

**AI-enhanced Filipino language learning: A mixed-methods evaluation
of effectiveness using achievement tests and technology acceptance**

Ana Aubrey Porlas*

Cebu Technological University – Carmen Campus, Cebu, Philippines

ABSTRACT

This mixed-methods study examined the effectiveness of an AI-enhanced Filipino language learning platform among tertiary-level students in a state university in Cebu, Philippines. Framed within a convergent parallel mixed-methods design, the study employed a quasi-experimental pretest-posttest nonequivalent group design for the quantitative strand and thematic analysis of learner reflections and focus group discussions (FGDs) for the qualitative strand. Eighty-two (82) students participated over a 12-week semester, with 42 assigned to the AI-enhanced experimental group and 40 to a conventional instruction control group. The Technology Acceptance Model (TAM) was operationalized through a validated 9-item Likert survey administered to experimental group participants at semester end. Results revealed statistically significant post-test gains across all five assessed domains — reading comprehension, grammar and structure, vocabulary, writing proficiency, and listening skills — in favor of the experimental group (overall mean gain: +16.5 vs. +8.2 for the control; $p < .001$). Effect sizes were uniformly large (Cohen's $d = 0.73$ – 1.23). TAM survey results indicated high overall acceptance ($M = 4.21$, $SD = 0.80$; Cronbach's $\alpha = .87$), with Perceived Usefulness emerging as the strongest subscale ($M = 4.29$). Multiple regression analysis confirmed PU as the strongest significant predictor of post-test achievement ($\beta = .41$, $p < .001$), followed by Perceived Ease of Use ($\beta = .26$) and Prior Technology Exposure ($\beta = .18$). Qualitative findings yielded five themes: heightened engagement and intrinsic motivation, appreciation for personalized feedback, reduced language anxiety, connectivity and technical barriers, and concerns regarding cultural authenticity of platform content. The study contributes timely empirical evidence to the nascent literature on AI-enhanced language learning in Philippine higher education, while foregrounding equity and cultural representation as unresolved design imperatives.

Keywords: Artificial intelligence in education, Filipino language learning, technology acceptance model, mixed-methods research, achievement testing, language anxiety, Philippine higher education, adaptive learning.

Date Submitted: June 19, 2026

Date Accepted: June 25, 2026

Date Published: July 2, 2026

*Corresponding author / Email: anaaubrey.porlas@ctu.edu.ph
DOI: <http://doi.org/10.69651/PIJHSS05031265>

Recommended citation:

Porlas, A. A. (2026). AI-enhanced Filipino language learning: A mixed-methods evaluation of effectiveness using achievement tests and technology acceptance. *Pantao (International Journal of the Humanities and Social Sciences)* 5 (3), 181-191. <http://doi.org/10.69651/PIJHSS05031265>

**Pinahusay na pagkatuto ng wikang Filipino sa pamamagitan ng artificial intelligence:
Isang mixed-methods na ebalwasyon ng pagiging epektibo gamit ang
mga pagsusulit sa pagganap at pagtanggap sa teknolohiya**

Ana Aubrey Porlas*

Cebu Technological University – Carmen Campus, Cebu, Philippines

ABSTRAK

Ang pag-aaral na ito na gumamit ng pinagsamang kuwantitatibo at kuwalitatibong pamamaraan (mixed-methods) ay nagsuri sa pagiging epektibo ng isang platapormang pinahusay ng Artificial Intelligence (AI) para sa pagkatuto ng wikang Filipino sa mga mag-aaral sa antas tersiyarya sa isang pampublikong unibersidad sa Cebu, Pilipinas. Isinagawa ang pananaliksik gamit ang convergent parallel mixed-methods na disenyo, kung saan ang kuwantitatibong bahagi ay gumamit ng quasi-experimental na pretest-posttest nonequivalent group design, samantalang ang kuwalitatibong bahagi ay gumamit ng tematikong pagsusuri sa mga repleksiyon ng mga mag-aaral at mga Focus Group Discussion (FGD). Kabuuang walumpu't dalawa (82) na mag-aaral ang lumahok sa loob ng labindalawang (12) linggong semestre, kung saan apatnapu't dalawa (42) ang kabilang sa eksperimental na pangkat na gumamit ng AI-enhanced na plataporma at apatnapu (40) naman ang kabilang sa pangkat na kontrol na tumanggap ng tradisyunal na paraan ng pagtuturo. Isinakatuparan ang Technology Acceptance Model (TAM) sa pamamagitan ng isang balidado na siyam (9) na item na Likert na talatanungan na ibinigay sa mga kalahok ng eksperimental na pangkat sa pagtatapos ng semestre. Ipinakita ng mga resulta ang estadistikal na makabuluhang pagtaas ng mga marka sa post-test sa lahat ng limang sinuring larangan, kabilang ang pag-unawa sa binasa, gramatika at estruktura, talasalitaan, kasanayan sa pagsulat, at kasanayan sa pakikinig, na pawang pumabor sa eksperimental na pangkat. Ang kabuuang mean gain ng eksperimental na pangkat ay +16.5 kumpara sa +8.2 ng pangkat na kontrol ($p < .001$). Ang mga sukat ng epekto (effect sizes) ay pare-parehong mataas (Cohen's $d = 0.73-1.23$). Ipinakita rin ng resulta ng TAM survey ang mataas na antas ng pagtanggap ng mga kalahok sa paggamit ng AI-enhanced na plataporma ($M = 4.21$, $SD = 0.80$; Cronbach's $\alpha = .87$), kung saan lumitaw ang Perceived Usefulness bilang may pinakamataas na mean ($M = 4.29$). Pinatunayan ng multiple regression analysis na ang Perceived Usefulness ang may pinakamalakas at makabuluhang ambag sa pagganap sa post-test ($\beta = .41$, $p < .001$), na sinundan ng Perceived Ease of Use ($\beta = .26$) at Prior Technology Exposure ($\beta = .18$). Mula sa kuwalitatibong pagsusuri, lumitaw ang limang pangunahing tema: mas mataas na antas ng pakikilahok at likas na motibasyon sa pagkatuto, pagpapahalaga sa personalisadong pagbibigay ng puna o feedback, pagbawas sa pagkabalisa sa paggamit ng wika, mga hamong dulot ng konektibidad at mga suliraning teknikal, at mga pangamba hinggil sa pagiging autentiko ng kulturang kinakatawan sa nilalaman ng plataporma. Nag-aambag ang pag-aaral na ito ng napapanahon at empirikal na ebidensiya sa umuusbong na literatura tungkol sa AI-enhanced na pagkatuto ng wika sa mataas na edukasyon sa Pilipinas, habang binibigyang-diin ang mga usapin ng pagkakapantay-pantay at wastong representasyong pangkultura bilang mahahalagang aspektong dapat pang higit na mapaunlad sa pagdidisenyo ng mga teknolohiyang pang-edukasyon.

