

# MANUAL DE EJECUCIÓN EMPÍRICA

## EXPERIMENTO CPO-001

*Aislamiento, Medición y Validación de la Señal de Proceso en la Intuición Experta*

**AUTOR Y  
DISEÑADOR**  
**Joseph  
Castillo**

*Arquitecto del  
experimento*

**VERSIÓN**  
**2.0 — Operativa**  
*Lista para ejecución en campo*

**HIPÓTESIS A FALSEAR**  
**H<sub>1</sub> CPO**  
*Señal tácita en trayectoria de  
proceso*

**ESTADO**  
**ACTIVO**  
*Ejecución  
autorizada*

**TASA DE ACIERTO  
REQUERIDA**  
**> 70%**  
*p < 0.05 — test binomial  
unilateral*

**ÍNDICE KAPPA REQUERIDO**  
**K > 0.60**  
*Acuerdo inter-juez — Cohen's Kappa*

**CONDICIÓN DE REFUTACIÓN**  
**≤ 60%**  
*H<sub>0</sub> no rechazada — rediseño de  
arquitectura*

## OBJETIVO EMPÍRICO

*Declaración operativa de lo que este experimento prueba*

Yo, Joseph Castillo, diseño este experimento con un único objetivo técnico: demostrar o refutar que la trayectoria de proceso de un agente experto contiene una señal estructural distingible de la trayectoria de un agente novato, medible de forma cuantitativa, sin acceso al resultado final de la tarea.

Este es un experimento de detección de señal, no un experimento de evaluación de calidad. No pregunto si el experto produce mejores resultados que el novato —eso es trivial y conocido. Pregunto si el camino que el experto recorre hacia cualquier resultado es estructuralmente distingible del camino que recorre el novato, y si esa distinción puede ser identificada por un observador con conocimiento del dominio que únicamente tiene acceso a los datos de proceso —tiempos, secuencias, retrocesos, profundidad de exploración— y nunca al outcome.

### ENUNCIADO FORMAL DEL OBJETIVO:

Aislar la señal tácita de proceso en registros de actividad de agentes con niveles de destreza verificablemente distintos, bajo condición de ceguera total al resultado, y determinar si dicha señal es suficientemente robusta como para ser identificada con consistencia por jueces independientes con conocimiento del dominio.

RESULTADO BUSCADO: Tasa de acierto > 70%,  $p < 0.05$ ,  $\kappa > 0.60$  entre jueces.

RESULTADO DE REFUTACIÓN: Tasa  $\leq 60\%$ , independiente del valor de  $\kappa$ .

### I.1 Lo que este experimento NO prueba

Establezco con precisión los límites de este diseño experimental para evitar sobre interpretaciones de sus resultados. Este experimento no prueba que el saber tácito pueda ser transferido entre agentes. No prueba que los indicadores VSP sean suficientes para construir un clasificador automático. No prueba la generalización de la señal a otros dominios. Estas son hipótesis derivadas que requieren fases de validación posteriores y autónomas. Este experimento prueba únicamente la condición necesaria: la señal existe y es detectable por observación humana informada.

## II

**DISEÑO DE LA MUESTRA***Perfiles exactos de Generadores y Jueces — criterios de selección y exclusión***II.1 Los Generadores: quiénes producen los registros de proceso**

Designo como Generadores a los agentes cuyas trayectorias de proceso constituyen el material empírico del experimento. Requiero exactamente cinco Generadores divididos en dos grupos de destreza verificada. El criterio de destreza no es auto reportado: debe ser verificable mediante evidencia externa objetiva definida a continuación.

