

1. Marco Metodológico para el Paper Académico

Versión: 1.9

Fecha: Junio 2026

Protocolo IA-Socrático — USACH

Cambios v1.8 → v1.9 (dictamen IP — 2026-06-12): - §5.3: Agregada H_{cohorte} (homogeneidad de cohorte) como variable exploratoria secundaria: similitud semántica media entre pares de rastros $M1$ de estudiantes distintos, por clase, trayectoria $C1 \rightarrow C5$, calculada por `AGENT_TRAYECTORIA`. Motivación: Liu et al. (2024, arXiv:2401.06816, preprint) — la homogeneización inducida por uso directo de IA persiste al retirar la herramienta. Doble función: réplica conceptual en dominio ingenieril y autocontrol del andamiaje socrático. Definición canónica en Glosario v1.10.

Cambios v1.7 → v1.8 (decisión IP — 2026-06-02): - Convención `AGENT_` para todos los agentes automatizados. Renombrado: Analista SFL → `AGENT_ANALISTA_SFL`. - Agregados `AGENT_SESSION` (intra-sesión, ejecuta `DD_16`) y `AGENT_TRAYECTORIA` (inter-sesión, heatmap longitudinal). - §4: Filas “Feedback al alumno” e “Informe analítico” ahora referencian `AGENT_SESSION`. Agregada fila “Informe longitudinal” (`AGENT_TRAYECTORIA`). - §5.1: Separada codificación primaria (`AGENT_SESSION`) de codificación longitudinal (`AGENT_TRAYECTORIA`). Ambas invocan `AGENT_ANALISTA_SFL` como motor SFL compartido.

Cambios v1.6 → v1.7 (decisión IP — 2026-06-02): - §4: Fila “Feedback al alumno” actualizada: el contenido es SFL de proceso (no dimensional). Tabla de ejemplos SÍ/NO agregada. - §5.1: Codificación primaria ahora especifica que el feedback describe hábitos cognitivos observados sin revelar niveles $D1-D4$ ni las dimensiones medidas. Esta distinción protege la validez interna del estudio. - §5.1: Agregada tabla de ejemplos de feedback SFL de proceso vs. feedback dimensional prohibido. - §4: Agregada Tabla 4b con distribución de momentos $M1-M4$ por clase (no todas las clases tienen los 4 momentos).

Cambios v1.5 → v1.6 (revisión SFL — Ingrid Westhoff Podestá, 2026-05-31): - §2: Agregada fila “Operacionalización de $D1-D4$ ” con fundamentación SFL y referencia a Rúbrica v1.6. - §5.1: Codificación para el paper ahora declara que los criterios se basan en recursos lingüísticos observables (SFL). Referencia a Rúbrica v1.6 §2 y Anexo A. - §5.2: “Análisis de discurso” expandido con definiciones operacionales de cambio superficial (variación léxica) y cambio estructural (incorporación de recursos lógico-semánticos). Referencia a `doc_inv_SFL_Analisis_v1.1.md`. - §6: Confiabilidad expandida — los criterios lingüísticos SFL aumentan la replicabilidad de la codificación. - §8: Agregada limitación 8 — asimetría de géneros entre $M1$ (género cerrado, 9 secciones) y $M4$ (género abierto). Δ_{intra} debe interpretarse con cautela. - §9: Marco teórico del paper ahora incluye SFL como marco de análisis del razonamiento técnico.

Cambios v1.4 → v1.5 (inspección IP 2026-05-30): - “Prueba definitiva” reemplazada por “transferencia cercana como evidencia de internalización” en §9 (estructura del paper). - Limitación §8.6 agregada: observador = ejecutor de la intervención (debilita independencia de triangulación). - Limitación §8.7 agregada: asimetría de comportamiento

NEUTRO vs C1-C4 (amenaza a validez interna declarada).

Cambios v1.3 → v1.4 (auditoría IP 2026-05-24): 1. Timeline DD_24 corregido: fases alineadas con protocolos v1.1 (PLAN 15-37, transición 37-40, BUILD 40-72).

