketika sebagian besar peserta telah menjawab benar pada pre-test, ruang peningkatan pada post-test menjadi sempit sehingga muncul efek plafon, terutama pada materi penulisan ilmiah dan Google Earth Engine; kondisi ini diketahui dapat membiaskan estimasi besaran perubahan (Wang et al., 2008). Nilai p yang mendekati batas signifikansi menunjukkan arah perubahan yang konsisten, namun ukuran sampel berpasangan dan jumlah butir yang terbatas menahan kekuatan uji. Dengan demikian, temuan ini tidak dibaca sebagai bukti bahwa kapasitas menulis tidak berkembang, melainkan sebagai petunjuk bahwa instrumen pilihan ganda kurang peka terhadap perubahan yang diharapkan dari klinik penulisan.

4.2 Penguasaan penulisan ilmiah dan materi penunjang

Analisis per domain memperlihatkan pola yang tidak seragam. Pada domain inti, penguasaan konsep penulisan ilmiah bergerak dari 83,6% menjadi 82,3% dengan selisih -1,3 poin persentase yang tidak signifikan ($p = 0,591$). Pada materi penunjang, domain AI dan StoryMaps meningkat dari 63,9% menjadi 67,8% ($p = 0,327$), sedangkan Google Earth Engine meningkat dari 85,2% menjadi 92,1% dan signifikan secara statistik ($p = 0,006$). Perbandingan visual disajikan pada Gambar 1 dan ringkasan angkanya pada Tabel 3.



Gambar 1. Perbandingan persentase penguasaan pada butir yang identik

Tabel 3. Perbandingan penguasaan berdasarkan domain materi

Domain	Pre (%)	Post (%)	Perubahan	p-value	Makna
Menulis Ilmiah	83,6	82,3	-1,3 p.p.	0,591	Belum signifikan
AI dan StoryMaps	63,9	67,8	+3,8 p.p.	0,327	Belum signifikan
Google Earth Engine	85,2	92,1	+6,9 p.p.	0,006	Signifikan

Penurunan kecil pada penulisan ilmiah tidak dapat langsung dimaknai sebagai penurunan kompetensi. Pengetahuan dasar peserta mengenai objektivitas tulisan, penggunaan data dan fakta, serta perbedaan artikel dan opini sudah tinggi sejak awal, sehingga selisih kecil lebih mungkin mencerminkan variasi konsentrasi dan kehati-hatian menjawab daripada kemunduran pemahaman. Lebih mendasar lagi, butir pilihan ganda mengukur pengetahuan deklaratif tentang kaidah menulis, sedangkan kapasitas menulis menuntut keterampilan prosedural yang baru tampak ketika penulis merencanakan, menyusun, dan merevisi teks (Flower & Hayes, 1981). Kesenjangan antara mengetahui kaidah dan mampu menerapkannya inilah yang menjadi alasan utama pendekatan berbasis draf. Implikasinya, klinik lanjutan sebaiknya tidak lagi mengukur konsep dasar, melainkan menilai kemampuan menyusun rumusan masalah, memilih metode, mengembangkan argumen, dan merevisi draf artikel dengan bantuan rubrik (Hyland, 2016).



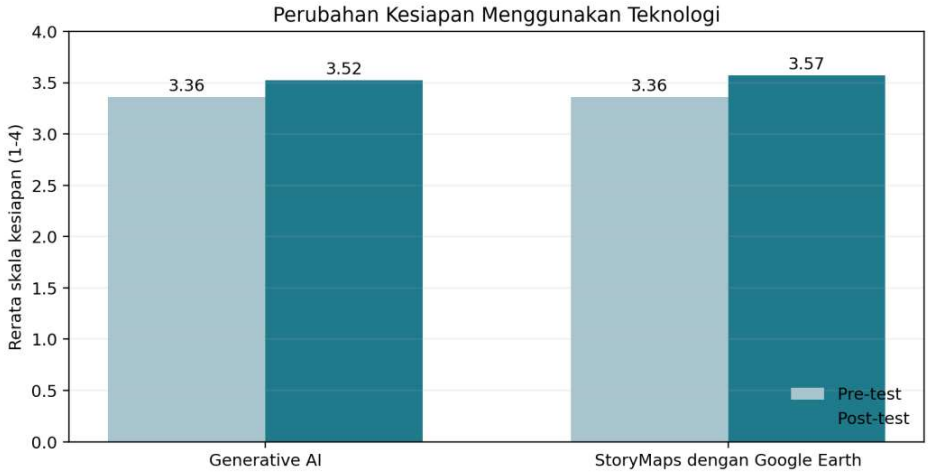
Gambar 2. Pelaksanaan kegiatan pelatihan