GRUPO	N	CRITERIO DE EXPERTICIA	CRITERIO DE EXCLUSIÓN	CÓDIGO ASIGNADO
GRUPO A — EXPERTO	2 o 3	Mínimo 5 años de práctica activa y continua en el dominio. Desempeño verificable por terceros (resultados documentados, historial de logros, reconocimiento por pares). No basta la autoevaluación.	Menos de 5 años activos. Práctica discontinua. Incapacidad de demostrar evidencia externa de nivel.	GEN-E01, GEN-E02, GEN-E03
GRUPO B — NOVATO	2 o 3	Menos de 12 meses de exposición al dominio. Conocimiento básico de las reglas operativas, pero sin destreza consolidada ni automatismos verificables.	Más de 18 meses de práctica. Conocimiento de nivel intermedio demostrado. Haber sido entrenado por los Generadores expertos.	GEN-N01, GEN-N02, GEN-N03

### ⚠ REGLA DE ANONIMIZACIÓN — OBLIGATORIA:

Los códigos GEN-E01, GEN-E02, etc., son los únicos identificadores que circulan fuera de esta tabla.

Los nombres reales de los Generadores no aparecen en ningún documento que llegue a los Jueces.

El Coordinador del experimento es la única persona que conoce la correspondencia código-nombre.

Esta correspondencia se revela únicamente después de que todos los Jueces hayan entregado sus hojas de respuesta completas y cerradas.

## II.2 Los Jueces: quiénes evalúan los registros de proceso

Designo como Jueces a los cinco evaluadores ciegos que clasificarán los registros de proceso sin acceso al resultado de las tareas. El perfil del Juez es crítico: debe tener conocimiento suficiente del dominio para leer los registros de proceso con comprensión, pero no puede conocer personalmente a ningún Generador.

CRITERIO	REQUISITO	VERIFICACIÓN	CONSECUENCIA DE INCUMPLIMIENTO
Conocimiento de dominio	Suficiente para interpretar una secuencia de acciones en el dominio elegido sin asistencia. No requiere ser experto.	Prueba de comprensión: el Juez debe describir correctamente el propósito de al menos 3 acciones del dominio seleccionadas al azar.	Exclusión del panel. Reemplazo obligatorio antes de iniciar Fase de Evaluación.
Independencia personal	El Juez no conoce personalmente a ningún Generador. Sin relación laboral, familiar o social directa en los 24 meses previos.	Declaración jurada simple. Verificación cruzada por el Coordinador con los Generadores.	Exclusión automática. Los datos de ese Juez se invalidan completamente.
Ceguera al resultado	El Juez nunca recibe información sobre si la tarea fue completada exitosamente, correctamente, o con qué calidad.	El Strip de Proceso entregado al Juez está auditado por el Coordinador antes de la distribución para confirmar ausencia de información de outcome.	Si el Juez recibió información de resultado accidentalmente, sus datos no se incluyen en el análisis Kappa.
Independencia entre jueces	Los Jueces no se comunican entre sí durante la fase de evaluación. Clasificaciones realizadas de forma completamente autónoma.	Los Jueces reciben y devuelven sus hojas en canales separados. No se realizan sesiones grupales de evaluación.	Si se detecta comunicación, los datos de ambos Jueces involucrados se invalidan.
Disponibilidad temporal	El Juez puede completar la evaluación de los 6 registros en una sola sesión de no más de 90 minutos sin interrupciones.	Confirmación explícita de disponibilidad antes de recibir los materiales.	Si el Juez interrumpe la sesión y la retoma, se registra el tiempo de interrupción. Interrupciones > 30 minutos invalidan la sesión.



## PROTOCOLO DE AISLAMIENTO DE VARIABLES — DOBLE CIEGO PARCIAL

*Cómo garantizo que los Jueces evalúan exclusivamente el proceso*

Mi protocolo de aislamiento de variables es el núcleo de la integridad experimental. Si un Juez recibe cualquier información sobre el outcome de la tarea —directa o indirecta— el experimento pierde su condición de ceguera y los datos de ese Juez se invalidan. El diseño de doble ciego parcial que establezco funciona del siguiente modo: los Generadores conocen que están siendo observados (no hay engaño sobre la existencia del experimento) pero no conocen qué métricas específicas se están extrayendo. Los Jueces no conocen a los Generadores y no reciben información de resultado bajo ninguna circunstancia.