Susing salita: Artipisyal na intelihensiya, pagkatuto ng wikang Filipino, modelo ng pagtanggap sa teknolohiya, mixed-methods na pananaliksik, pagsusulit sa pagganap, pagkabalisa sa wika, mataas na edukasyon sa Pilipinas, adaptibong pagkatuto.

PANIMULA

Ang integrasyon ng artipisyal na katalinuhan sa pormal na pagtuturo ng wika ay nakaakit ng walang kapantay na atensyon ng mga iskolar at institusyon, ngunit ang mga aplikasyon nito ay nananatiling nakatuon sa Ingles at iba pang mga wikang pandaigdigang nangingibabaw (Godwin-Jones, 2022; Lyu et al., 2025; Wang et al., 2024). Ang Filipino — bilang konstitusyonal na itinalagang pambansang wika at midyum ng pagtuturo sa mga paaralang Pilipino — ay nasa mapanghimok na posisyon: ito ay sabay-sabay na isang wika ng pambansang pagkakakilanlan at, para sa milyun-milyong mag-aaral mula sa Visayas, Mindanao, at iba pang mga rehiyon ng lingguwistikong minoridad, isang natutunang ikalawang wika na ang panlipunang mga stake ay puno ng mga tanong ng rehiyunal na pagkabilang at kultural na kapital (Tupas, 2015; Ulla et al., 2024). Sa kabila ng mga kumplikasyon na ito, ang empirikong pananaliksik sa teknolohiyang pinahusay na pag-aaral ng Filipino sa antas tersyaryo ay nananatiling manipis, at ang mga imbestigasyon na kinasasangkutan ng mga AI-based na plataporma ay halos wala sa peer-reviewed na literatura (Dela Cruz et al., 2025).

Sa pandaigdigang antas, ang base ng katibayan para sa AI-assisted na pagkuha ng ikalawang wika (SLA) ay lumaki nang malaki. Ang mga meta-analysis ni Lyu et al. (2025), Wang et al. (2024), at Li et al. (2025) ay magkasama na nag-uulat ng katamtaman hanggang malaking positibong epekto para sa chatbot-assisted at AI-mediated na pag-aaral ng wika sa iba't ibang konteksto ng L2, na may partikular na lakas na dokumentado sa pagkuha ng bokabularyo, kasanayan sa pagsasalita, at feedback sa pagsulat. Ang mga parallel na imbestigasyon na nag-aaplay ng Technology Acceptance Model (TAM; Davis, 1989) sa mga AI-pinahusay na setting ng edukasyon ay palaging nakakahanap na ang Nakitang Pagiging Kapaki-pakinabang at Nakitang Kaghinhawahan sa Paggamit ay mga makabuluhang tagahula ng parehong hangarin sa pag-adopt at, lalong lumalaki, aktwal na mga resulta ng pagkatuto (Dahri et al., 2024; Yim, 2024). Gayunpaman, karamihan sa mga pag-aaral na ito ay nakaangkla sa mga konteksto ng EFL sa Silangang Asya o Kanlurang Europa; ang kanilang generalizability sa mga underresourced na lingguwistikong kapaligiran ng Timog-Silangang Asya ay hindi tiyak.

Ang Pilipinas ay nag-aalok ng partikular na instructive na case study sa ilang dahilan. Una, ang sektor ng mas mataas na edukasyon ng bansa ay sumasailalim sa pinabilis na digital na pagbabago. Pangalawa, ang Filipino bilang paksa sa kolehiyo ay may dalawahang pedagogikal na layunin: upang mapabuti ang kakayahang pangkomunikatibo sa pambansang wika at upang mapaunlad ang kritikal na pakikipag-ugnayan sa kultural na pamana ng Pilipinas. Pangatlo, at kritikal, ang digital divide ay nananatiling matindi: isang survey ng industriya noong 2023 ay natuklasan na 61% ng mga guro sa Pilipinas ay natukoy ang mga gastos sa pag-access ng teknolohiya bilang pangunahing hadlang (Instructure, 2023).

Paglalahad ng suliranin

Sinasagot ng pag-aaral na ito ang mga kondisyon na ito sa pamamagitan ng pag-ebalwasyon ng isang lokal na binuo na plataporma para sa pag-aaral ng Filipino na pinahusay ng AI. Tatlong mga tanong sa pananaliksik ang gumagabay sa inkwiri:

1. Sa anong lawak ang mga mag-aaral sa pangkat na pinahusay ng AI ay nakamit ang mas mataas na iskor sa post-test kumpara sa pangkat ng kontrol sa limang larangan ng Filipino?
2. Paano naiintindihan ng mga mag-aaral ng pangkat ng AI ang plataporma sa mga tuntunin ng mga konstrukto ng TAM?

3. Anong mga karanasang qualitative — motibasyon, emosyonal, at istrukturan — ang iniuulat ng mga mag-aaral sa panahon at pagkatapos ng interbensyon?

