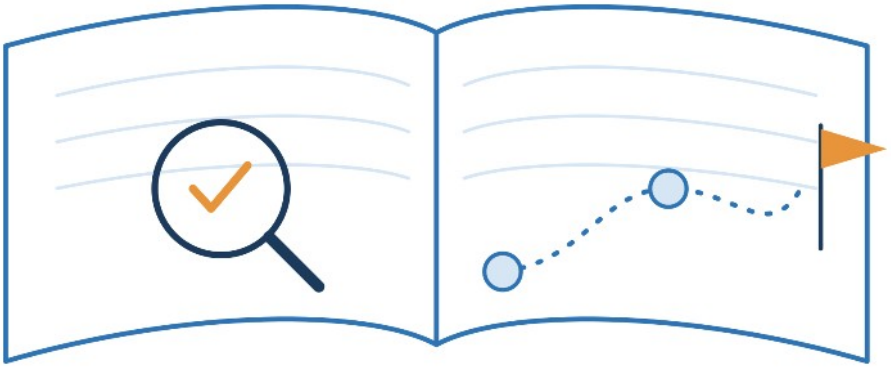
Scrum Manager
Uncovering better ways

Guía

Discovery y validación con clientes

Desarrollo de los temas del State of the art



Actualizado a junio de 2026

Scrum Manager® · Skill Arena

Sobre este documento

Este documento desarrolla en profundidad los temas del mapa de referencia (State of the art) para descubrir y validar con clientes antes de construir, que está disponible en: Skill Arena.

Úsalo como material de consulta para mantener y contrastar tu práctica con el conocimiento profesional actual.

Se complementa con la plataforma de entrenamiento y evaluación en Skill Arena. En el área Customer Discovery puedes realizar pruebas de entrenamiento para contrastar y mejorar tu nivel de conocimiento y si lo deseas también puedes obtener un diploma que acredita curricularmente la solvencia y vanguardia profesional en esta área.

Estado del conocimiento

El conocimiento del discovery y la validación con clientes es un campo relativamente asentado: su núcleo conceptual lleva años probado y evoluciona de forma lenta y progresiva, más por refinamiento de la práctica que por rupturas. Aún así, algunas técnicas concretas siguen ganando forma a medida que su ejecución se generaliza. En los distintos apartados del documento, las etiquetas (ESTABLECIDO, EN CONSOLIDACIÓN, EMERGENTE) ayudan a identificar la madurez de cada concepto:

ESTABLECIDO consenso asentado; conocimiento que se da por necesario.

EN CONSOLIDACIÓN gana adopción con rapidez; aún no universal pero ya relevante.

EMERGENTE frontera reciente; alta relevancia y alta volatilidad.

Índice

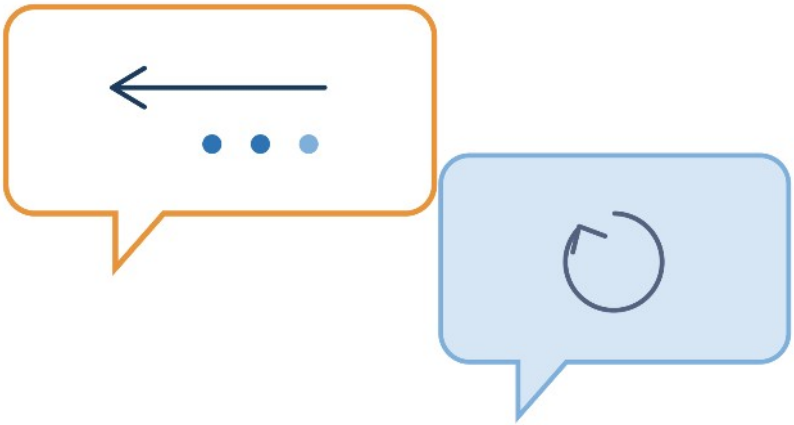
Capítulos del desarrollo, en el orden del State of the art.

BLOQUE A · CONVERSAR Y CAPTURAR SEÑAL

Capítulo 1. La entrevista que aprende, no la que confirma

ESTABLECIDO

Una entrevista de descubrimiento explora el comportamiento pasado y el contexto real del usuario; no es una encuesta oral ni una venta encubierta.



La entrevista útil explora el pasado observable; la que confirma vuelve sobre la propia hipótesis.

Introducción

«Talk to your customers», dice todo el mundo; casi nadie explica cómo hacerlo sin dejarse engañar. Ese es el punto de partida de Rob Fitzpatrick en The Mom Test: hasta tu madre te mentirá para protegerte, así que el problema no es la gente, sino las preguntas que le hacemos. Una entrevista de descubrimiento bien hecha no busca que el cliente apruebe tu idea, sino aprender cómo es su vida, qué hace hoy y dónde se atasca. Este capítulo establece el propósito real de la entrevista y la frontera entre una pregunta que genera señal y una que genera ruido.

1.1 El propósito real: aprender, no validar

Una entrevista de descubrimiento no es una encuesta oral ni una sesión de venta disfrazada. Su finalidad es explorar el comportamiento pasado y el contexto real del usuario para entender sus necesidades genuinas. La diferencia entre una entrevista útil y una inútil está en una pregunta previa: ¿buscas aprender algo nuevo o confirmar lo que ya crees? Antes de redactar cualquier guion conviene poder completar la frase «queremos aprender X y lo sabremos si observamos Y». Sin esa claridad, la conversación se vuelve difusa y los datos no son accionables.

La regla temporal. Las preguntas que generan señal apuntan al pasado observable —«cuéntame la última vez que tuviste este problema»—; las que generan ruido apuntan a futuros hipotéticos —«¿usarías esto?»— o a valoraciones abstractas —«¿qué opinas de mi idea?»—. El pasado se recuerda; el futuro se inventa por cortesía.

1.2 Preguntas de comportamiento frente a preguntas de opinión

Las preguntas que generan evidencia útil se centran en comportamiento pasado específico, no en intenciones futuras ni preferencias. «¿Qué intentaste para resolverlo y qué pasó?» o «muéstrame cómo lo haces ahora» revelan hechos; «¿te gustaría tener esta función?» o «¿no sería mejor centralizarlo todo?» empujan a respuestas de cortesía, piden estimaciones especulativas o sugieren la respuesta. Pedir opiniones sobre el futuro confunde, además, descubrimiento con venta.