Pada domain AI dan StoryMaps, peserta relatif memahami konsep halusinasi AI dan fungsi umum StoryMaps, tetapi masih keliru membedakan platform Generative AI berbasis teks, mengenali batas kemampuan StoryMaps, memahami alasan pedagogis pemilihan Google Earth versi web, serta menerapkan komponen kerangka KORE dalam menyusun perintah. Bagi penulis artikel, kelemahan ini berdampak langsung karena ketepatan perintah dan verifikasi keluaran menentukan apakah Generative AI membantu atau justru menyisipkan informasi keliru ke dalam draf (Kasneci et al., 2023; Dwivedi et al., 2023). Temuan ini sejalan dengan gagasan literasi AI yang menempatkan kemampuan evaluasi kritis di atas keterampilan operasional (Ng et al., 2021; Chiu et al., 2023).

Peningkatan paling kuat pada Google Earth Engine terutama ditemukan pada pemahaman NDVI dan Urban Heat Island. Penyajian konsep yang dikaitkan langsung dengan data satelit dan fenomena perkotaan tampak mudah dihubungkan dengan pengetahuan geografi yang telah dimiliki peserta, sebagaimana potensi yang dilaporkan dalam kajian pemanfaatan platform tersebut (Gorelick et al., 2017; Tamiminia et al., 2020). Dari perspektif penulisan, kemampuan ini penting karena menyediakan data empiris wilayah DKI Jakarta yang dapat dikembangkan menjadi bahan artikel ilmiah.

4.3 Perubahan kesiapan menggunakan teknologi penunjang penulisan

Selain pengetahuan, evaluasi mengukur kesiapan peserta menggunakan teknologi pada skala satu sampai empat. Rerata kesiapan menggunakan Generative AI meningkat dari 3,36 menjadi 3,52 dengan p = 0,060, sedangkan kesiapan membuat StoryMaps dengan Google Earth meningkat dari 3,36 menjadi 3,57 dan signifikan (p = 0,014). Pada kedua indikator, proporsi peserta yang memberi penilaian positif naik dari 85,2% menjadi 93,4%. Perubahan tersebut ditampilkan pada Gambar 2 dan Tabel 4.



Gambar 3. Perubahan rerata kesiapan peserta menggunakan teknologi pembelajaran

Tabel 4. Perubahan kesiapan menggunakan teknologi (skala 1–4)

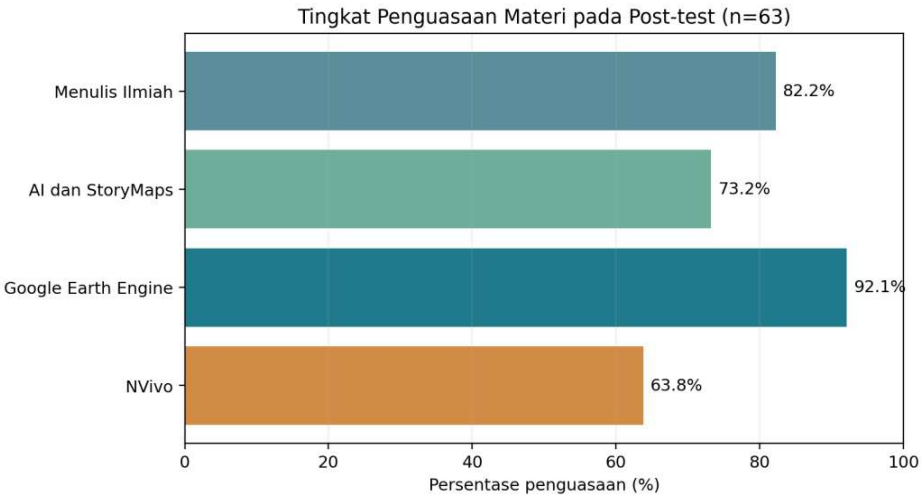
Indikator kesiapan	Pre	Post	Respons positif	p-value	Makna
Menggunakan Generative AI	3,36	3,52	85,2% → 93,4%	0,060	Belum signifikan
Membuat StoryMaps	3,36	3,57	85,2% → 93,4%	0,014	Signifikan