METODOLOHIYA

Gumamit ang pag-aaral na ito ng convergent parallel mixed-methods na disenyo (Creswell at Plano Clark, 2018), kung saan ang mga quantitative at qualitative na datos ay nakolekta nang sabay-sabay, sinuri nang hiwalay, at isinama sa panahon ng interpretasyon. Gumamit ang quantitative strand ng quasi-experimental na pretest-posttest na disenyo ng hindi katumbas na pangkat na may buo na mga seksyon ng klase, samantalang gumamit ang qualitative strand ng mga nakasulat na journal ng repleksyon na nakolekta sa mga linggong 4, 8, at 12, at isang 90-minutong FGD na isinagawa sa katapusan ng semestre na may purposively selected na sub-sample ($n = 12$) mula sa pangkat na pang-eksperimento.

Ang mga kalahok ay 82 na mag-aaral ng kolehiyo mula sa dalawang buo na seksyon ng Filipino 101 (Komunikasyon sa Akademikong Filipino) na naka-enroll sa isang unibersidad ng estado sa Cebu sa ikalawang semestre ng Taong Akademiko 2023–2024. Ang Seksyon A ($n = 42$) ay random na pinili bilang pangkat na pang-eksperimento; ang Seksyon B ($n = 40$) ay nagsilbing pangkat ng kontrol. Isang power analysis na isinagawa bago ang koleksyon ng datos ($\alpha = .05$, lakas = .80, katamtamang laki ng epekto $d = 0.50$) ay nagtantiya ng minimum na sample na 34 bawat pangkat; ang nakamit na sample na 42 at 40 ay nalampasan ang threshold na ito. Sa larangan ng kasarian, binubuo ang Pangkat AI ($n=42$) ng 18 lalaki (42.9%) at 24 babae (57.1%), habang binubuo ang Pangkat Kontrol ($n=40$) ng 17 lalaki (42.5%) at 23 babae (57.5%); sa kabuuan ($N=82$), 35 (42.7%) ang lalaki at 47 (57.3%) ang babae. Sa antas ng pag-aaral, sa Pangkat AI ay may 11 (26.2%) sa Unang Taon, 15 (35.7%) sa Ikalawang Taon, at 16 (38.1%) sa Ikatlong Taon; sa Pangkat Kontrol ay may 10 (25.0%) sa Unang Taon, 14 (35.0%) sa Ikalawang Taon, at 16 (40.0%) sa Ikatlong Taon; sa kabuuan, 21 (25.6%) ang nasa Unang Taon, 29 (35.4%) ang nasa Ikalawang Taon, at 32 (39.0%) ang nasa Ikatlong Taon. Hinggil sa naunang karanasan sa teknolohiya, ang Pangkat AI ay binubuo ng 19 (45.2%) na may mataas na karanasan, 16 (38.1%) na may katamtamang karanasan, at 7 (16.7%) na may mababang karanasan; ang Pangkat Kontrol naman ay binubuo ng 18 (45.0%) na may mataas, 15 (37.5%) na may katamtaman, at 7 (17.5%) na may mababang karanasan; sa kabuuan, 37 (45.1%) ang may mataas, 31 (37.8%) ang may katamtaman, at 14 (17.1%) ang may mababang naunang karanasan sa teknolohiya. Ang paunang antas ng pagsubok (SD) ay 59.3 (8.8) para sa Pangkat AI, 59.7 (8.6) para sa Pangkat Kontrol, at 59.5 (8.7) sa kabuuan. Mahalagang banggitin na ang mga seksyon, at hindi ang mga indibidwal na mag-aaral, ang random na itinalaga sa mga kondisyon ng pang-eksperimento at kontrol, at ang paunang antas ng pagsubok ay sumasalamín sa hindi weighted na average ng limang pre-test na iskor sa larangan sa sukat na 100 punto.

Ang AI-pinahusay na plataporma na ginamit sa pag-aaral na ito ay isang lokal na binuo na web application na nag-iintegrate ng tatlong pangunahing functional na module: isang adaptive na module ng pagbabasa at pag-unawa sa Filipino na may NLP-based na pagsusuri; isang conversational na module ng pagsasanay sa Filipino gamit ang large language model (LLM) API para sa real-time na interaksyon at corrective feedback; at isang automated na module ng pagsusuri sa pagsulat na nagbibigay ng domain-specific na feedback batay sa rubric sa gramatika, koherensya, at hanay ng leksikal. Ang pagpapaunlad ng plataporma ay sumusunod sa mga prinsipyo ng participatory na disenyo, na may Content Validity Index na .91. Ang mga mag-aaral sa pangkat na pang-eksperimento ay inaasahang mag-log ng minimum na tatlong one-hour na session bawat linggo, samantalang sinunod ng pangkat ng kontrol ang karaniwang syllabus ng Filipino 101. Ang panahon ng pag-aaral ay sumasaklaw sa 12 linggo.

Hinggil sa mga instrumento ng koleksyon ng datos, ang researcher-developed na Filipino Language Achievement Test (FLAT) ay ibinigay sa parehong pangkat bilang pre-test