### III.1 Construcción del Strip de Proceso — el único material que llega al Juez

El Strip de Proceso es el documento estandarizado que contiene exclusivamente la información de trayectoria que mis Jueces evaluarán. Su construcción es un proceso de depuración activa: se parte del registro completo de la sesión y se elimina sistemáticamente toda información que pueda revelar el outcome.

#### I CONTENIDO INCLUIDO EN EL STRIP DE PROCESO:

- ✓ Secuencia numerada de acciones en orden cronológico estricto.
- ✓ Tiempo en segundos entre acción consecutiva (delta-t por paso).
- ✓ Indicación binaria de retroceso: S/N. Sin razón del retroceso.
- ✓ Número de alternativas exploradas antes de cada decisión irreversible.
- ✓ Duración total de la sesión en segundos.
- ✓ Número de pausas superiores a 15 segundos (sin contenido de la pausa).

#### I CONTENIDO EXPLÍCITAMENTE EXCLUIDO DEL STRIP DE PROCESO:

- X El resultado final de la tarea — NUNCA incluido bajo ninguna circunstancia.
- X Evaluación de calidad del proceso o del resultado.
- X Identidad o código del Generador.

- X Comparaciones con otros Generadores.
- X Cualquier información temporal o contextual que permita identificar al individuo.
- X Imágenes, capturas de pantalla o fragmentos del trabajo realizado.

## III.2 Protocolo de construcción del Strip — procedimiento paso a paso

1. El Coordinador extrae el log de eventos completo de la sesión de grabación.
2. El Coordinador elimina manualmente todas las celdas de contenido de las acciones, preservando únicamente el tipo de acción (acción-A, acción-B, retroceso, pausa) sin descripción de su contenido.
3. El Coordinador realiza una revisión de auditoría: lee el Strip completo preguntándose explícitamente: '¿Puede un lector de este Strip inferir si la tarea fue completada exitosamente?' Si la respuesta es posiblemente sí, el Strip requiere depuración adicional.
4. El Coordinador asigna al Strip un código de registro aleatorio de cuatro dígitos (REG-XXXX) que no guarda ninguna relación con el código del Generador.
5. El Strip finalizado se almacena con su código REG-XXXX. La tabla de correspondencia código-Generador permanece únicamente en poder del Coordinador en documento sellado.
6. Los seis Strips (correspondientes a las seis sesiones evaluables del experimento) se numeran REG-0001 a REG-0006 en orden aleatorio, no en orden de grabación.

### III.3 Protocolo de entrega a los Jueces — control de contaminación

PASO	ACCIÓN	QUIÉN EJECUTA	CONTROL DE INTEGRIDAD
1	Cada Juez recibe los 6 Strips en sobre cerrado o archivo digital encriptado. Todos los Jueces reciben los mismos 6 Strips, en el mismo orden (REG-0001 a REG-0006).	Coordinador	El Coordinador confirma por escrito que entregó materiales idénticos a todos los Jueces.
2	Junto con los Strips, el Juez recibe únicamente: la hoja de instrucción (texto estándar) y la hoja de respuesta VSP. Nada más.	Coordinador	Lista de chequeo de materiales entregados firmada por Juez y Coordinador.
3	El Juez ejecuta la evaluación de forma autónoma, en una sola sesión, sin interrupciones superiores a 30 minutos.	Juez	El Juez registra hora de inicio y hora de cierre de su sesión en la hoja de respuesta.
4	El Juez devuelve la hoja de respuesta completada al Coordinador por el mismo canal de entrega, sin comunicarla a otros Jueces.	Juez	El Coordinador registra la hora de recepción de cada hoja. Las hojas se almacenan sin ser revisadas hasta que las cinco estén recibidas.
5	El Coordinador abre la tabla de correspondencia código-Generador solo después de recibir las cinco hojas de respuesta de todos los Jueces.	Coordinador	El sobre o archivo con la tabla de correspondencia tiene sello de integridad que documenta que no fue abierto antes de este momento.