Cambios v1.2 → v1.3 (alineación completa con Premisas de Diseño v1.0, 2026-05-19): 1. Muestra: $n \approx 20$ (P4); ausencias como dato, no exclusión (DD_36). 2. Nombre canónico del curso (P4). 3. Nueva sección 3: descripción completa de la intervención (DD_13, DD_24, DD_19, DD_25/DD_26, DD_27, DD_31, DD_35, DD_39). 4. Feedback diferenciado: alumno (proceso) y profesor (informe analítico) (DD_16). 5. Arquitectura de plataforma: Astro + n8n + PostgreSQL (DD_15).

1.1 1. Título oficial tentativo

Deuda cognitiva en educación técnica: un protocolo de intervención con chatbot socrático para el rediseño del aprendizaje universitario en contexto de IA generativa

1.1.1 Alternativas descartadas

Opción A — Descriptiva/metodológica (anterior): > Rediseño del aprendizaje técnico universitario en contexto de IA generativa: un protocolo socrático para capturar proceso de razonamiento

Opción B — Posicionamiento crítico: > No prohíbas la IA, rediseña la clase: evidencia de un protocolo socrático que captura razonamiento técnico en ingeniería

Motivo de selección: la Opción oficial nombra el problema (deuda cognitiva) antes que la solución, protege la validez metodológica al no sobrevender generalización, y es consistente con el Documento Rector v1.1.

1.2 2. Diseño de investigación

Elemento	Especificación
Tipo	Investigación-acción con diseño cuasi-experimental de medidas repetidas
Paradigma	Mixto (cualitativo predominante, con métricas cuantitativas de apoyo)

Elemento	Especificación
Unidad de análisis	Trayectoria de razonamiento del estudiante a través de 5 sesiones
Variable independiente	Secuencia metodológica del protocolo: rastro en papel → contraste socrático personalizado → revisión profesional de un entregable imperfecto → transferencia autónoma con herramienta disponible pero sin andamiaje
Variable dependiente	Desplazamiento cognitivo en 4 dimensiones (D1–D4)
Operacionalización de D1-D4	Las dimensiones se miden como recursos lingüísticos observables fundamentados en la Lingüística Sistémico-Funcional (Halliday & Matthiessen, 2014): D1 (complejidad causal) = nexos causales explícitos y nominalizaciones; D2 (especificidad técnica) = participantes con modificación numérica y unidades; D3 (consciencia epistémica) = operadores modales y marcadores evidenciales; D4 (decisión bajo incertidumbre) = verbos de decisión en primera persona y cláusulas condicionales de consecuencia. Ver Rúbrica Longitudinal v1.6, §2 y Anexo A
Control	El rastro inicial de Clase 1 (línea base intra-sujeto)

1.3 3. Participantes

Ítem	Descripción
Población	Estudiantes de Ingeniería Industrial, USACH
Asignatura	Laboratorio de Máquinas y Equipos Industriales (14362-0-L-1)
Muestra	Cohorte completa del semestre (censo de conveniencia)
Criterio de inclusión	Asistencia obligatoria a Clase 1 (produce el rastro inicial = control intra-sujeto). Sin rastro de C1 el caso es analíticamente inútil aunque cumpla 4/5.

Ítem	Descripción
Tratamiento de ausencias	Las ausencias son dato, no exclusión (DD_36). El análisis cuantitativo central usa casos con asistencia completa (C1-C5). Estudiantes con sesiones faltantes no se excluyen de la muestra : pasan a una sección de análisis exploratorio donde el hueco se analiza como hallazgo (“alumno que faltó a C3 muestra X en C4”). No se pierde muestra.
Consentimiento	Autorización de uso investigativo de registros de clase: firma física pre-C1 + confirmación digital en primer acceso a plataforma

1.4 3b. Descripción de la intervención

1.4.1 Principio de puesta en escena pedagógica

La intervención se ejecuta como un laboratorio de diagnóstico industrial, no como una demostración tecnológica. El estudiante no entra primero a una herramienta de IA: entra a una instalación con síntomas, sensores imperfectos, operadores humanos, datos incompletos y una decisión que debe defender. La infraestructura digital existe para capturar y presionar el razonamiento, pero la escena pedagógica se organiza alrededor del sistema técnico bajo incertidumbre.

La IA se interpreta metodológicamente como presión cognitiva subordinada: no enseña máquinas por el estudiante, sino que obliga al estudiante a pensar como técnico frente a una máquina. Por eso la evidencia principal no es la conversación con el chatbot, sino el desplazamiento observable en la calidad del razonamiento técnico entre rastro inicial, contraste, revisión profesional y transferencia.