at post-test. Sinasaklaw ng FLAT ang limang larangan: pag-unawa sa binabasa (20 aytem), gramatika at estruktura ng pangungusap (20 aytem), bokabularyo (20 aytem), kakayahan sa pagsulat (nakaisgored na rubric, 20 puntos), at mga kasanayang pakikinig (20 aytem), para sa kabuuang posibleng iskor na 100. Ang instrumento ay napatunayan ng apat na guro ng Filipino (CVI = .91) at sinubukan sa 30 mag-aaral, na nagbibigay ng Cronbach's α na .86. Bukod dito, ang pagtanggap ng teknolohiya ay sinuri gamit ang isang 9-aytem na adapted na survey ng TAM batay sa seminal na instrumento ni Davis (1989), na inayos sa tatlong subscale: Nakitang Pagiging Kapaki-pakinabang (3 aytem), Nakitang Kaginhawahan sa Paggamit (3 aytem), at Hangarin sa Paggamit (3 aytem). Ang mga tugon ay nairehistro sa isang 5-puntong sukat ng Likert (1 = Lubos na Hindi Sumasang-ayon, 5 = Lubos na Sumasang-ayon), at ipinakita ng adapted na survey ang katanggap-tanggap na internal na konsistensya (kabuuang Cronbach's α = .87). Ang survey na ito ay ibinigay lamang sa pangkat na pang-eksperimento sa katapusan ng linggong 12. Para naman sa mga qualitative na instrumento, ang mga datos ay nakolekta mula sa dalawang pinagkukunan: mga istrukturadong journal ng repleksyon na isinumite ng lahat ng 42 mag-aaral ng pangkat na pang-eksperimento sa mga linggong 4, 8, at 12, gamit ang mga open-ended na prompt na nagta-target ng mga karanasan sa pagkatuto, mga motibasyon, at mga hamon; at isang 90-minutong FGD na audio-recorded at transcribed verbatim, na isinagawa sa Filipino at Ingles na may 12 purposively selected na kalahok.

Sa pagsusuri ng datos, ang mga datos na quantitative ay sinuri gamit ang IBM SPSS Statistics Version 28. Ang pagkakapantay-pantay ng pre-test ng pangkat ay kinumpirma sa pamamagitan ng mga independent samples t-test sa lahat ng limang larangan (lahat ng $p > .05$). Ang mga pagkakaiba sa post-test sa pagitan ng mga pangkat ay sinuri gamit ang mga independent samples t-test na may mga laki ng epekto na kinalkula bilang Cohen's d , samantalang ang multiple linear regression ay ginamit upang matukoy ang mga makabuluhang tagahula ng kabuuang iskor sa post-test. Ang mga datos na qualitative naman ay sinuri nang tematikal ayon sa anim na yugto ng pamamaraan ni Braun at Clarke (2006), kung saan ang dalawang mananaliksik ay nagsagawa ng independyenteng pag-code sa datos, na nakamit ang katanggap-tanggap na inter-rater reliability (Cohen's κ = .78).

RESULTA AT PAGATATALAKAY

Ang seksyong ito ay naglalahad ng mga resulta at talakayan ng pag-aaral na isinagawa kasama ang 82 mag-aaral ng kolehiyo mula sa dalawang buo na seksyon ng Filipino 101, na hinati sa pangkat na pang-eksperimento ($n = 42$) at pangkat ng kontrol ($n = 40$), sa ilalim ng isang convergent parallel mixed-methods na disenyo na gumamit ng quasi-experimental na pretest-posttest na disenyo ng hindi katumbas na pangkat para sa quantitative strand at mga journal ng repleksyon at FGD para sa qualitative strand. Ang Filipino Language Achievement Test (FLAT) ang ginamit na pangunahing instrumento sa pagsukat ng tagumpay sa pag-aaral sa limang larangan, samantalang ang adapted na Technology Acceptance Model (TAM) survey ang sumukat sa pagtanggap ng teknolohiya ng mga mag-aaral sa pangkat na pang-eksperimento. Ang mga datos na quantitative ay sinuri gamit ang independent samples t-test, Cohen's d para sa laki ng epekto, Pearson correlation, at multiple linear regression sa pamamagitan ng IBM SPSS Statistics Version 28, samantalang ang mga datos na qualitative ay sinuri nang tematikal ayon sa anim na yugto ng pamamaraan ni Braun at Clarke (2006). Ang talakayan na sumusunod ay nakabatay nang buo sa mga datos na nakolekta at ang mga resulta ay direktang inihahanay sa mga layunin ng pag-aaral upang masuri ang epekto ng AI-pinahusay na pagtuturo sa tagumpay sa wikang Filipino.

Kinumpirma ng mga independent samples t-test na walang makabuluhang pagkakaiba sa estadistika sa pagitan ng pang-eksperimento at pangkat ng kontrol sa alinman sa limang larangan ng pre-test (lahat ng p na halaga $> .05$), na binabalidahan ang pagkakapantay-pantay ng baseline sa pagitan ng dalawang pangkat bago ang interbensyon. Sa pag-unawa sa binabasa, ang Pangkat AI ay nagtala ng pre-test na $M = 61.4$ ($SD = 8.2$) at post-test na $M = 78.9$ ($SD = 7.1$), kumpara sa Pangkat Kontrol na $M = 62.1$ ($SD = 7.9$) sa pre-test at $M = 70.3$ ($SD = 8.4$) sa post-test, na nagresulta sa pagtaas na $+17.5$ para sa Pangkat AI at $+8.2$ para sa Pangkat Kontrol ($p = .003^{**}$). Sa gramatika at estruktura ng pangungusap, ang Pangkat AI ay umangat mula sa $M = 57.8$ ($SD = 9.4$) tungong $M = 74.2$ ($SD = 8.6$), samantalang ang Pangkat Kontrol ay umangat mula sa $M = 58.3$ ($SD = 8.7$) tungong $M = 66.8$ ($SD = 9.1$), na may pagtaas na $+16.4$ kumpara sa $+8.5$ ($p = .001^{**}$). Sa bokabularyo, ang Pangkat AI ay umangat mula sa $M = 63.2$ ($SD = 7.6$) tungong $M = 80.1$ ($SD = 6.8$), samantalang ang Pangkat Kontrol ay umangat mula sa $M = 62.7$ ($SD = 8.1$) tungong $M = 71.4$ ($SD = 7.3$), na may pagtaas na $+16.9$ kumpara sa $+8.7$ ($p = .002^{**}$). Sa kakayahan sa pagsulat, ang Pangkat AI ay umangat mula sa $M = 54.6$ ($SD = 10.1$) tungong $M = 69.3$ ($SD = 9.2$), samantalang ang Pangkat Kontrol ay umangat mula sa $M = 55.2$ ($SD = 9.8$) tungong $M = 62.1$ ($SD = 10.4$), na may pagtaas na $+14.7$ kumpara sa $+6.9$ ($p = .018^{*}$). Sa kasanayang pakikinig, ang Pangkat AI ay umangat mula sa $M = 59.3$ ($SD = 8.9$) tungong $M = 76.6$ ($SD = 7.4$), samantalang ang Pangkat Kontrol ay umangat mula sa $M = 60.1$ ($SD = 8.3$) tungong $M = 68.9$ ($SD = 8.8$), na may pagtaas na $+17.3$ kumpara sa $+8.8$ ($p = .004^{**}$). Sa kabuuang iskor, ang Pangkat AI ay umangat mula sa $M = 59.3$ ($SD = 8.8$) tungong $M = 75.8$ ($SD = 7.8$), samantalang ang Pangkat Kontrol ay umangat mula sa $M = 59.7$ ($SD = 8.6$) tungong $M = 67.9$ ($SD = 8.8$), na may pagtaas na $+16.5$ kumpara sa $+8.2$ ($p = .001^{**}$). Ipinapahiwatig ng mga resultang ito na ipinakita ng pangkat na pang-eksperimento ang mga makabuluhang kalamangan sa estadistika kumpara sa pangkat ng kontrol sa lahat ng limang larangan, kung saan ang pinakamalaking ganap na pagkakaiba sa mean ay naobserbahan sa pag-unawa sa binabasa (AI post-test $M = 78.9$ kumpara sa kontrol $M = 70.3$) at mga kasanayang pakikinig ($M = 76.6$ kumpara sa $M = 68.9$).