**IV****MATRIZ DE EXTRACCIÓN VSP***Instrucciones de medición de los 5 Indicadores de Señal de Proceso*

Establezco la Matriz de Extracción VSP como el instrumento central de medición cuantitativa del experimento. Los cinco indicadores que la componen deben ser calculados de forma estandarizada sobre el log de eventos de cada sesión. A continuación, defino con precisión la fórmula de extracción de cada indicador, la unidad de medida, el umbral orientativo que separa el perfil de experto del perfil de novato, y la fuente de datos en el registro.

**VSP-  
1****LATENCIA DEL PRIMER MOVIMIENTO***Tiempo entre exposición al problema y primera acción significativa*

CAMPO	DEFINICIÓN
Definición	Intervalo de tiempo en segundos entre el momento $t_0$ (inicio formal de la sesión, confirmado por el Coordinador en el log) y el momento $t_1$ (primera acción que modifica el estado del problema de forma irreversible).
Unidad de medida	Segundos (resolución mínima: 1 segundo).
Fórmula de extracción	$VSP-1 = t_1 - t_0$ . Donde $t_0$ = timestamp del evento SESSION_START y $t_1$ = timestamp del primer evento clasificado como ACTION_IRREVERSIBLE en el log.
Umbral orientativo	$VSP-1 < 8$ segundos → perfil experto (reconocimiento de patrón). $VSP-1$ entre 8–20 segundos → perfil ambiguo. $VSP-1 > 20$ segundos → perfil novato (exploración sin reconocimiento) o $< 3$ segundos → impulsividad sin base.
Advertencia de medición	Si el Generador verbaliza explícitamente que está leyendo instrucciones al inicio, el Coordinador registra el final de la lectura como $t_0$ efectivo, no el inicio de la sesión.

**VSP-  
2****TASA DE RETROCESO POR FASE***Frecuencia y distribución temporal de correcciones*

CAMPO	DEFINICIÓN
Definición	Número de eventos de retroceso (acciones que revierten o anulan una acción anterior) dividido por el tiempo transcurrido en cada fase del proceso, expresado como retrocesos por minuto.
Unidad de medida	Retrocesos / minuto por fase. Las fases se definen como: Fase Inicial (0–33% del tiempo total), Fase Media (33–67%), Fase Final (67–100%).
Fórmula de extracción	$VSP-2(\text{fase}) = N \text{ retrocesos}(\text{fase}) / T(\text{fase})_{\text{en minutos}}$ . Índice de distribución: $D = VSP-2(\text{initial}) / VSP-2(\text{final})$ . $D > 1$ indica perfil experto (retrocesos tempranos). $D < 1$ indica perfil novato (retrocesos tardíos).
Umbral orientativo	Ratio $D > 1.5 \rightarrow$ perfil experto marcado (detecta errores antes de comprometerse). Ratio $D$ entre 0.7–1.5 $\rightarrow$ perfil ambiguo. Ratio $D < 0.7 \rightarrow$ perfil novato marcado (retrocede después de comprometerse).
Advertencia de medición	Un retroceso técnico (deshacer acción por limitación de la herramienta, no por decisión cognitiva) debe ser excluido del conteo. El Coordinador marca estos eventos durante la transcripción del log con código <b>UNDO_TECHNICAL</b> para distinguirlos de <b>UNDO_COGNITIVE</b> .