1.3 La trampa del sesgo de confirmación

Cuando el equipo ya tiene una solución en mente, surge la tendencia a diseñar preguntas que la confirmen. Un objetivo como «validar si los usuarios quieren la nueva pantalla de alta» lleva la solución embebida y empuja hacia respuestas que la apoyen. Las señales de alerta son claras: el objetivo nombra una funcionalidad concreta, usa verbos como «validar», «confirmar» o «demostrar» respecto de una solución, o define el éxito como un número mínimo de respuestas positivas. La corrección es reformular hacia la exploración: «entender cómo se registran hoy, dónde se atascan y qué intentan conseguir».

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Formular el objetivo de aprendizaje de una entrevista como «queremos aprender X y lo sabremos si observamos Y», antes de redactar preguntas.
- Distinguir una pregunta de comportamiento pasado de una de opinión o intención futura, y reescribir la segunda en términos de la primera.
- Detectar en un objetivo o guion las señales de sesgo de confirmación (solución nombrada, verbos de validación, éxito definido como respuestas positivas).
- Reformular un objetivo cargado de solución hacia un objetivo de exploración abierta.
- Reconocer cuándo una entrevista ha mezclado descubrimiento con venta y separar ambas conversaciones.

Errores y antipatrones frecuentes

Preguntar «¿te gustaría X?». Parece directo, pero pide una predicción de cortesía sin valor. Reformula hacia el pasado: «¿cuándo fue la última vez que…?».

Mezclar venta con descubrimiento. La presión por demostrar el valor del producto contamina la señal. En discovery no se presentan soluciones; se aprende del problema.

Objetivo vago («recoger feedback»). Sin una hipótesis clara, los datos no son accionables. Define qué supuesto exploras y qué evidencia lo validaría.

Salir sin aprendizajes accionables. Preguntas demasiado genéricas producen respuestas vagas. Prepara preguntas de seguimiento para profundizar en cada fricción.

Para profundizar

· Rob Fitzpatrick — The Mom Test

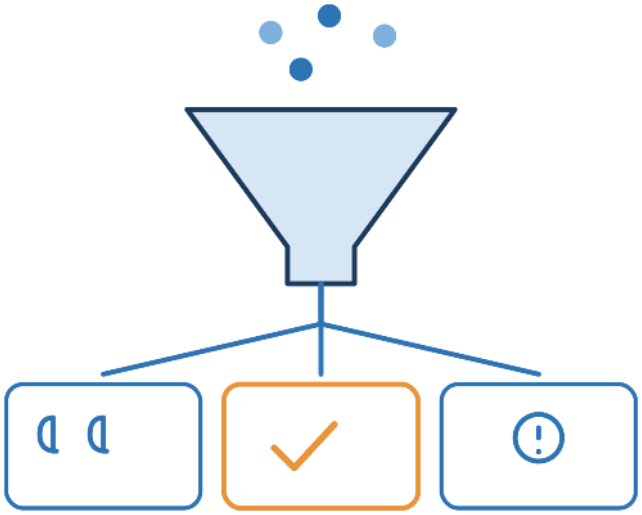
· Teresa Torres — Customer interviews

BLOQUE A · CONVERSAR Y CAPTURAR SEÑAL

Capítulo 2. Notas de calidad: separar hecho, cita e interpretación

ESTABLECIDO

La síntesis posterior solo es tan buena como las notas que la alimentan; separar dato e interpretación desde la captura es lo que la hace posible.



Capturar por separado cita, hecho e interpretación evita contaminar los datos desde el inicio.

Introducción

Una entrevista excelente con notas pobres equivale a una entrevista perdida. La memoria de los detalles cualitativos decae en horas, y lo que no quedó registrado con precisión se reconstruye luego con sesgo. Este capítulo desarrolla cómo capturar señal de calidad durante y después de la conversación, de modo que la síntesis del capítulo 3 parta de material fiable y no de impresiones.

2.1 La jerarquía de la captura

Durante la entrevista se captura en un orden de prioridad: primero citas textuales —palabras exactas entre comillas, el dato más objetivo—; después contexto específico —quién, cuándo, en qué situación—, que permite segmentar y priorizar; luego interpretaciones, claramente marcadas como inferencias propias; y por último las incertidumbres pendientes que requieren más exploración. Mezclar estos niveles desde el inicio contamina los datos originales y hace inviable una síntesis rigurosa.

La prueba del adjetivo. Si tus notas suenan a adjetivos —«le gustó», «buena actitud», «contento con los tiempos»— probablemente falta evidencia concreta. «Paso dos horas al día en reuniones de seguimiento» es señal; «está contento» es una valoración sin sustento verificable.

2.2 Separar observación de interpretación

Anotar interpretaciones es legítimo y útil, siempre que estén marcadas como tales. Un sistema de etiquetas sencillo funciona bien: una marca para la cita textual, otra para la observación factual (qué hizo, qué herramientas usó), otra para la interpretación sobre significado o emoción, y otra para la incertidumbre que requiere seguimiento. Así, al sintetizar, se distingue lo que el cliente dijo e hizo de lo que el entrevistador supuso, y se evita tomar la inferencia por dato.

2.3 El procesamiento posterior inmediato

La memoria de los matices cualitativos se evapora deprisa. Reservar quince o veinte minutos justo después de cada entrevista para completar las notas, separar hechos de interpretaciones, identificar los tres a cinco puntos más relevantes para la hipótesis y anotar las preguntas de seguimiento es la práctica que sostiene todo el análisis. Esperar días para cerrar las notas equivale a perder información que ya no se podrá recuperar.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Ordenar la captura por prioridad: cita textual, contexto específico, interpretación marcada e incertidumbre pendiente.
- Clasificar una nota concreta como cita, observación factual o interpretación, y reescribir una interpretación encubierta de dato.
- Reconocer una nota vaga (basada en adjetivos) y convertirla en una observación verificable con contexto.
- Aplicar un sistema de etiquetas que separe dato de inferencia a lo largo de una entrevista.
- Cerrar el procesamiento posterior de una entrevista en el plazo en que la memoria conserva los matices.