Upang masukat ang magnitude ng mga pagkakaibang ito, kinalkula ang mga laki ng epekto bilang Cohen's d para sa bawat larangan at para sa kabuuang mean na iskor. Sa pag-unawa sa binabasa, ang pagkakaiba ng mean na 8.60 at pinagsamang SD na 7.75 ay nagbunga ng $d = 1.11$, na minarkahang malaki. Sa gramatika at estruktura, ang pagkakaiba ng mean na 7.40 at pinagsamang SD na 8.85 ay nagbunga ng $d = 0.84$, na minarkahang malaki. Sa bokabularyo, ang pagkakaiba ng mean na 8.70 at pinagsamang SD na 7.05 ay nagbunga ng $d = 1.23$, ang pinakamalaking laki ng epekto sa buong pag-aaral at minarkahang malaki. Sa kakayahan sa pagsulat, ang pagkakaiba ng mean na 7.20 at pinagsamang SD na 9.80 ay nagbunga ng $d = 0.73$, ang pinakamaliit sa lahat ng larangan at minarkahang katamtaman-malaki. Sa kasanayang pakikinig, ang pagkakaiba ng mean na 7.70 at pinagsamang SD na 8.10 ay nagbunga ng $d = 0.95$, na minarkahang malaki. Para sa kabuuang iskor, ang pagkakaiba ng mean na 7.90 at pinagsamang SD na 8.30 ay nagbunga ng $d = 0.95$, na minarkahang malaki. Gamit ang interpretasyon ni Cohen (1988), kung saan ang maliit ay 0.20 – 0.49 , ang katamtaman ay 0.50 – 0.79 , at ang malaki ay ≥ 0.80 , lahat ng larangan ay nagbunga ng malaking mga laki ng epekto ($d \geq 0.73$), at ang mga magnitude ng epekto na ito ay lumagpas sa kabuuang katamtamang epekto ($d = 0.608$) na iniulat ni Lyu et al. (2025).

Nakolekta ang datos ng TAM survey mula sa lahat ng 42 kalahok sa pangkat na pang-eksperimento sa katapusan ng semestre. Sa antas ng aytem, ang konstruktong Nakitang Pagiging Kapaki-pakinabang (NK), na may subscale Cronbach's $\alpha = .83$, ay nagtala ng kabuuang subscale na $M = 4.29$ ($SD = .78$). Sa loob nito, ang aytem na "nakatutulong sa mas mabilis na pagkatuto ng Filipino" ay tinugunan ng 52.4% SN, 35.7% SA, 7.1% N, 4.8% HD, at 0.0% SHD, na may $M = 4.31$ ($SD = .83$); ang aytem na "ginagawang mas may kaugnayan ang pagkatuto" ay tinugunan ng 45.2% SN, 40.5% SA, 9.5% N, 4.8% HD, at 0.0% SHD, na

may $M = 4.26$ ($SD = .81$); at ang aytem na "nagpapabuti ng aking mga resulta sa pagkatuto" ay tinugunan ng 47.6% SN, 38.1% SA, 9.5% N, 4.8% HD, at 0.0% SHD, na may $M = 4.29$ ($SD = .84$). Ang konstruktong Nakitang Kaginhawahan sa Paggamit (KP), na may subscale Cronbach's $\alpha = .79$, ay nagtala ng kabuuang subscale na $M = 4.14$ ($SD = .79$). Ang aytem na "madaling mag-navigate sa platform" ay tinugunan ng 40.5% SN, 42.9% SA, 11.9% N, 4.8% HD, at 0.0% SHD, na may $M = 4.19$ ($SD = .80$); ang aytem na "malinaw ang mga tagubilin" ay tinugunan ng 38.1% SN, 45.2% SA, 11.9% N, 4.8% HD, at 0.0% SHD, na may $M = 4.17$ ($SD = .81$); at ang aytem na "hindi nangangailangan ng labis na pagsisikap" ay tinugunan ng 35.7% SN, 42.9% SA, 14.3% N, 7.1% HD, at 0.0% SHD, na may $M = 4.07$ ($SD = .90$). Ang konstruktong Hangarin sa Paggamit (HP), na may subscale Cronbach's $\alpha = .81$, ay nagtala ng kabuuang subscale na $M = 4.19$ ($SD = .84$). Ang aytem na "plano kong gamitin ito sa susunod na semestre" ay tinugunan ng 42.9% SN, 38.1% SA, 11.9% N, 7.1% HD, at 0.0% SHD, na may $M = 4.17$ ($SD = .89$); at ang aytem na "irerekomenda ko ito sa aking mga kaklase" ay tinugunan ng 45.2% SN, 35.7% SA, 14.3% N, 4.8% HD, at 0.0% SHD, na may $M = 4.21$ ($SD = .85$). Ang kabuuang iskor ng TAM ay $M = 4.21$ ($SD = .80$), na may kabuuang Cronbach's $\alpha = .87$, na nagpapahiwatig ng lubos na positibong saloobin ng mga mag-aaral sa pagtanggap ng teknolohiya, kung saan ang NK ang nagrehistro ng pinakamataas na iskor ng subscale ($M = 4.29$). Ang pagsusuri ng korelasyon ng Pearson ay nagpakita ng makabuluhang positibong mga relasyon sa pagitan ng mga konstrukto ng TAM at mga iskor sa post-test: NK at mga iskor sa post-test ($r = .58$, $p < .001$), KP at mga iskor sa post-test ($r = .42$, $p < .01$), at HP at mga iskor sa post-test ($r = .46$, $p < .01$).