1.4.2 Estructura de cada sesión (DD_24)

Cada clase de 80 minutos (P4) sigue una secuencia de 6 fases:

Fase	Tiempo	Actividad
0 — Encuadre	0-5 min	Devolución de feedback de la clase anterior; aclaración de dudas
1 — Rastro en papel	5-15 min	Alumno produce diagnóstico/solución a mano, sin IA; sube foto vía “Subir imagen” (DD_38)

Fase	Tiempo	Actividad
2 — PLAN socrático	15-37 min	Chatbot cuestiona el razonamiento del alumno sin generar contenido
3 — Transición	37-40 min	Profesor activa modo BUILD desde la interfaz (DD_19); intervención verbal grupal
4 — BUILD	40-72 min	Chatbot genera entregable con errores; alumno evalúa en formato libre (DD_25/DD_26)
5 — Cierre	72-80 min	Profesor activa reflexión Δ _intra (DD_30) desde la interfaz

Nota Clase 5: No sigue esta estructura; el alumno accede al chatbot en modo NEUTRO con un caso de transferencia nuevo.

1.4.3 Entregables BUILD por clase (DD_13)

Clase	Entregable que genera el chatbot	Nivel de errores deliberados (DD_8)
C1	— (solo PLAN socrático básico)	—
C2	Informe técnico de diagnóstico	Obvios (fácilmente detectables)
C3	Reporte de análisis de tendencias SCADA	Sutiles (requieren conocimiento técnico)
C4	Protocolo de emergencia + análisis económico	Profesionales (parecen correctos, requieren criterio experto)
C5	— (chatbot NEUTRO, no genera documento)	—

1.4.4 Comportamiento defensivo del chatbot (DD_27)

Cuando el alumno señala un error en el documento BUILD, el chatbot no confirma: **defiende su posición con contraargumentos técnicos**. El alumno debe sostener su postura con evidencia. El chatbot cede solo ante argumentación sólida. Esto cambia la naturaleza de la interacción: no es corrección pasiva sino debate técnico.

1.4.5 Evaluación del entregable BUILD (DD_25/DD_26)

El alumno evalúa el documento en formato libre dentro del chat (no con checklist ni formulario externo). La calidad de su revisión es dato primario del estudio.

1.4.6 Escenarios técnicos acumulativos (DD_39)

C1-C4 usan escenarios del mismo Centro Acuático (piscina semi-olímpica) con complejidad creciente. C5 presenta un caso industrial nuevo (torre de enfriamiento) para medir transferencia.

1.4.7 Controles docentes

El profesor controla la transición PLAN→BUILD (DD_19) y la activación de la reflexión de cierre (DD_30) desde la interfaz. El alumno no puede activar estos cambios. No hay límite de turnos en el chat (DD_31); la cantidad es dato orgánico.

1.4.8 Preparación pre-clase (DD_35)

El día anterior a cada sesión, el alumno recibe por WhatsApp un mensaje con el vocabulario técnico clave del caso, para nivelar el punto de partida.

1.4.9 Arquitectura de plataforma (DD_15)

La plataforma es un desarrollo propio del IP compuesto por tres capas: - **Frontend:** Astro (desplegado en Cloudflare) — interfaz del alumno y del profesor - **Orquestador:** n8n — gestiona flujos (onboarding, inyección de datos al chatbot, envío de feedback por WhatsApp, análisis de imagen) - **Persistencia:** PostgreSQL — almacena conversaciones, rastros digitalizados, feedback, métricas

El LLM se consume vía API (sin fine-tuning con datos de alumnos). Los datos no salen de la infraestructura del proyecto.