Errores y antipatrones frecuentes

Notas vagas sin contexto. Sin quién, cuándo y en qué situación, no se pueden derivar insights accionables ni segmentar después.

Mezclar datos con interpretaciones que parecen datos. Produce conclusiones no verificables. Marca explícitamente cita, observación e interpretación.

Registrar solo lo que confirma la hipótesis. Es sesgo de confirmación que invalida el discovery. Pregúntate: «¿anoté algo que no me gustó escuchar?».

Anotar solo conclusiones para «ahorrar tiempo». Se pierde la evidencia y se vuelve imposible revisar. Captura primero el dato crudo; concluye después.

Para profundizar

· Teresa Torres — Customer interviews y síntesis

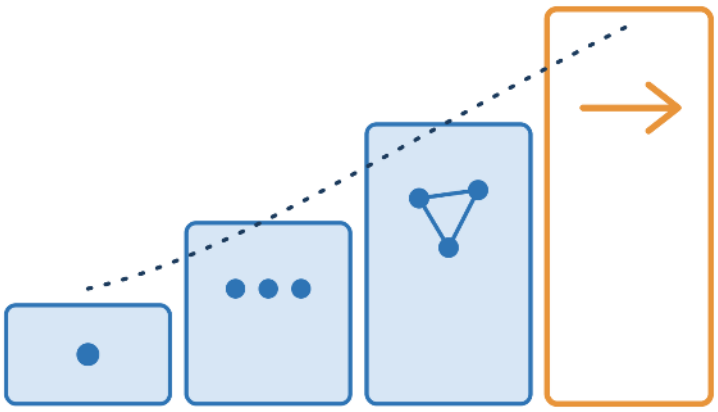
· Rob Fitzpatrick — The Mom Test

BLOQUE B · DE LOS DATOS A LA COMPRENSIÓN

Capítulo 3. Síntesis: de la observación al insight accionable

ESTABLECIDO

El análisis cualitativo asciende de la observación al patrón, de este al insight y del insight a la implicación; confundir los niveles desvía decisiones.



La síntesis asciende de la observación al patrón, al insight y a la implicación accionable.

Introducción

Tener muchas conversaciones no es lo mismo que aprender de ellas. La síntesis es el trabajo de convertir notas dispersas en comprensión accionable, y es donde se concentra buena parte del valor diferencial del profesional. Este capítulo desarrolla la jerarquía del análisis cualitativo y la estructura de un insight que de verdad guíe una decisión de producto.

3.1 Los cuatro niveles del análisis

El análisis cualitativo asciende por niveles que se sostienen unos en otros. La observación es un dato de una fuente individual («Ana tarda treinta minutos buscando documentos»). El patrón es la recurrencia detectada en varias fuentes («seis de diez entrevistados mencionan fricción con la búsqueda»). El insight es la comprensión nueva que conecta el patrón con una causa o motivación («pierden tiempo porque la estructura de carpetas no refleja cómo piensan los proyectos»). Y la implicación es la acción que se deriva («explorar una navegación por proyectos además de carpetas»). El error más común es llamar insight a una cita interesante o a un patrón sin interpretar.

La prueba del insight. Un insight válido responde siempre: «¿qué aprendimos que no sabíamos y qué significa para el producto?». Si no hay causa ni implicación, todavía no es un insight: es una observación o, como mucho, un patrón.

3.2 De citas a patrones: agrupar y cuantificar

El primer paso de la síntesis es agrupar observaciones similares por temas emergentes: revisar todas las notas sin categorizar de entrada, agrupar las que tratan del mismo fenómeno, contar cuántas fuentes distintas mencionan cada tema y ordenar por frecuencia y relevancia. La cuantificación es crucial: «siete de doce entrevistados mencionaron fricción con el tiempo de espera» es mucho más útil que «varios usuarios mencionaron problemas de velocidad». La regla práctica es buscar el patrón en al menos tres a cinco fuentes antes de afirmarlo.

3.3 Las excepciones como señal de segmento

Cuando la mayoría coincide pero algunos casos no encajan, descartarlos como ruido es un error. Conviene analizar qué tienen en común los casos diferentes: ¿otro rol, otro uso, otro contexto organizativo? Las excepciones suelen revelar segmentos con necesidades distintas o casos de uso no considerados. La pregunta clave no es «¿cómo ignoro a los que no encajan?», sino «¿qué tienen en común entre sí estos casos que no encajan?».

3.4 La estructura de un insight bien formulado

Un insight eficaz sigue una estructura: «observamos que [patrón cuantificado] porque [interpretación de causa]; esto implica que deberíamos [acción concreta]». «Observamos que cinco de ocho usuarios mencionan fricción con contenido demasiado largo porque sus jornadas fragmentadas no permiten sesiones extensas; deberíamos evaluar un formato de microlearning» es accionable. «Los usuarios necesitan una experiencia más personalizada» es genérico: ni patrón cuantificado, ni causa, ni implicación.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Clasificar una afirmación en uno de los cuatro niveles: observación, patrón, insight o implicación.
- Distinguir un insight válido de una cita interesante o de un patrón sin interpretar, aplicando la prueba «qué aprendimos y qué significa».
- Agrupar observaciones por tema y cuantificar la recurrencia en número de fuentes distintas, exigiendo un mínimo razonable antes de afirmar un patrón.
- Tratar los casos que no encajan como posible señal de segmento, buscando qué tienen en común entre sí.
- Formular un insight con la estructura patrón cuantificado + causa + implicación accionable.

Errores y antipatrones frecuentes

Lista de citas sin agrupar. No permite ver patrones ni priorizar. Agrupa por temas y cuantifica la recurrencia.

Generalizar desde un caso aislado. Una sola voz puede priorizar la necesidad de una minoría. Verifica el patrón en varias fuentes antes de afirmarlo.

Insight sin implicación práctica. Si no guía una decisión, no es accionable. Añade el «esto significa que deberíamos…».