**VSP-**  
**3****PROFUNDIDAD DE EXPLORACIÓN PREVIA AL COMPROMISO***Alternativas consideradas antes de decisiones irreversibles*

CAMPO	DEFINICIÓN
Definición	Número promedio de alternativas explícitamente exploradas antes de cada Punto de Decisión Irreversible (PDI) del proceso. Un PDI se define como cualquier acción que requeriría más de 30 segundos de trabajo para ser revertida completamente.
Unidad de medida	Alternativas exploradas por PDI (número adimensional). Promediado sobre todos los PDI de la sesión.
Fórmula de extracción	Para cada PDI_k: registrar N_k = número de acciones clasificadas como ACTION_EXPLORATORY ejecutadas entre el PDI anterior y PDI_k. VSP-3 = $( \sum N_k ) / K$ , donde K es el número total de PDI en la sesión.
Umbral orientativo	VSP-3 entre 1.5–4.0 → perfil experto (exploración dirigida, poda eficiente del espacio). VSP-3 > 7.0 → perfil novato (exploración sin dirección, búsqueda aleatoria). VSP-3 < 1.0 → impulsividad (compromiso sin exploración previa; puede ser experto en tareas de alta familiaridad o novato impulsivo).
Advertencia de medición	La distinción entre ACTION_EXPLORATORY y ACTION_COMMITTED debe establecerse por el Coordinador en la fase de transcripción, no durante la sesión. Criterio: una acción es EXPLORATORY si el Generador puede revertirla en menos de 10 segundos sin costo significativo.

**VSP-  
4****VELOCIDAD DE ABANDONO DE CALLEJONES***Tiempo de permanencia en sub-secuencias que terminan en retroceso*

CAMPO	DEFINICIÓN
Definición	Para cada sub-secuencia de acciones que termina en un evento UNDO_COGNITIVE (retroceso cognitivo), el tiempo transcurrido desde la primera acción de esa sub-secuencia hasta el evento UNDO_COGNITIVE que la cierra. Promediado sobre todas las subsecuencias de retroceso de la sesión.
Unidad de medida	Segundos por sub-secuencia de callejón. Promediado sobre K_callejones sub-secuencias.
Fórmula de extracción	Para cada callejón C_k: VSP-4_k = t(UNDO_COGNITIVE_k) – t(primer_acción_C_k). VSP-4 = ( Σ VSP-4_k ) / K_callejones. Si K_callejones = 0 (ningún retroceso cognitivo), VSP-4 = 0 y se reporta como N/A para esta sesión.
Umbral orientativo	VSP-4 < 25 segundos → perfil experto (abandono temprano, reconocimiento anticipatorio de fracaso). VSP-4 entre 25–60 segundos → perfil ambiguo. VSP-4 > 60 segundos → perfil novato (persistencia en callejones confirmados).
Advertencia de medición	Sessions con K_callejones < 2 producen un VSP-4 estadísticamente poco representativo. El Coordinador registra el valor pero lo señala con bandera BAJA_CONFIANZA_VSP4 en la hoja de datos. Este flag no invalida la sesión pero se reporta en el análisis.

**VSP-****5****COHERENCIA ESTRUCTURAL DE LA SECUENCIA***Medición del orden lógico del proceso*

CAMPO	DEFINICIÓN
Definición	Medida del grado en que la secuencia de acciones de una sesión sigue el esquema de fases esperado para el dominio (apertura → desarrollo → consolidación → cierre). Cuantificado como el porcentaje de acciones clasificadas como 'en fase correcta' sobre el total de acciones de la sesión.
Unidad de medida	Porcentaje de acciones en fase correcta (0–100%). Para cada dominio, el Coordinador define previamente el esquema de fases esperado en un documento de referencia de dominio (DRD) firmado antes de iniciar las sesiones de captura.
Fórmula de extracción	Para cada acción $a_k$ en la sesión: clasificar como IN_PHASE (acción pertenece a la fase esperada según el DRD) o OUT_PHASE. $VSP-5 = (N_{IN\_PHASE} / N_{total\_acciones}) \times 100$ .
Umbral orientativo	$VSP-5 > 72\% \rightarrow$ perfil experto (esquema procedimental interiorizado). $VSP-5$ entre 50–72% → perfil ambiguo. $VSP-5 < 50\% \rightarrow$ perfil novato (secuencia sin estructura reconocible).
Advertencia de medición	El Documento de Referencia de Dominio (DRD) debe ser completado y firmado por el Coordinador ANTES de que comience la primera sesión de captura. Modificar el DRD después de ver los datos es una violación de integridad experimental.