Signifikannya peningkatan kesiapan StoryMaps merupakan temuan penting karena kesiapan merupakan prasyarat penerapan di sekolah dan berkaitan dengan keyakinan diri guru, yang pada ranah

menulis memengaruhi motivasi dan capaian (Pajares, 2003). Ketika guru menilai suatu perangkat mudah diakses, relevan dengan materi, dan dapat digunakan bersama murid, peluang adopsinya meningkat, sesuai penjelasan kerangka penerimaan teknologi mengenai peran ekspektasi usaha dan kondisi pendukung (Venkatesh et al., 2012). Kesiapan menggunakan Generative AI yang belum signifikan menunjukkan bahwa penerimaan terhadap AI membutuhkan pengalaman praktik lebih banyak, khususnya pada penyusunan perintah, verifikasi fakta, etika penggunaan, dan pendokumentasian bantuan AI dalam draf (Kasneci et al., 2023).

4.4 Tingkat penguasaan pada akhir kegiatan

Berdasarkan seluruh 23 butir post-test, rerata skor peserta adalah 17,76 atau 77,2% dari skor maksimum. Penguasaan tertinggi terdapat pada Google Earth Engine sebesar 92,1%, diikuti penulisan ilmiah 82,2%, AI dan StoryMaps 73,2%, serta NVivo 63,8%. Penguasaan akhir pada penulisan ilmiah tergolong relatif tinggi, sehingga fokus pendampingan berikutnya dapat bergeser dari konsep dasar ke keterampilan menyusun dan merevisi draf. Sebaran penguasaan disajikan pada Gambar 3 dan Tabel 5.



Gambar 4. Tingkat penguasaan materi berdasarkan seluruh butir post-test

Tabel 5. Tingkat penguasaan materi pada akhir pelatihan (n = 63)

Materi	Rerata skor	Skor maksimum	Penguasaan
Menulis Ilmiah	4,11	5	82,2%
AI dan StoryMaps	5,86	8	73,2%
Google Earth Engine	4,60	5	92,1%
NVivo	3,19	5	63,8%
Total	17,76	23	77,2%

Sumber: Data analisis, 2026

Analisis butir memperlihatkan enam area yang masih memerlukan penguatan, yaitu penerapan kerangka KORE dalam menyusun perintah AI, penggunaan Matrix Coding Query untuk membandingkan tema antar-kasus, pemaknaan references sebagai jumlah kutipan yang telah dikodekan, sinkronisasi atribut data dengan Cases, alasan pedagogis penggunaan Google Earth versi web, serta batas fungsi StoryMaps sebagai media naratif geospasial dan bukan perangkat analisis statistik. Keenam area tersebut dapat

dijadikan sasaran umpan balik terfokus pada siklus klinik berikutnya, misalnya melalui lembar diagnostik yang menautkan setiap area dengan bagian draf artikel yang relevan.

4.5 Pembahasan

Secara keseluruhan, kegiatan menghasilkan dampak yang berbeda pada tiap domain. Materi yang bersifat visual, kontekstual, dan langsung berkaitan dengan fenomena wilayah menunjukkan peningkatan paling kuat, sedangkan materi yang menuntut penguasaan prosedur berkembang lebih lambat. Pola ini konsisten dengan kerangka TPACK yang menyatakan bahwa penguasaan teknologi hanya bermakna ketika terhubung dengan pengetahuan konten dan pedagogi yang telah dimiliki guru (Mishra & Koehler, 2006). Google Earth Engine memiliki keterhubungan yang kuat dengan konsep geografi sehingga materi baru dapat dilekatkan pada struktur pengetahuan yang sudah mapan. Penulisan ilmiah, sebaliknya, merupakan keterampilan lintas konten yang tidak cukup diukur melalui butir konsep, sehingga stabilitas skor pada domain ini lebih tepat dibaca sebagai pengetahuan dasar yang sudah memadai, bukan sebagai tanda bahwa kapasitas menulis telah terbentuk.