Ignorar los casos contradictorios. Se pierde información sobre posibles segmentos. Documenta las excepciones y explora qué las diferencia.

Para profundizar

· Teresa Torres — Product discovery

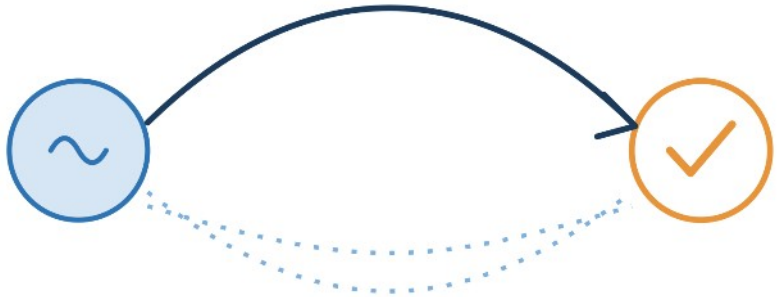
· Scrum Manager — Propietario del producto

BLOQUE B · DE LOS DATOS A LA COMPRENSIÓN

Capítulo 4. Enmarcar el problema con Jobs-to-be-Done

ESTABLECIDO

Definir el problema como «falta de funcionalidad» esconde la solución dentro del problema; el enmarcado JTBD describe el progreso buscado con independencia de la solución.



JTBD describe el progreso que busca el usuario, dejando abiertas varias soluciones posibles.

Introducción

El modo en que se enuncia un problema condiciona todas las soluciones que se podrán imaginar después. El error más caro y frecuente en producto es definir el problema como «falta de X funcionalidad», porque cierra la exploración antes de empezar. Este capítulo desarrolla el enmarcado de problemas con el marco Jobs-to-be-Done y la disciplina de bajar del síntoma a la causa.

4.1 La trampa de definir el problema como ausencia de solución

Cuando llega una petición como «necesitamos exportación a PDF», la tendencia es asumir que el problema es la ausencia de esa función. Pero «los usuarios no tienen un botón de exportar a PDF» embebe la solución y cierra la indagación. Las preguntas correctas son otras: ¿para qué necesitan el PDF?, ¿a quién se lo envían?, ¿qué hacen con él después? Quizá lo que necesitan es compartir datos con terceros, y un enlace compartible sea mejor que un PDF. La reformulación canónica: «el usuario intenta lograr [objetivo] pero actualmente [fricción] porque [causa]».

4.2 La estructura del Job-to-be-Done

Un JTBD bien formulado describe el progreso que busca el usuario con independencia de la solución: «cuando [situación específica], quiero [progreso o motivación], para [resultado esperado]», y documenta las alternativas actuales. «Cuando coordino entregables entre tres equipos remotos, quiero visibilidad del estado real de cada tarea, para anticipar bloqueos antes del deadline; hoy uso email más una hoja compartida» es específico y observable. «Cuando necesito gestionar mi trabajo, quiero una herramienta fácil, para ser más productivo» es genérico y circular.

Los cuatro componentes. Situación (el disparador concreto, no «cuando trabajo»), motivación o progreso (el cambio de estado buscado, no la funcionalidad), resultado final (el beneficio último) y alternativas actuales (cómo lo resuelve hoy, que define el estándar a superar y revela las restricciones reales).

4.3 Del síntoma a la causa raíz

El enmarcado correcto exige excavar bajo el síntoma visible. Se identifica el síntoma que el usuario

reporta («el sistema es lento»), se exploran las causas con preguntas de seguimiento («¿qué intentabas hacer cuando lo notaste lento?»), se define la necesidad en términos de resultado («encontrar información concreta rápido para decidir») y se documenta el contexto. A menudo «lento» revela un problema de findability —encontrar—, no de rendimiento técnico: optimizar la base de datos no resolvería una frustración de búsqueda.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Detectar cuándo un problema está enunciado como «falta de funcionalidad» y reescribirlo sin prescribir la solución.
- Formular un Job-to-be-Done con la estructura situación + motivación + resultado, evitando el enunciado genérico y circular.
- Identificar los cuatro componentes del enmarcado, incluida la alternativa actual que el usuario ya emplea.
- Separar síntoma, causa y necesidad subyacente en una petición de usuario, llegando a la causa raíz mediante preguntas de seguimiento.
- Comprobar que un enmarcado deja abiertas varias soluciones posibles en lugar de una sola.

Errores y antipatrones frecuentes

Problema como falta de feature. «No tienen exportación a PDF» prescribe la solución. Reformula hacia el objetivo: «necesitan compartir datos con externos de forma consultable».

JTBD circular. «Quiero gestionar proyectos para ser más productivo» no dice nada. Especifica situación concreta y resultado observable.

Ignorar las alternativas actuales. Siempre hay una: email, Excel, papel, o ignorar el problema. Documentarla revela el estándar real a superar.

Saltar a la solución. «El problema es que necesitan un dashboard» ya decidió. Describe la fricción sin prescribir cómo resolverla.

Para profundizar

· Strategyn — Jobs-to-be-Done

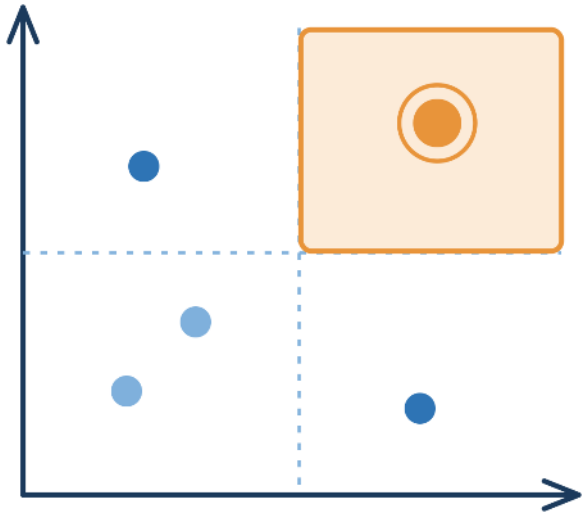
· Teresa Torres — Opportunity solution trees

BLOQUE C · DISEÑAR EXPERIMENTOS PARA APRENDER

Capítulo 5. Mapeo de supuestos: deseabilidad, viabilidad, factibilidad

ESTABLECIDO

Todo proyecto descansa sobre supuestos de tres tipos; priorizar por impacto e incertidumbre evita validar lo fácil y dejar sin probar lo crítico.