## IV.6 Hoja maestra de extracción VSP — completar por sesión

El Coordinador completa una hoja maestra por cada una de las 15 sesiones (5 Generadores × 3 sesiones). Esta hoja es el documento fuente de todos los análisis posteriores.

CAMPO	VALOR EXTRAÍDO	FLAG DE CALIDAD
Código de sesión (GEN-XX × Sesión A/B/C)	_____	—
Código de registro asignado (REG-XXXX)	_____	—
VSP-1: Latencia primer movimiento (segundos)	_____	NORMAL / EXCLUIDO_LECTURA
VSP-2: Ratio D de distribución de retrocesos	_____	NORMAL / BAJA_CONFIANZA
VSP-3: Profundidad promedio de exploración por PDI	_____	NORMAL / K_PDI_BAJO
VSP-4: Tiempo promedio por callejón (segundos)	_____	NORMAL / BAJA_CONFIANZA_VSP4 / N_A
VSP-5: Coherencia estructural (%)	_____	NORMAL / DRD_REFERENCIA
Nivel de incertidumbre de la tarea (1–5, valorado al inicio)	_____	—
Calidad del resultado final (0–10, valorado por Coordinador)	_____	SOLO PARA CORRELACIÓN POST-HOC
Observaciones del Coordinador	_____	—

**V****CRITERIO DE ÉXITO Y CONDICIÓN DE REFUTACIÓN***La matemática de la validación — cómo leo los datos*

Establezco tres métricas independientes que deben ser calculadas sobre los datos del experimento. Cada una responde una pregunta distinta. Las tres juntas constituyen la condición de validación de  $H_1$ . La ausencia de cualquiera de las tres métricas en los resultados hace el análisis incompleto.

**V.1 Métrica 1: Tasa de Acierto Global**

La tasa de acierto global mide si los Jueces en conjunto clasificaron correctamente las sesiones de expertos y novatos. Se calcula sobre el total de clasificaciones realizadas por todos los Jueces sobre todos los registros.

**FÓRMULA:**

$$TA = C_{\text{correctas}} / C_{\text{totales}}$$

**Donde:**

$C_{\text{correctas}}$  = número total de clasificaciones correctas (todos los Jueces, todos los registros)

$C_{\text{totales}} = N_{\text{jueces}} \times N_{\text{registros}} = 5 \times 6 = 30$  clasificaciones totales

UMbral de Éxito:  $TA > 0.70$

**TEST ESTADÍSTICO — Prueba Binomial Unilateral:**

$H_0: p_{\text{real}} \leq 0.50$  (rendimiento no superior al azar)

$H_1: p_{\text{real}} > 0.50$

Valor crítico de  $p$  requerido:  $p < 0.05$

Cálculo:  $P(X \geq k | n=30, p=0.50)$  donde  $k$  = número de aciertos observados

Si p-value calculado < 0.05, rechazo  $H_0$  y concluyo que la tasa es significativamente superior al azar.

## V.2 Métrica 2: Coeficiente Kappa de Cohen ( $\kappa$ )

El Kappa de Cohen mide el acuerdo entre pares de Jueces más allá del acuerdo esperado por azar. Esta métrica responde si los Jueces aciertan por la misma razón o por razones distintas. Un  $\kappa$  alto con una TA alta indica señal robusta y consistente. Una TA alta con  $\kappa$  bajo indica que los Jueces aciertan por heurísticas personales distintas, no por una señal compartida.