1.5 4. Instrumentos de recolección de datos

Instrumento	Clase	Qué captura
M1 — Rastro en papel (foto subida vía plataforma → AI Vision → BD)	1–5	Razonamiento pre-IA: línea base individual por sesión
M2 — Log de interacción PLAN (chatbot socrático/adversarial)	1–4	Trayectoria de refinamiento cognitivo bajo presión
M3 — Log de interacción BUILD (alumno evalúa documento con errores en chat)	2–4	Capacidad evaluativa: detección de errores, calidad argumentativa, aceptación/rechazo

Instrumento	Clase	Qué captura
M4 — Reflexión de cierre DD_30 (chatbot pide qué cambiaría del rastro inicial; logs en PostgreSQL)	1–5	Cierre cognitivo: evidencia oficial de Δ_{intra} (M4 – M1)
Indicador DD_29 (push “¿lo firmarías?”)	2–4	Dato binario: ¿el alumno aceptó sin cuestionar? ¿Recapacitó tras el push?
Log chatbot neutro	5	Interacción autónoma: ¿reproduce metodología sin guía?
Encuesta final (Likert + abiertas, 4 ejes)	5	Autopercepción de cambio en relación con IA
Feedback al alumno (AGENT_SESION → WhatsApp, automático sin revisión humana)	1–5	Retroalimentación de proceso (no de correctitud); dato secundario. El contenido es SFL de proceso, nunca dimensional: describe hábitos cognitivos observados (cómo conectó variables, cómo revisó hipótesis bajo contradicción) sin revelar niveles D1-D4 ni las dimensiones medidas. Ver ejemplos en §5.1.
Informe analítico al profesor (AGENT_SESION → interfaz docente)	1–5	Análisis consolidado del desempeño del grupo para toma de decisiones pedagógicas
Informe longitudinal (AGENT_TRAYECTORIA → interfaz docente)	Post-C2 a C5	Δ_{inter} , heatmap de trayectorias, detección de estancamientos/regresiones/saltos, comparación M1_C1→C5
Notas de observación docente	1–5	Comportamiento en aula, intervenciones, indicadores no digitales

Tabla 4b. Distribución de momentos M1-M4 por clase

Clase	M1 (Papel)	M2 (Chat PLAN)	M3 (Chat BUILD)	M4 (Cierre DD_30)	Total
C1	✓	✓ (PLAN básico)	—	✓	3 momentos
C2	✓	✓ (PLAN)	✓ (BUILD)	✓	4 momentos

Clase	M1 (Papel)	M2 (Chat PLAN)	M3 (Chat BUILD)	M4 (Cierre DD_30)	Total
C3	✓	✓ (PLAN)	✓ (BUILD)	✓	4 momentos
C4	✓	✓ (PLAN adversarial)	✓ (BUILD)	✓	4 momentos
C5	✓	✓ (NEUTRO)	—	✓	3 momentos

C1 no tiene BUILD por ser línea base (DD_21). C5 no tiene BUILD ni PLAN: chatbot NEUTRO para medir transferencia sin andamiaje. M3 (BUILD) no se usa para Δ_{intra} : es indicador independiente de capacidad evaluativa (D3, DD_27).

1.6 5. Procedimiento de análisis

1.6.1 5.1 Codificación (dual: IA + humano)

Codificación primaria (tiempo real, formativa): AGENT_SESION invoca a AGENT_ANALISTA A_SFL sobre M1, M2, M3 y M4 automáticamente post-sesión para codificar D1-D4. Genera: (a) feedback de proceso al alumno vía WhatsApp, (b) informe analítico al profesor. No requiere revisión humana. Datos en PostgreSQL.

Codificación longitudinal (inter-sesión): AGENT_TRAYECTORIA invoca a AGENT_ANALISTA _SFL sobre los rastros M1 de sesiones consecutivas. Calcula $\Delta_{inter} = M1_N - M1_{(N-1)}$ por dimensión. Detecta patrones: estancamientos ($\Delta_{inter} = 0 \times 2$ sesiones), regresiones ($\Delta_{inter} < 0$), saltos ($\Delta_{inter} \geq +2$). Genera heatmap de trayectorias individuales. Se dispara post-C2, C3, C4 y C5 (acumulativo).

Contenido del feedback al alumno: El feedback es **SFL de proceso, nunca dimensional**. Describe hábitos cognitivos observados sin revelar niveles D1-D4 ni las dimensiones que se miden. Esta distinción es metodológicamente necesaria: si el alumno supiera qué se está midiendo, sus respuestas en sesiones siguientes podrían adaptarse a la métrica (características de demanda), inflando artificialmente Δ_{inter} y destruyendo la validez interna del estudio.