Se ataca primero el supuesto de alto impacto y alta incertidumbre, aunque sea incómodo de probar.

Introducción

Antes de diseñar cualquier experimento conviene saber qué hay que probar y en qué orden. El mapeo de supuestos, formalizado por David Bland y Alexander Osterwalder en Testing Business Ideas, es la herramienta que hace explícitos los supuestos de un proyecto y los prioriza según lo que de verdad pone en riesgo la iniciativa. Este capítulo desarrolla los tres tipos de supuesto y la matriz de priorización.

5.1 Los tres tipos de supuesto

Todo proyecto nuevo descansa sobre supuestos de tres categorías. La deseabilidad responde «¿lo quieren?»: el problema es real y prioritario, los usuarios cambiarán de su herramienta actual, están dispuestos a pagar el precio. La viabilidad responde «¿es negocio sostenible?»: el coste de adquisición es menor que el valor de vida del cliente, los márgenes son positivos. La factibilidad responde «¿podemos construirlo?»: el equipo tiene las capacidades técnicas, la infraestructura soporta la carga, se puede integrar con lo existente.

5.2 El error de priorizar lo fácil sobre lo crítico

Un patrón habitual es priorizar los supuestos técnicos (factibilidad) porque son cómodos de probar internamente, dejando sin explorar los de mercado. Pero si nadie pagará el precio, la capacidad de construirlo es irrelevante. La secuencia correcta para la mayoría de productos nuevos es deseabilidad primero, viabilidad después y factibilidad al final. Invertir en factibilidad antes de validar la demanda lleva a un producto técnicamente sólido que nadie quiere.

Los supuestos del liderazgo. Los más peligrosos suelen ser los que nadie cuestiona porque vienen de la dirección: «los usuarios cambiarán sin resistencia», «el boca a boca bastará para crecer». Incorporarlos al mapa los hace visibles y evaluables; evitarlos por jerarquía perpetúa puntos ciegos

que pueden hundir el proyecto.

5.3 La matriz de impacto e incertidumbre

Los supuestos se priorizan en dos dimensiones. El impacto si es falso mide la gravedad para el proyecto: alto si la iniciativa no tiene sentido sin él (la disposición a pagar), medio si exige ajustes, bajo si es solucionable con cambios menores. La incertidumbre mide cuán seguros estamos: alta sin datos ni experiencia, baja con evidencia que lo respalde. Se ataca primero el cuadrante de alto impacto y alta incertidumbre, aunque sea incómodo de testear, porque es donde un error sale más caro.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Clasificar un supuesto como de deseabilidad, viabilidad o factibilidad.
- Reformular un supuesto vago en un enunciado testeable y preciso.
- Ubicar un supuesto en la matriz de impacto-si-es-falso por incertidumbre y justificar su prioridad.
- Defender la secuencia deseabilidad → viabilidad → factibilidad frente al reflejo de validar primero lo técnico.
- Hacer explícitos los supuestos implícitos del liderazgo o de la cultura del equipo para incorporarlos al mapa.

Errores y antipatrones frecuentes

Validar lo fácil, no lo crítico. Los supuestos letales quedan sin probar. Usa la matriz impacto × incertidumbre para decidir el orden.

Supuestos implícitos nunca listados. Producen puntos ciegos que hunden el proyecto. Haz una sesión explícita de mapeo con todos los roles.

Prioridad invertida. Probar factibilidad antes que deseabilidad lleva a construir algo que nadie quiere. Sigue la secuencia correcta.

Supuestos vagos no testeables. Lo que no se puede validar ni invalidar no sirve. Reformula con criterios medibles.

Para profundizar

· David Bland — Assumptions Mapping (Strategyzer)

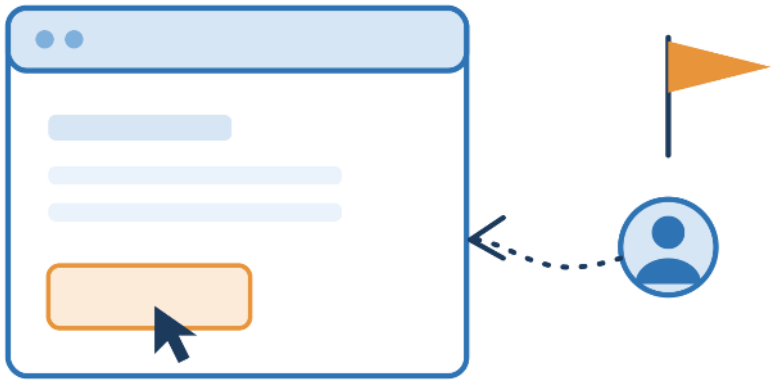
· Bland & Osterwalder — Testing Business Ideas

BLOQUE C · DISEÑAR EXPERIMENTOS PARA APRENDER

Capítulo 6. Tests de prototipo: tareas como objetivos, no instrucciones

EN CONSOLIDACIÓN

Un test de prototipo genera señal de usabilidad solo si las tareas son objetivos, el moderador es neutral y los participantes son externos y representativos.



El test da señal solo si la tarea es un objetivo y el usuario, externo, la persigue sin guía.

Introducción

Probar un prototipo con usuarios parece sencillo, pero la mayoría de los tests producen señal contaminada por defectos de diseño del propio test. Este capítulo desarrolla las cuatro decisiones que separan un test que mide el diseño de uno que mide otra cosa: cómo se plantean las tareas, quién modera, a quién se recluta y qué se observa. La teoría está asentada; su ejecución disciplinada todavía no es universal, de ahí su estado en consolidación.

6.1 Tareas como objetivos, no como instrucciones

El error más común es diseñar tareas que indican cómo completarlas, eliminando la señal de descubribilidad. «Pulsa el botón verde de arriba a la derecha, rellena el formulario y guarda con el icono de disquete» evalúa si el usuario sigue instrucciones. «Crea un proyecto llamado Prueba Q1» revela si el diseño es intuitivo, dónde se confunde y qué modelo mental aplica. La regla es plantear la tarea como objetivo, sin indicar el método, y observar el camino que toma.