### FÓRMULA DE KAPPA DE COHEN (para cada par de Jueces i, j):

$$\kappa(i,j) = (P_o - P_e) / (1 - P_e)$$

Donde:

$P_o$  = proporción de acuerdo observado entre Juez\_i y Juez\_j sobre los 6 registros

$P_e$  = proporción de acuerdo esperado por azar

### Cálculo de $P_e$ (clasificación binaria: EXPERTO / NOVATO):

$$P_e = [p_E(i) \times p_E(j)] + [p_N(i) \times p_N(j)]$$

Donde  $p_E(i)$  = proporción de registros clasificados como EXPERTO por Juez\_i

$$p_N(i) = 1 - p_E(i)$$

### KAPPA GLOBAL DEL EXPERIMENTO:

$\kappa_{global}$  = media aritmética de  $\kappa(i,j)$  para todos los pares posibles

Con 5 Jueces:  $C(5,2) = 10$  pares  $\rightarrow \kappa_{global} = (\sum \kappa(i,j)) / 10$

UMBRAL DE ÉXITO:  $\kappa_{global} > 0.60$

**Tabla de interpretación:**

- $\kappa < 0.20 \rightarrow$  Acuerdo leve. La señal no es robusta.
- $\kappa 0.20\text{--}0.40 \rightarrow$  Acuerdo discreto. Señal débil o contaminada.
- $\kappa 0.40\text{--}0.60 \rightarrow$  Acuerdo moderado. Señal real con ruido significativo.
- $\kappa 0.60\text{--}0.80 \rightarrow$  Acuerdo sustancial. ✓ Umbral CPO-001. Señal robusta.
- $\kappa > 0.80 \rightarrow$  Acuerdo casi perfecto. Resultado excepcional.

**V.3 Métrica 3: Correlación Incertidumbre-Señal**

Esta métrica prueba la hipótesis derivada  $H_1^b$ : que la señal de proceso es más fuerte en sesiones de alta incertidumbre que en sesiones de baja incertidumbre. Se calcula comparando las tasas de acierto de los Jueces sobre registros de sesiones con distinto nivel de incertidumbre.

**PROCEDIMIENTO:**

1. Clasificar los 6 registros en dos grupos según su nivel de incertidumbre:

GRUPO\_BAJA\_INC: Registros de sesiones con nivel 1–2 en la escala de incertidumbre.

GRUPO\_ALTA\_INC: Registros de sesiones con nivel 4–5 en la escala de incertidumbre.

2. Calcular  $TA(\text{baja})$  = tasa de acierto de los Jueces sobre GRUPO\_BAJA\_INC

Calcular  $TA(\text{alta})$  = tasa de acierto de los Jueces sobre GRUPO\_ALTA\_INC

3. Índice de Amplificación por Incertidumbre (IAI):

$$\text{IAI} = TA(\text{alta}) / TA(\text{baja})$$

**INTERPRETACIÓN:**

$\text{IAI} > 1.10 \rightarrow H_1^b$  apoyada: la señal es más fuerte en alta incertidumbre.

$\text{IAI } 0.90\text{--}1.10 \rightarrow$  Sin diferencia. La señal es uniforme por nivel.

$\text{IAI} < 0.90 \rightarrow$  Inverso al esperado. La señal es más fuerte en baja incertidumbre.

NOTA: Esta métrica es secundaria. No invalida  $H_1$  si TA y  $\kappa$  se cumplen.

Solo determina dónde priorizar la captura en el sistema CPO.