Lo que SÍ dice el feedback (SFL de proceso)	Lo que NO dice (dimensional, prohibido)
---	---

“Conectaste varias variables pero tus cadenas causales eran de un solo eslabón. La próxima vez intenta trazar el camino completo: ¿qué causó qué, y eso qué provocó después?”

“Estás en D1 nivel 2. Necesitas más nexos causales para llegar a nivel 3.”

Lo que SÍ dice el feedback (SFL de proceso)	Lo que NO dice (dimensional, prohibido)
“Cuando el chatbot te mostró datos contradictorios, revisaste tu hipótesis. Ese hábito te va a servir en el caso de la torre de enfriamiento.”	“Tu D3 mejoró de nivel 1 a nivel 2. Sigue así.”
“Tus variables estaban bien identificadas pero sin valores. En ingeniería, un dato sin número es una opinión.”	“D2 nivel 1: te faltan participantes con modificación numérica.”

El primer tipo enseña a razonar. El segundo enseña a pasar un test.

Codificación para el paper (investigativa): 1. Dos codificadores humanos independientes asignan niveles (1–4) en cada dimensión, sobre datos codificados (sin nombres). Los criterios de codificación se basan en **recursos lingüísticos observables** derivados de la Lingüística Sistémico-Funcional (Halliday & Matthiessen, 2014): cada nivel de D1-D4 se define por la presencia o ausencia de marcadores lingüísticos específicos (nexos causales, nominalizaciones, operadores modales, marcadores evidenciales, cláusulas condicionales, verbos de decisión). Ver Rúbrica Longitudinal v1.6, §2 columna “Criterio de codificación lingüística” y Anexo A para definiciones operacionales y ejemplos 2. Se calcula acuerdo inter-codificador con **Cohen’s κ ponderado** (lineal o cuadrático, por la naturaleza ordinal de la rúbrica: un desacuerdo $1 \leftrightarrow 4$ pesa más que $1 \leftrightarrow 2$). Umbral ≥ 0.80 (substantial-strong, Landis & Koch) para reporte publicable; $\kappa \geq 0.70$ acepta solo análisis exploratorio 3. Desacuerdos se resuelven por consenso con un tercer evaluador 4. **Dato metodológico adicional:** se calcula κ entre AGENT_ANALISTA_SFL y codificadores humanos. Si $\kappa_{IA-humano} \geq 0.80$, es hallazgo publicable sobre viabilidad de codificación automatizada de razonamiento; si queda entre 0.70 y 0.79, se reporta solo como análisis exploratorio

1.6.2 5.2 Análisis cualitativo

- **Análisis de contenido temático** de rastros iniciales y producciones finales
- **Comparación constante** entre sesiones (progresión individual)
- **Análisis de discurso** de interacciones con chatbot, fundamentado en la Lingüística Sistémico-Funcional (Halliday & Matthiessen, 2014). Se distinguen dos tipos de cambio: **cambio superficial** (variación léxica sin incorporación de nuevos recursos lógico-semánticos — el estudiante cambia palabras pero no la estructura del argumento) y **cambio estructural** (incorporación de nuevos recursos de las metafunciones ideacional o interpersonal — el estudiante añade nexos causales, nominalizaciones, marcadores evidenciales o cláusulas condicionales que no estaban presentes en su rastro inicial). Ver `do_c_inv_SFL_Analisis_v1.1.md` para el marco completo
- **Casos extremos:** análisis profundo de estudiantes con máximo y mínimo desplazamiento

1.6.3 5.3 Análisis cuantitativo (descriptivo-exploratorio)

Dado un n esperado de ≈ 20 (cohorte completa de una asignatura, máximo ~ 20 alumnos), el componente cuantitativo se reporta como **descriptivo-exploratorio**, no inferencial confirmatorio.