6.2 Moderador neutral y participantes representativos

Cuando el diseñador del prototipo modera, introduce sesgos sutiles: interés emocional en el éxito, guía inconsciente hacia el «camino correcto», interpretación que atribuye la confusión al usuario y no al diseño, y ayuda excesiva que elimina señal. La solución es separar roles: un moderador neutral conduce y el creador observa sin intervenir. Igual de crítico es reclutar participantes externos y representativos del segmento: probar con compañeros o amigos —que conocen el contexto y tienen incentivos para ser amables— compromete la señal de raíz.

Observar comportamiento, no pedir opiniones. El criterio de éxito no es «que diga que sí», sino métricas observables: dónde hace clic primero, dónde se detiene, cuánto tarda, qué elementos no encuentra. «¿Te gusta el diseño?» pertenece a la venta, no al test.

6.3 La fidelidad se ajusta a la pregunta

El nivel de fidelidad del prototipo debe corresponder al objetivo del test. Para validar si un concepto

se entiende, bastan bocetos o wireframes; para explorar un flujo de navegación, un prototipo clicable de fidelidad media; para evaluar la usabilidad de interacciones concretas, alta fidelidad; para testear la propuesta de valor, una landing page. Construir alta fidelidad para todo desperdicia tiempo: usa la fidelidad mínima que responda a tu pregunta de investigación.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Reescribir una tarea de test redactada como instrucción paso a paso en una tarea planteada como objetivo.
- Justificar por qué el moderador no debe ser quien diseñó el prototipo y asignar correctamente los roles de moderación y observación.
- Distinguir un participante representativo del segmento de uno conveniente, y explicar cómo el segundo compromete la señal.
- Definir criterios de éxito observables (clics, pausas, tiempo, elementos no encontrados) en lugar de preguntas de opinión.
- Elegir el nivel de fidelidad del prototipo adecuado al objetivo del test.

Errores y antipatrones frecuentes

Tareas guiadas. Si las instrucciones mencionan ubicaciones o elementos concretos, se mide obediencia, no usabilidad. Reformula como objetivo.

El moderador es el creador. «Modero yo porque conozco mejor el diseño» introduce sesgo. Asigna un moderador neutral; el diseñador observa.

Participantes internos. Compañeros y amigos conocen el contexto y quieren agradar. Recluta usuarios externos del segmento objetivo.

Pedir opiniones. «¿Te gusta?» o «¿qué te parece?» no son señal de usabilidad. Diseña tareas que generen comportamiento observable.

Para profundizar

· Steve Krug — usability testing

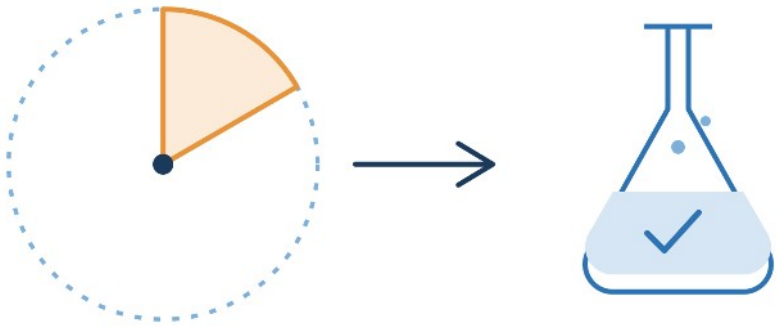
· David Bland — Experiment library (Strategyzer)

BLOQUE C · DISEÑAR EXPERIMENTOS PARA APRENDER

Capítulo 7. El MVP como experimento, no como producto reducido

EN CONSOLIDACIÓN

Un MVP real es el mínimo necesario para aprender algo específico, no una versión pequeña de todo; a menudo se valida mejor sin escribir código.



El MVP es lo mínimo necesario para aprender algo concreto, no una versión pequeña de todo.

Introducción

Pocos términos de producto se malinterpretan tanto como MVP. Entendido como «producto con menos funcionalidades», produce backlogs disfrazados de experimento. Entendido como instrumento de aprendizaje, ahorra meses de desarrollo. Este capítulo desarrolla el MVP como experimento, las alternativas no-producto y por qué «no escala» es un argumento irrelevante en etapa temprana. La práctica gana adopción pero aún convive con el reflejo de construir primero, de ahí su estado en consolidación.

7.1 MVP como experimento, no como producto reducido

Un MVP real es el mínimo necesario para aprender algo específico, no el mínimo para tener un producto. «Incluye registro, perfil, notificaciones, dashboard e integraciones; hipótesis: los usuarios usarán la plataforma; criterio: publicar en seis semanas» es un backlog de componentes con una hipótesis genérica y sin criterio de validación. «Hipótesis: los equipos de finanzas pagarán el precio por automatizar la reconciliación; MVP: servicio concierge manual para cinco clientes durante dos semanas; criterio: tres de cinco piden continuar y pagar» es un experimento. El primero construye para «ver qué pasa»; el segundo aprende antes de construir.

7.2 Alternativas no-producto

Antes de escribir código conviene preguntarse si la hipótesis se puede validar con algo más ligero. El MVP concierge entrega el servicio manualmente, a la vista del usuario, y valida la demanda real y la disposición a pagar antes de automatizar. El Mago de Oz simula un sistema automático con operación humana oculta y valida la experiencia sin invertir en desarrollo. La landing page describe la propuesta y captura intención —email, reserva, prepago— antes de construir nada. El single feature es el mínimo código para testear una hipótesis concreta. La pregunta guía es «¿cuál es lo mínimo para aprender X?», no «¿cuál es el mínimo producto funcional?».

El falso problema de «no escala». Argumentar contra las alternativas manuales porque «no escalan» desperdicia la oportunidad de aprender rápido y barato. Si nadie quiere el servicio manual, no hay nada que escalar; si lo quieren, habrás validado antes de invertir en automatización.