Sebaliknya, NVivo dan penyusunan perintah terstruktur menuntut rangkaian tindakan berurutan, mulai dari mengimpor data, membuat Cases, melakukan pengodean, menjalankan query, hingga membaca visualisasi. Keterampilan semacam ini sulit terbentuk dalam satu sesi dan memerlukan latihan terbimbing yang berulang (Woods et al., 2016; Dhakal, 2022). Hal serupa berlaku bagi keterampilan menulis: pemodelan oleh fasilitator yang diikuti pengurangan bantuan secara bertahap sebagaimana konsep *scaffolding* (Wood et al., 1976), serta umpan balik yang diarahkan pada tugas, proses, dan regulasi diri (Hattie & Timperley, 2007), memerlukan waktu lebih panjang daripada satu hari. Temuan ini menegaskan pentingnya durasi memadai dan pendampingan berkelanjutan sebagai ciri pengembangan keprofesian yang berdampak (Desimone, 2009; Darling-Hammond et al., 2017; Kennedy, 2016).

Tidak signifikannya perubahan skor keseluruhan tidak berarti kegiatan tidak bermanfaat. Pertama, skor awal peserta sudah tinggi sehingga ruang peningkatan sempit. Kedua, analisis dibatasi pada 13 butir yang setara. Ketiga, sebagian hasil justru tercermin pada peningkatan kesiapan dan penguasaan domain tertentu. Keempat, tes pilihan ganda hanya menjangkau tingkat pembelajaran dalam kerangka Kirkpatrick dan Kirkpatrick (2006) dan belum menangkap keterampilan praktis menyusun artikel, menyusun StoryMaps, mengoperasikan NVivo, atau menjalankan analisis Google Earth Engine, maupun kualitas draf yang dihasilkan peserta. Karena itu, evaluasi berikutnya perlu dilengkapi penilaian produk dan kinerja, misalnya rubrik penilaian draf sebelum dan sesudah umpan balik, serta penelusuran perubahan perilaku menulis dan publikasi dalam jangka yang lebih panjang (Guskey, 2002).

Berdasarkan temuan tersebut, tindak lanjut yang disarankan meliputi (a) siklus klinik berbasis draf yang berulang dengan satu artikel yang benar-benar diselesaikan oleh setiap peserta atau kelompok kecil, (b) pendampingan bertahap untuk NVivo dan Generative AI melalui demonstrasi hingga revisi produk, (c) penguatan konteks lokal DKI Jakarta seperti banjir, urban heat island, dan perubahan tutupan lahan sebagai sumber topik artikel, (d) evaluasi autentik berbasis rubrik dan portofolio draf, (e) pembentukan komunitas praktik MGMP sebagai kelompok menulis sejawat yang menjaga keberlanjutan setelah pelatihan (Lee & Boud, 2003; Wenger, 1998), serta (f) perbaikan instrumen agar pre-test dan post-test disusun dari kisi-kisi yang setara.

5. KESIMPULAN (Conclusions)

Draft-based writing clinic bagi guru anggota MGMP Geografi DKI Jakarta meningkatkan penguasaan peserta pada 13 butir setara dari 79,7% menjadi 82,7%, namun peningkatan tersebut belum signifikan secara statistik ($p = 0{,}062$) karena skor awal yang telah tinggi. Pengetahuan konseptual tentang penulisan ilmiah relatif stabil pada tingkat tinggi (83,6% menjadi 82,3%; $p = 0{,}591$). Peningkatan yang signifikan terjadi pada domain Google Earth Engine, yaitu dari 85,2% menjadi 92,1% ($p = 0{,}006$), dan pada

kesiapan membuat StoryMaps dengan Google Earth, yaitu dari rerata 3,36 menjadi 3,57 (p = 0,014). Kesiapan menggunakan Generative AI meningkat dari 3,36 menjadi 3,52 tetapi belum signifikan (p = 0,060). Pada akhir kegiatan, penguasaan tertinggi terdapat pada Google Earth Engine dan terendah pada NVivo sebesar 63,8%.