- Media y desviación estándar de niveles por dimensión y por clase
- Δ_{intra} y Δ_{inter} por dimensión (medianas, rangos intercuartílicos)
- **Wilcoxon emparejado para diferencias entre sesiones:** solo se reporta acompañado de (a) tamaño de efecto r de Wilcoxon o d de Cohen para muestras pareadas, y (b) declaración explícita de poder estadístico. Para detectar un efecto medio ($d \approx 0.5$) con $\alpha = 0.05$ y poder 0.80, se requieren ~ 35 pares; con $n < 35$, p-values se interpretan como exploratorios.
- **H_cohorte — homogeneidad de cohorte (variable exploratoria secundaria):** similitud semántica media entre todos los pares de rastros M1 de estudiantes distintos, por clase (embeddings de texto, similitud coseno), calculada por AGENT_TRAYECTORIA sobre los rastros ya digitalizados, sin instrumentos nuevos ni carga adicional al estudiante. Motivación: Liu et al. (2024, arXiv:2401.06816, preprint) documentaron experimentalmente que el uso directo de IA generativa homogeneiza las producciones y que esa homogeneización persiste al retirar la herramienta. En este estudio cumple doble función: (a) réplica conceptual de ese hallazgo en dominio ingenieril; (b) autocontrol del diseño — detectar si el propio andamiaje socrático comprime la diversidad de razonamiento de la cohorte. Cautela interpretativa: la ficha M1 es un género cerrado de 9 secciones que impone similitud estructural de base, por lo que H_cohorte se interpreta por su trayectoria relativa C1→C5, nunca por su valor absoluto. No participa del cálculo de Δ_{intra} ni Δ_{inter} y no genera feedback a alumnos ni al profesor durante el piloto. Definición canónica en Glosario v1.10
- Visualización: heatmap de trayectorias individuales

1.7 6. Validez y confiabilidad

Criterio	Estrategia
Validez interna	Diseño intra-sujeto (cada estudiante es su propio control); secuencia fija; captura antes/después
Validez externa	Documentación completa del protocolo para replicación; caso técnico transferible a otras ingenierías
Credibilidad	Triangulación de fuentes (rastros + chatbot + observación docente); member checking

Criterio	Estrategia
Transferibilidad	Descripción densa del contexto; protocolo público y replicable
Confiabilidad	Codificación doble con $\kappa \geq 0.80$ (ponderado); audit trail del análisis. Los criterios de codificación se basan en recursos lingüísticos observables (SFL), lo que aumenta la replicabilidad: dos codificadores que cuentan nexos causales o identifican marcadores evidenciales llegan al mismo resultado con más frecuencia que dos codificadores que interpretan “calidad del razonamiento” por intuición
Confirmabilidad	Todos los datos brutos preservados; decisiones de codificación documentadas

1.8 7. Consideraciones éticas

- Consentimiento informado entendido como autorización de uso investigativo de registros de clase (firma física pre-C1 + confirmación digital), con derecho a retirar esa autorización sin consecuencia académica
- Anonimización de datos en publicación (código alfanumérico)
- El piloto NO afecta calificaciones del curso
- Datos almacenados en PostgreSQL centralizada con acceso restringido; AGENT_ANALISTA_SFL procesa sin revisión humana pero no transmite a terceros ni entrena con los datos
- LLM del chatbot usado vía API (sin fine-tuning con datos de alumnos)
- **Engaño leve declarado:** En C2-C4, el documento BUILD contiene errores deliberados sin conocimiento del alumno (se presenta como “informe basado en tu trabajo”). Justificación: simula condición profesional real; no genera daño; debriefing post-C5 explica el mecanismo
- Aprobación del comité de ética de la Facultad (tramitar antes del inicio)

1.9 8. Limitaciones anticipadas

1. **Muestra por conveniencia:** una sola cohorte, una sola asignatura
2. **Efecto Hawthorne:** los estudiantes saben que participan en un piloto
3. **Sin grupo control:** diseño intra-sujeto mitiga parcialmente, pero no hay comparación con clase sin protocolo