7.3 Elegir el tipo según la hipótesis

Cada hipótesis pide un tipo de MVP. Para «¿hay interés inicial?», una landing con captura de email. Para «¿pagarían por esto?», un concierge con servicio manual. Para «¿la experiencia propuesta resuelve el problema?», un Mago de Oz. Para «¿esta funcionalidad concreta aporta valor?», un single feature con el mínimo código. Elegir el instrumento según lo que se quiere aprender —y no por costumbre— es lo que mantiene el MVP como experimento y no como primera entrega.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Distinguir un MVP entendido como experimento de aprendizaje de uno entendido como producto reducido o lista de funcionalidades.
- Formular la hipótesis principal de un MVP y la señal concreta que la validaría o invalidaría, antes de construir.
- Elegir entre concierge, Mago de Oz, landing page o single feature según la hipótesis que se quiere validar.
- Rebatir el argumento de «no escala» en etapa temprana, explicando por qué la escalabilidad es entonces irrelevante.
- Eliminar de un MVP los elementos que no contribuyen al aprendizaje buscado.

Errores y antipatrones frecuentes

MVP como lista de funcionalidades. Provoca scope creep y pérdida de velocidad. Define primero la hipótesis y luego el mínimo para probarla.

Sin hipótesis definida. Sin ella no se pueden interpretar los resultados. Formula «creeremos que X si observamos Y».

Sin criterio de éxito previo. Permite racionalizar cualquier resultado a posteriori. Fija los umbrales antes de lanzar.

Descartar alternativas manuales. Lleva a invertir cuando un test manual habría bastado. Pregunta siempre «¿podemos validar esto sin código?».

Para profundizar

· Bland & Osterwalder — Testing Business Ideas

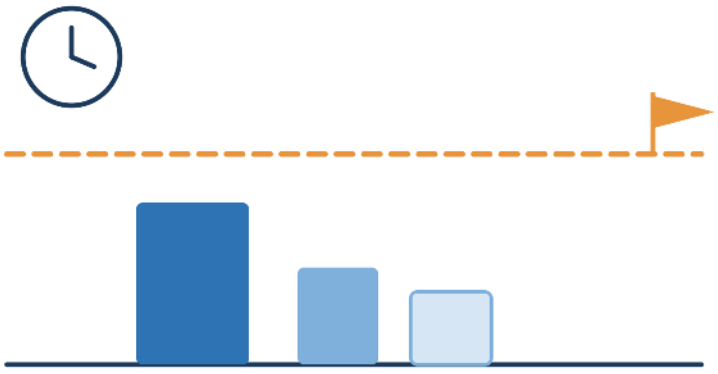
· Teresa Torres — Assumption testing

BLOQUE D · DECIDIR CON EVIDENCIA

Capítulo 8. Criterios de éxito y señales definidos de antemano

ESTABLECIDO

El criterio de éxito se fija antes de lanzar el experimento; decidir después «qué nos parece bueno» es racionalizar cualquier resultado.



El umbral de éxito se fija antes de lanzar, no después de ver los resultados.

Introducción

Un experimento sin criterio previo de éxito no es un experimento: es una historia que se contará a conveniencia cuando lleguen los datos. Este capítulo desarrolla el principio de definición previa, la jerarquía de métricas y los límites que impone el tamaño de muestra, que es donde el discovery se vuelve riguroso o deja de serlo.

8.1 El principio de definición previa

El criterio de éxito debe definirse antes de lanzar, no después de ver los resultados. Decidir a posteriori «qué nos parece bueno» es una forma informal de p-hacking que permite presentar casi cualquier resultado como éxito. «Mediremos visitas, tiempo en página, clics y conversiones, y decidiremos luego qué nos parece bueno» no es criterio. «Métrica primaria: conversión ≥5 % con al menos 200 visitas; guardrail: rebote <70 %; definido el 15 de enero, antes del lanzamiento» sí lo es, porque no deja espacio para ajustar el listón según convenga.

8.2 La jerarquía de métricas

Un experimento bien diseñado distingue funciones. La métrica primaria determina éxito o fracaso: debe ser una sola, con umbral específico. Las métricas secundarias aportan contexto sin cambiar la decisión. Los guardrails son umbrales que no deben cruzarse aunque la primaria salga bien, para detectar daño colateral. Y las métricas de diagnóstico ayudan a explicar el porqué de cara a la iteración. La confusión típica —tener varias «métricas primarias»— diluye el criterio y habilita el cherry-picking posterior.

Tamaño de muestra y efecto detectable. Con muestras pequeñas, las fluctuaciones normales parecen señales. Antes de lanzar conviene calcular qué tamaño de efecto se puede detectar de forma fiable con la muestra esperada; un umbral de «cualquier mejora» con n=50 confunde ruido con señal.

8.3 El problema del grupo de control

Sin grupo de control no se puede atribuir un cambio a la intervención: un aumento de sesiones tras

activar una función podría deberse a estacionalidad, otras campañas o factores externos. El diseño mínimo para atribuir causalidad incluye un grupo de tratamiento, un grupo de control equivalente y aleatorización en la asignación. Sin esos elementos, los resultados son correlacionales, no causales, y conviene decirlo así.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Definir el criterio de éxito de un experimento —métrica primaria, umbral numérico y fecha— antes de lanzarlo.
- Clasificar una métrica como primaria, secundaria, guardrail o de diagnóstico, y exigir una única métrica primaria.
- Reconocer cuándo el tamaño de muestra no permite detectar el efecto buscado y ajustar las expectativas en consecuencia.
- Identificar la ausencia de grupo de control y el salto injustificado de correlación a causalidad.
- Detectar el p-hacking informal de fijar o mover el umbral después de ver los datos.

Errores y antipatrones frecuentes

Definir el umbral tras ver resultados. Es p-hacking que justifica cualquier cosa. Documenta los criterios con fecha anterior al lanzamiento.

Medir muchas métricas sin priorizar. Permite elegir la favorable a posteriori. Define una única métrica primaria clara.

Umbral de «cualquier mejora». Confunde ruido con señal. Calcula el efecto mínimo detectable dada la muestra.

Sin grupo de control. Lleva a conclusiones de causalidad injustificadas. Diseña con grupo de control cuando sea posible.