Upang matukoy ang mga makabuluhang tagahula ng tagumpay sa post-test, isinagawa ang multiple linear regression na may kabuuang iskor sa post-test bilang variable ng kinalabasan at limang tagahula na sabay-sabay na ipinasok: NK, KP, Naunang Karanasan sa Teknolohiya, Antas ng Pag-aaral, at Kasarian. Ang kabuuang modelo ay makabuluhan sa estadistika, $F(5, 36) = 11.37$, $p < .001$, na nagpapaliwanag ng 61.2% ng pagbabago sa mga iskor sa post-test ($R^2 = .612$, Adjusted $R^2 = .587$, Durbin-Watson = 1.94). Ang konstante ay $B = 24.31$ ($SE = 4.12$, $t = 5.90$, $p < .001$, 95% CI [16.10, 32.51]). Ang Nakitang Pagiging Kapakipakinabang ang naging pinakamalakas na makabuluhang tagahula, na may $B = 5.84$ ($SE = 1.23$, $\beta = .41$, $t = 4.75$, $p < .001^{**}$, 95% CI [3.39, 8.29], VIF = 1.43). Sumunod ang Nakitang Kaginhawahan sa Paggamit, na may $B = 3.47$ ($SE = 1.18$, $\beta = .26$, $t = 2.94$, $p = .005^{**}$, 95% CI [1.13, 5.81], VIF = 1.38), at Naunang Karanasan sa Teknolohiya, na may $B = 2.19$ ($SE = 0.94$, $\beta = .18$, $t = 2.33$, $p = .023^{*}$, 95% CI [0.32, 4.06], VIF = 1.12). Ang Antas ng Pag-aaral ($B = 1.03$, $SE = 0.88$, $\beta = .09$, $t = 1.17$, $p = .246$, 95% CI [-0.72, 2.78], VIF = 1.09) at Kasarian ($B = 0.64$, $SE = 1.02$, $\beta = .05$, $t = 0.63$, $p = .532$, 95% CI [-1.38, 2.66], VIF = 1.04) ay hindi makabuluhang hinulaan ang tagumpay. Lahat ng halaga ng VIF ay mas mababa sa 5, na nagpapahiwatig ng kawalan ng multicollinearity sa modelo.

Ang tematikal na pagsusuri ng mga journal ng repleksyon at mga transcript ng FGD ay naglabas ng limang pangunahing tema. Ang nangingibabaw na tema, Pakikipag-ugnayan at Motibasyon (T1, $n = 38$, 90.5%), ay kumakatawan sa isang qualitative na pagbabago sa affective na pakikipag-ugnayan ng mga mag-aaral sa Filipino, na may kalahok na sumulat: "Hindi ko inasahan na magiging masaya ang pag-aaral ng Filipino. Naghahanap na ko ng dagdag na leksyon pagkatapos ng klase." Ang Halaga ng Personalisadong Feedback (T2, $n = 35$, 83.3%) ay tumutukoy sa agarang, individualized na feedback bilang isang natatangi at pinahahalagahang katangian, ayon sa kalahok na nagsabing: "Gusto ko na kita-kita agad ang aking mga pagkakamali. Hindi ka mahihiyang magkamali kasi ikaw lang at ang AI ang nakakaalam." Ang Pagbaba ng Pagkabalisa sa Wika (T3, $n = 33$, 78.6%) ay tahasang tumutukoy sa nabawasang pagkabalisa, partikular sa mga mag-aaral mula sa Visayas, gaya ng

inilarawan ng kalahok: "Mas comfortable akong mag-practice sa AI kaysa sa harap ng klase. Hindi ko natatakot na mapagalitan." Ang Mga Hadlang sa Koneksyon at Teknolohiya (T4, $n = 18$, 42.9%) ay nagtaas ng mga alalahanin tungkol sa koneksyon ng internet, ayon sa kalahok: "Minsan mabagal ang internet sa aming lugar kaya hindi ko magamit ang platform. Nakaka-frustrate kung kailangan mong gamitin ito sa bahay." Ang Mga Alalahanin sa Katutubong Kultura (T5, $n = 14$, 33.3%) ay nagpahayag ng alalahanin na ang nilalaman ng plataporma ay nagbibigay ng kagustuhan sa Filipino na nakasentro sa Maynila, ayon sa kalahok: "Ang mga halimbawa sa platform ay parang mas para sa mga taga-Maynila. Sana may mga salitang Bisaya-Filipino din para mas relate kami."