4. **Generalización limitada:** caso técnico específico (hidráulica) puede no transferir a otras disciplinas sin adaptación
5. **Dependencia del chatbot:** resultados pueden variar según plataforma y configuración del modelo
6. **Observador = ejecutor de la intervención:** el docente que completa la ficha de observación (Protocolo de Observación Docente v1.3) es el mismo que activa BUILD, activa el cierre e interviene verbalmente durante la sesión. Esto debilita la independencia de la triangulación observacional. La ficha se concentra en comportamientos no digitalizables (afecto, insight, comentarios espontáneos) para minimizar el solapamiento con los datos que el dashboard ya captura digitalmente. Una alternativa futura sería un observador independiente o grabación de audio
7. **Asimetría de comportamiento del chatbot entre C1-C4 y C5:** en C1-C4, el chatbot redirige al alumno a su hoja cuando pide datos y añade una pregunta de re-enganche (“¿revisaste tu hoja?”). En C5 (NEUTRO), la redirección es plana, sin pregunta de re-enganche, porque no hay presión socrática. Esta diferencia de comportamiento es intencional (P5, DD_17, DD_20), pero cualquier diferencia observada en la trayectoria C1→C5 puede deberse parcialmente a esta asimetría de diseño, no solo a la internalización del método. Se declara como amenaza a la validez interna
8. **Asimetría de géneros entre M1 y M4:** M1 (ficha Pre-AI) es un género cerrado de 9 secciones obligatorias que fuerza la activación de las metafunciones ideacional, interpersonal y textual. M4 (cierre DD_30) es un género abierto de respuesta libre donde el estudiante elige qué metafunción activar. Las escalas D1-D4 se aplican a ambos momentos, pero la exigencia estructural no es equivalente: un nivel bajo en M4 puede reflejar que la pregunta no forzó ese recurso lingüístico, no que el estudiante no lo posee. $\Delta_{intra} = M4 - M1$ debe interpretarse con cautela. Esta limitación se fundamenta en el análisis SFL del protocolo (`doc_inv_SFL_Analisis_v1.1.md §4`)

1.10 9. Estructura propuesta del paper

1. Introducción

- Problema: IA generativa y deuda cognitiva en educación superior
- Brecha: las respuestas actuales (prohibir, ignorar) no resuelven el problema de diseño instruccional
- Contribución: protocolo socrático que captura proceso de razonamiento

2. Marco teórico

- Deuda cognitiva (concepto propuesto)
- Método socrático y presión cognitiva
- Transferencia como evidencia de aprendizaje profundo
- IA como herramienta pedagógica vs. sustituto cognitivo
- Lingüística Sistémico-Funcional como marco de análisis del razonamiento técnico (Halliday & Matthiessen, 2014): el razonamiento se manifiesta en recursos lingüísticos ob-

servables — nexos causales (D1), densidad léxica (D2), modalidad epistémica (D3), actos de habla deónticos (D4)

3. Metodología

- Diseño del estudio
- Participantes y contexto
- El protocolo: 5 sesiones, secuencia, roles
- Instrumentos y procedimiento de análisis

4. Resultados

- Trayectorias de desplazamiento cognitivo (D1–D4)
- Análisis por clase: qué produce el mayor salto
- Casos ilustrativos (máximo vs. mínimo desplazamiento)
- Evidencia de transferencia (Clase 5)

5. Discusión

- ¿El protocolo previene la deuda cognitiva?
- ¿Qué fase es más crítica para el aprendizaje?
- Implicaciones para el diseño instruccional universitario
- Limitaciones y futuras investigaciones

6. Conclusiones

- El protocolo es ejecutable, medible y replicable
- La secuencia importa: rastro antes, IA después
- La transferencia cercana como evidencia de internalización

1.11 10. Revistas objetivo (por orden de ajuste)

Revista	Área	Factor de impacto	Justificación
<i>Computers & Education</i>	Ed-Tech	~12	Máximo impacto en tecnología educativa
<i>Studies in Higher Education</i>	Educ. Superior	~4	Enfoque en pedagogía universitaria
<i>International Journal of Engineering Education</i>	Ing. + Educ.	~1.5	Público exacto del piloto

Revista	Área	Factor de impacto	Justificación
<i>Teaching in Higher Education</i>	Educ. Superior	~3	Innovación pedagógica
<i>Revista Educación en Ingeniería (ACOFI)</i>	Regional	—	Visibilidad latinoamericana, español

1.12 11. Cronograma de publicación

Fase	Fecha estimada	Entregable
Aprobación comité de ética	Mayo–Junio 2026	Acta de aprobación
Ejecución del piloto	Mayo–Junio 2026	Datos brutos de 5 sesiones
Codificación y análisis	Julio 2026	Dataset codificado, κ ponderado calculado
Primer borrador	Agosto 2026	Manuscrito completo
Revisión interna	Septiembre 2026	Versión corregida
Envío a revista	Octubre 2026	Submission

Cronograma original (v1.0) declaraba fechas 2025; reprogramado 12 meses en v1.1 por desfase observado. Si el calendario real difiere, actualizar antes del envío a comité.