Dengan demikian, penguatan kapasitas penulisan ilmiah guru geografi pada tahap awal tampak paling nyata pada materi penunjang yang kontekstual dan aplikatif serta pada kesiapan menggunakan teknologi, sedangkan perkembangan kapasitas menulis itu sendiri belum dapat disimpulkan karena tes pilihan ganda tidak menangkap kualitas draf. Disarankan agar kegiatan berikutnya dirancang sebagai siklus klinik berbasis draf dengan pendampingan bertahap dan konteks kewilayahan DKI Jakarta, dilengkapi evaluasi autentik berupa rubrik produk dan portofolio, serta menggunakan instrumen pre-test dan post-test yang disusun dari kisi-kisi setara agar besaran perubahan dapat ditafsirkan lebih kuat. Pembentukan komunitas praktik antar-anggota MGMP diperlukan untuk menjaga keberlanjutan penerapan hasil pelatihan di sekolah.

6. UCAPAN TERIMA KASIH (Acknowledgement)

Penulis mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat Universitas Negeri Jakarta atas dukungan pendanaan kegiatan, serta kepada pengurus dan anggota MGMP Geografi DKI Jakarta atas partisipasi aktif selama pelaksanaan pelatihan dan pengisian instrumen evaluasi.

7. DAFTAR PUSTAKA (References)

Aitchison, C., & Lee, A. (2006). Research writing: Problems and pedagogies. Teaching in Higher Education, 11(3), 265–278.

Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., … Brisco, B. (2020). Google Earth Engine cloud computing platform for remote sensing big data applications: A comprehensive review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5326–5350.

Cargill, M., & O'Connor, P. (2013). Writing scientific research articles: Strategy and steps (2nd ed.). Chichester: Wiley-Blackwell.

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118.

Cope, M. P., Mikhailova, E. A., Post, C. J., Schlautman, M. A., & Carbajales-Dale, P. (2018). Developing and evaluating an ESRI story map as an educational tool. Natural Sciences Education, 47(1), 1–9.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development. Palo Alto, CA: Learning Policy Institute.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. Educational Researcher, 38(3), 181–199.

Dhakal, K. (2022). NVivo. Journal of the Medical Library Association, 110(2), 270–272.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., … Wright, R. (2023). Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management, 71, 102642.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition and Communication, 32(4), 365–387.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27.

Guskey, T. R. (2002). Professional development and teacher change. Teachers and Teaching, 8(3), 381–391.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112.

Hyland, K. (2016). Teaching and researching writing (3rd ed.). New York: Routledge.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. Language Teaching, 39(2), 83–101.

Kamler, B., & Thomson, P. (2014). Helping doctoral students write: Pedagogies for supervision (2nd ed.). London: Routledge.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274.

Kennedy, M. M. (2016). How does professional development improve teaching? Review of Educational Research, 86(4), 945–980.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). Evaluating training programs: The four levels (3rd ed.). San Francisco, CA: Berrett-Koehler.

Lee, A., & Boud, D. (2003). Writing groups, change and academic identity: Research development as local practice. Studies in Higher Education, 28(2), 187–200.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054.

Murray, R. (2013). Writing for academic journals (3rd ed.). Maidenhead: Open University Press.

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. Reading & Writing Quarterly, 19(2), 139–158.

Peraturan Menteri Negara Pendayagunaan Aparatur Negara dan Reformasi Birokrasi Nomor 16 Tahun 2009 tentang Jabatan Fungsional Guru dan Angka Kreditnya.

Strachan, C., & Mitchell, J. (2014). Teachers' perceptions of Esri story maps as effective teaching tools. Review of International Geographical Education Online, 4(3), 195–220.

Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., & Brisco, B. (2020). Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS Journal of Photogrammetry and Remote Sensing, 164, 152–170.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178.

Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2008). Investigating ceiling effects in longitudinal data analysis. Multivariate Behavioral Research, 43(3), 476–496.

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge: Cambridge University Press.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100.

Woods, M., Paulus, T., Atkins, D. P., & Macklin, R. (2016). Advancing qualitative research using qualitative data analysis software (QDAS)? Reviewing potential versus practice in published studies using ATLAS.ti and NVivo, 1994–2013. Social Science Computer Review, 34(5), 597–617.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(39), 1–27.