Pag-interpret ng mga Pagtaas sa Tagumpay. Ang makabuluhang kalamangan sa estadistika sa post-test ng pangkat na pinahusay ng AI sa lahat ng limang larangan, na may malaking mga laki ng epekto mula $d = 0.73$ hanggang $d = 1.23$, ay kumakatawan sa isang nakakaakit na quantitative na natuklasan. Ang mga laki ng epekto na ito ay lumagpas sa kabuuang katamtamang epekto ($ES = 0.608$) na iniulat ng meta-analysis ni Lyu et al. (2025) at naaayon sa mga single-study na imbestigasyon. Ang comparatively na mas maliit na pagtaas ng differential sa kakayahan sa pagsulat ($d = 0.73$) ay sumasalamain sa cognitive complexity ng pagsulat bilang isang kasanayan na hindi gaanong agad na sumasagot sa interbensyon kumpara sa mga larangan tulad ng bokabularyo ($d = 1.23$). Sa kritikal na paraan, ang sariling makabuluhang pagtaas ng pangkat ng kontrol (+8.2 puntos sa kabuuan) ay nagpapakita na ang karaniwan ng pagtuturo ng Filipino ay nananatiling may tunay na pedagogikal na halaga. Dahil dito, ang mga natuklasan ay pinakamainam na mainterpreta bilang additive na epekto ng AI-mediated na pagsasanay sa ibabaw ng harapang pagtuturo, hindi bilang kapalit nito.

Mga Natuklasan ng TAM at Kanilang Predictive na Kahalagahan. Ang kabuuang iskor ng TAM na $M = 4.21$ at ang iskor ng subscale ng NK na $M = 4.29$ ay nagpapahiwatig ng lubos na pabor na pagtanggap ng teknolohiya. Ang natuklasan ng regression na ang NK ($\beta = .41$) ay napalampas ang KP ($\beta = .26$) bilang tagahula ng tagumpay sa post-test ay teoretikal na makabuluhan. Nagmumungkahi ito na ang nagtutulak sa mga resulta ng pagkatuto ay hindi lamang kaginhawahan sa interface kundi ang substantibong paniniwala ng mga mag-aaral na ang kasangkapan ay may halaga sa edukasyon. Ito ay naaayon sa expectancy-value theory at naaayon sa natuklasan ni Dahri et al. (2024) na ang NK ay makabuluhang nagsisilbing tagapamagitan ng relasyon sa pagitan ng paggamit ng AI at ng mga resulta ng metacognitive na pagkatuto.

Equity, Imprastruktura, at Kultural na Representasyon. Marahil ang pinaka-consequential na mga natuklasan para sa patakaran sa edukasyon ng Pilipinas ay hindi ang mga pagtaas sa tagumpay kundi ang mga qualitative na tema ng mga teknikal na hadlang at mga alalahanin sa kultural na awtentikasyon. Ang 42.9% na dalas ng hadlang sa koneksyon, na kumakatawan sa halos kalahati ng mga kalahok ng pangkat na pang-eksperimento, ay isang nakakalungkot na paalala na ang pagiging epektibo ng AI-pinahusay na pagtuturo ay condition-contingent. Ang alalahanin sa kultural na awtentikasyon na itinanghal ng 33.3% ng mga kalahok mula sa Visayas ay nagtatanging isang struktural na limitasyon. Kapag ang mga training corpora at kultural na senaryo ay nag-default sa Filipino na nakasentro sa Maynila, ang mga regional na mag-aaral ay napoposisyon bilang mga mag-aaral ng isang wika na hindi ganap na nabibilang sa kanila (Tupas, 2015).

Mga Limitasyon. Ang pag-aaral na ito ay may ilang mga limitasyon na kailangang itampok. Una, ang quasi-experimental na disenyo ay pumipigil sa malakas na causal na inferensya, sapagkat ang mga hindi nasukat na epekto ng guro, mga selection biases sa antas ng seksyon, at mga Hawthorne effect ay hindi maaaring ganap na kontrolin. Pangalawa, ang 12-linggong timeframe ay hindi nagpapahintulot ng mga konklusyon tungkol sa pangmatagalang pananatili. Pangatlo, ang qualitative strand ay isinagawa lamang sa mga kalahok ng pangkat na pang-eksperimento. Pang-apat, ang pag-aaral ay isinagawa sa isang

institusyon ng Cebu, na nagpipigil sa generalizability sa mga probinsyal at rural na mga setting. Panglima, ang lokal na binuo na prototype na plataporma ay naiiba sa mga commercially available na kasangkapan.

Sa kabuuan, ang mga natuklasan ng pag-aaral na ito, na nakatuon sa 82 mag-aaral sa pamamagitan ng quantitative at qualitative na mga lente, ay nagpapatunay na ang AI-pinahusay na pagtuturo ay nagbunga ng makabuluhang at malalaking pagtaas sa tagumpay sa wikang Filipino sa lahat ng limang larangan, na sinusupportahan ng malaking mga laki ng epekto ($d = 0.73$ hanggang $d = 1.23$) at lubos na positibong pagtanggap ng teknolohiya (Kabuuang TAM $M = 4.21$, $SD = .80$), kung saan ang Nakitang Pagiging Kapaki-pakinabang ang pinakamalakas na predictor ng tagumpay ($\beta = .41$, $p < .001$). Ang mga resultang ito ay direktang tumutugon sa layunin ng pag-aaral na suriin ang epekto ng AI-mediated na interbensyon sa tagumpay sa pagkatuto, at nagbibigay ng karagdagang empirikal na suporta sa larangan ng AI-assisted na pagtuturo ng wika sa konteksto ng Pilipinas, partikular sa pagpapatunay ng mahalagang papel ng nakitang pagiging kapaki-pakinabang bilang tagapamagitan ng resulta ng pagkatuto. Gayunpaman, ang mga qualitative na natuklasan tungkol sa mga hadlang sa koneksyon (42.9%) at mga alalahanin sa kultural na representasyon (33.3%) ay nagpapahiwatig na ang implementasyon ng naturang teknolohiya ay hindi neutral at nangangailangan ng pansin sa equity ng imprastruktura at kultural na pagiging angkop. Ang mga implikasyon at rekomendasyong ito, kasama ang mga limitasyon ng disenyo ng pag-aaral, ay binibigyang-diin sa sumusunod na kabanata bilang batayan para sa mga konklusyon at mungkahi para sa hinaharap na pananaliksik at praktika.