#### V.4 Tabla de decisión: qué concluyo según los resultados

ESCENARIO	TA	$\kappa$	IAI	CONCLUSIÓN	ACCIÓN INMEDIATA
<b>VALIDACIÓN COMPLETA</b>	> 70% $p < 0.05$	> 0.60	> 1.10	$H_1, H_1^a$ y $H_1^b$ confirmadas. La señal existe, es robusta y se amplifica en alta incertidumbre.	Implementar Capa 1 VSP automática. Publicar resultado como preprint. Iniciar Fase 1.
<b>VALIDACIÓN PARCIAL A</b>	> 70% $p < 0.05$	> 0.60	0.90–1.10	$H_1$ y $H_1^a$ confirmadas. $H_1^b$ no resuelta. La señal es robusta pero uniforme por nivel.	Implementar Capa 1. Diseñar sesiones de mayor rango de incertidumbre en Fase 2.
<b>VALIDACIÓN PARCIAL B</b>	> 70% $p < 0.05$	< 0.40	—	$H_1$ no confirmada robustamente. Los Jueces aciertan por razones distintas. La señal existe pero no es compartida.	No implementar Capa 1 aún. Analizar VSP por Juez. Replantear el instrumento de captura.
<b>RESULTADO AMBIGUO</b>	60–70%	cualquiera	—	Señal insuficiente para confirmar $H_1$ . El diseño de tarea o el dominio no expone el saber tácito con suficiente fuerza.	Repetir con tarea de mayor complejidad o dominio diferente. Aumentar N de Generadores.
<b>REFUTACIÓN — <math>H_0</math> NO RECHAZADA</b>	≤ 60%	cualquiera	—	$H_0$ no se rechaza. El proceso no contiene señal distingible suficiente en este dominio y diseño de tarea.	Registrar resultado con integridad. Replantear premisa de detección. No abandonar. Reformular.

## V.5 Análisis VSP post-hoc: mapa de indicadores

Independientemente del resultado en TA y  $\kappa$ , ejecuto el análisis VSP post-hoc sobre las hojas de respuesta de los Jueces. Este análisis identifica qué indicadores fueron utilizados con mayor frecuencia y cuál fue la tasa de acierto cuando cada indicador fue determinante. Este análisis es el puente directo entre los datos del experimento y las especificaciones de ingeniería de la Capa 1 automática del sistema CPO.

INDICADOR	FRECUENCIA DE USO POR JUECES	TASA DE ACIERTO CUANDO SE USÓ	ACUERDO ENTRE JUECES QUE LO USARON	¿CANDIDATO A MÉTRICA AUTOMÁTICA?
VSP-1 Latencia primer movimiento	___ / 30 clasificaciones	___ %	$\kappa =$ ___	SÍ / NO / CONDICIONAL
VSP-2 Tasa de retroceso por fase	___ / 30 clasificaciones	___ %	$\kappa =$ ___	SÍ / NO / CONDICIONAL
VSP-3 Profundidad de exploración	___ / 30 clasificaciones	___ %	$\kappa =$ ___	SÍ / NO / CONDICIONAL
VSP-4 Abandono de callejones	___ / 30 clasificaciones	___ %	$\kappa =$ ___	SÍ / NO / CONDICIONAL
VSP-5 Coherencia estructural	___ / 30 clasificaciones	___ %	$\kappa =$ ___	SÍ / NO / CONDICIONAL

## V.6 Tarjetas de estado del experimento — completar al concluir el análisis

TASA DE ACIERTO	<b>%</b> <small>resultado final</small>	UMBRAL <b>&gt; 70%, p &lt; 0.05</b>	PENDIENTE
KAPPA GLOBAL	<b>K =</b> <small>acuerdo inter-juez promedio</small>	UMBRAL <b>&gt; 0.60</b>	PENDIENTE
ÍNDICE IAI	<b> IAI  =</b> <small>amplificación por incertidumbre</small>	UMBRAL <b>&gt; 1.10</b>	PENDIENTE

## DECLARACIÓN FINAL DEL DISEÑADOR:

Yo, Joseph Castillo, declaro que este protocolo está diseñado para refutar mi propia hipótesis con igual

rigor que para confirmarla. He definido con precisión matemática la condición de fracaso antes de ejecutar el experimento. No modificaré los umbrales ni el procedimiento de análisis después de obtener los datos. Los resultados, cualesquiera que sean, se reportarán con integridad completa.

Este manual es el documento de registro oficial del diseño experimental CPO-001.

Cualquier desviación del protocolo durante la ejecución debe ser registrada como

DESVIACIÓN-[número] en el log del Coordinador y reportada en la sección de limitaciones del informe de resultados.

---

## MANUAL DE EJECUCIÓN EMPÍRICA · EXPERIMENTO CPO-001 v2.0

*Joseph Castillo · Arquitecto del experimento · 2025 · Documento de Operaciones de Laboratorio*