KONKLUSYON

Nagbibigay ang pag-aaral na ito ng methodologically multifaceted na larawan ng pag-aaral ng Filipino na pinahusay ng AI na sabay-sabay na nagpapatibay ng potensyal ng pagtuturo ng teknolohiya at matapat na tungkol sa mga kondisyon na naghihigpit dito. Ang malalaking laki ng epekto na dokumentado sa lahat ng limang larangan ng tagumpay ay nagtataglay ng makabuluhang katibayan na ang mga lokal na binuo na AI na plataporma ay maaaring magdagdag ng halaga sa pagtuturo ng Filipino. Laban sa pagtanggap na pagbabasa na ito, ang pag-aaral ay nagpipilit sa materiyalidad ng imprastrakturang hindi pagkakapantay-pantay at kahalagahan ng kultural na representasyon.

Para sa mga institusyon ng mas mataas na edukasyon ng Pilipinas at mga developer ng plataporma, inirerekomenda ng pag-aaral na ito ang: (a) mandatory na offline o mababang-bandwidth na mga mode ng functionality para sa lahat ng AI na kasangkapan sa pagkatuto ng wika; (b) sistematikong rehiyunal na lingguwistikong pag-audit ng nilalaman ng plataporma, na may mga proseso ng co-design na kinasasangkutan ng mga mag-aaral mula sa Visayas at Mindanao; (c) mga programang paunlad para sa guro na nagpoposisyon sa mga tagapagturo bilang mga pedagogically active na interpreter ng AI-generated na datos ng feedback; at (d) mga longitudinal na protocol ng ebalwasyon na sumusubaybay ng pananatili at paglipat sa loob ng kahit isang taong akademiko.

TALASANGGUNIAN

Behforouz, B., & Al Ghaithi, A. (2024). Artificial intelligence in second and foreign language learning. *Science & Education*, 33, 1033–1063. <https://doi.org/10.1007/s11191-022-00399-z>

- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), 8427. <https://doi.org/10.3390/app12178427>
- Belda-Medina, J., & Kokoška, V. (2023). Integrating artificial intelligence (AI) in language classes: Teachers' and students' perspectives. *Education Sciences*, 13(6), 549. <https://doi.org/10.3390/educsci13060549>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL*, 16(2), 113–131.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
- Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., & Aldraiweesh, A. (2024). Extended TAM-based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education. *Heliyon*, 10(8), e29317. <https://doi.org/10.1016/j.heliyon.2024.e29317>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Dela Cruz, L., Aguinaldo, I., Alzate, L. J., Camero, C., Abiado, K. R., & Gumpal, B. (2025). Fostering intercultural competence through culturally responsive pedagogy: Practices and perspectives in Philippine higher education institutions. *International Journal on Culture, History, and Religion*, 7(SI2), 36–50.
- Gaddi, J. A. G. (2024). Learning gap in Filipino sa Piling Larangan (Teknikal-Bokasyunal): An assessment. *International Research Journal of Science, Technology, Education, and Management*, 4(2), 88–95.
- Godwin-Jones, R. (2022). Scaling up and zooming in: Big data and personalization in language learning. *Language Learning & Technology*, 26(1), 1–15.
- Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, 50(5), 2572–2593.
- Instructure. (2023). *State of student success and engagement in higher education: Philippines country findings*. Instructure Inc.
- Jeon, J. (2022). Exploring a self-directed interactive app for informal EFL learning: A self-determination theory perspective. *Education and Information Technologies*, 27(4), 5767–5787.
- Koč, F. Ş. (2024). The development of listening and speaking skills in EFL via an artificially intelligent chatbot application. *Journal of Higher Education Theory and Practice*, 24(3), 361–400.
- Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language, Learning and Technology*, 27(1), 1–19.

- Li, S., Chen, J., & Zhang, M. (2025). Can generative AI chatbots promote second language acquisition? A meta-analysis. *Journal of Computer Assisted Learning*, 41(2), e70060. <https://doi.org/10.1111/jcal.70060>
- Lyu, Y., Kim, H., & Park, J. (2025). Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12668>
- Tai, T.-Y., & Chen, H. H.-J. (2023). The impact of intelligent personal assistants on adolescent EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 37(1–2), 1–28.
- Tai, T.-Y., & Chen, H. H.-J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 37(3), 433–460.
- Tupas, R. (2015). Inequalities of multilingualism: Challenges to mother tongue-based multilingual education. *Language and Education*, 29(2), 112–124.
- Ulla, M. B., Advincula, M. J. C., & Mombay, C. D. (2024). How can GenAI foster an inclusive language classroom? A critical language pedagogy perspective from Philippine university teachers. *Computers and Education: Artificial Intelligence*, 7, 100247.
- UNESCO. (2022). *Technology in education: A case study on the Philippines*. United Nations Educational, Scientific and Cultural Organization.
- Wang, F., Cheung, A. C. K., Neitzel, A. J., & Chai, C. S. (2024). Does chatting with chatbots improve language learning performance? A meta-analysis. *Review of Educational Research*. <https://doi.org/10.3102/00346543241255621>
- World Bank. (2022). *Digital transformation of Philippine higher education* (Report No. AUS0002964). The World Bank Group.
- Wu, X., & Li, R. (2024). Unraveling effects of AI chatbots on EFL learners' language skill development: A meta-analysis. *Asia-Pacific Education Research*.
- Yang, Y., Sun, H., Chai, Z., Jiang, G., & Estigoy, E. (2024). Usefulness, ease-of-use, and acceptance towards generative AI in language learning of non-language majors: A TAM-based study. *International Journal of Advanced Engineering Research and Science*, 11(6).
- Yim, R. (2024). Teachers' perceptions, attitudes, and acceptance of artificial intelligence educational learning tools. *Future in Educational Research*, 2(4), 318–345.