Para profundizar

· Teresa Torres — Assumption testing

· Bland & Osterwalder — Testing Business Ideas

BLOQUE D · DECIDIR CON EVIDENCIA

Capítulo 9. Calidad de la evidencia y control de sesgos

ESTABLECIDO

No toda evidencia pesa igual; cuatro sesgos la erosionan y varias señales de bajo coste se confunden con validación.

El comportamiento con compromiso real pesa más que las señales de bajo coste.

Introducción

Recoger datos no equivale a tener evidencia. Este capítulo desarrolla cómo ponderar la calidad de la evidencia, los cuatro sesgos que más la erosionan, las señales que parecen validación pero no lo son y la frontera práctica entre correlación y causalidad. Es el filtro que separa el discovery riguroso del teatro de validación.

9.1 Los cuatro sesgos críticos

Cuatro sesgos comprometen la evidencia. El de confirmación lleva a ver solo los datos que apoyan la creencia previa; se mitiga buscando activamente lo que la contradice. El de supervivencia analiza solo los casos exitosos —encuestar solo a quien completó el proceso—; se mitiga incluyendo a quienes abandonaron o fracasaron. El de selección usa una muestra no representativa —reclutar de la propia red—; se mitiga definiendo criterios de selección antes de reclutar. Y el de deseabilidad social produce respuestas que buscan agradar —«lo usaría», «lo pagaría»—; se mitiga preguntando por comportamiento pasado: «¿cuánto pagaste la última vez?» en lugar de «¿cuánto pagarías?».

9.2 Señales de evidencia débil

No toda señal valida. Los likes y comentarios en redes son de bajo coste y no implican uso ni disposición a pagar: 47 likes en un post no validan demanda. El feedback de la red propia —amigos, contactos, suscriptores— viene con incentivo social para ser positivo. Las aperturas de email indican buenos suscriptores, no demanda del nuevo producto. La evidencia sólida es otra: reservas con pago, registros de desconocidos, cartas de intención de clientes fuera de tu red, comportamiento observable y no declarado.

La pregunta de calidad. Ante cualquier dato que se presente como validación, conviene preguntar: «¿esto es comportamiento observable con coste para quien lo hace, o es una opinión de bajo coste?». La diferencia decide cuánto peso merece.

9.3 Correlación frente a causalidad

En producto es fácil confundir correlación con causalidad. «Tras añadir la función subieron las conversiones» no implica que la función las causara; «los usuarios que usan X retienen más» no implica que X cause retención —quizá los más comprometidos simplemente usan más funciones—. Para afirmar causalidad hacen falta grupo de control, aleatorización y control de factores de confusión. Sin ellos, lo honesto es decir «observamos correlación», que no es lo mismo que «X causa Y».

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Identificar cuál de los cuatro sesgos —confirmación, supervivencia, selección, deseabilidad social— afecta a una muestra o a un dato, y proponer su mitigación.
- Distinguir una señal de validación débil (likes, feedback de la red, aperturas) de una señal sólida (pago, registro de desconocidos, carta de intención, comportamiento observable).
- Reformular una pregunta de intención futura en una de comportamiento pasado para reducir la deseabilidad social.
- Diferenciar correlación de causalidad en un resultado concreto y reconocer cuándo falta un grupo de control.
- Pedir evidencia adicional cuando la disponible es insuficiente o está comprometida por la forma en que se obtuvo.

Errores y antipatrones frecuentes

Tomar likes como validación. «47 likes = demanda» confunde ruido con señal. Busca compromiso real: pago, tiempo invertido.

Seleccionar participantes sesgados. Encuestar solo a quien completó el onboarding excluye a quien abandonó. Incluye a los que fracasaron.

Ignorar señales contradictorias. Tratar a dos usuarios negativos como outliers pierde información. Investiga qué los diferencia.

Concluir causalidad sin control. «Tras X pasó Y, luego X causó Y» es un salto injustificado. Reconoce los límites: «observamos correlación».

Para profundizar

· David Bland — Strength of evidence (Strategyzer)

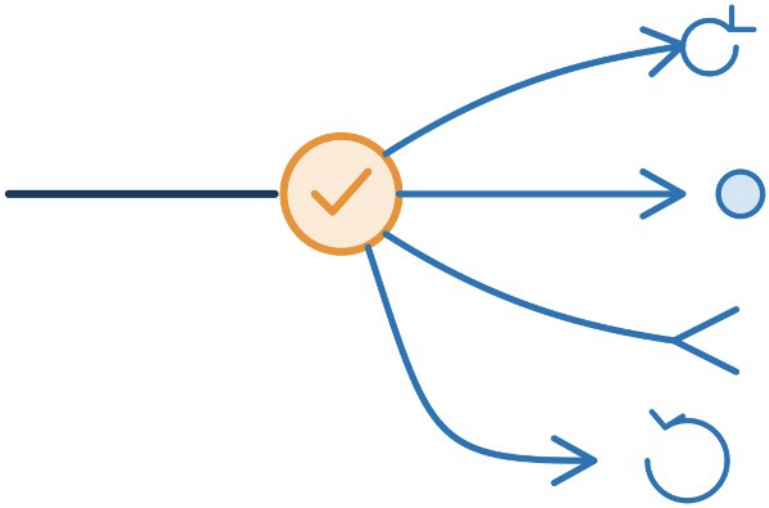
· Rob Fitzpatrick — The Mom Test

BLOQUE D · DECIDIR CON EVIDENCIA

Capítulo 10. Pivotar, iterar, segmentar o perseverar

EN CONSOLIDACIÓN

Ante la evidencia, las opciones no son binarias; la falacia del coste hundido es el sesgo más peligroso de la decisión.



Ante la evidencia, la decisión se abre en cuatro caminos, no en un sí o un no.

Introducción

El discovery culmina en una decisión, y decidir bien con evidencia mixta es una competencia propia. Este capítulo desarrolla el espectro completo de opciones ante un resultado, la falacia del coste hundido, qué exige un pivote responsable y un marco para decidir cuando la evidencia no es concluyente. Su estado es en consolidación porque, aunque el marco está aceptado, su uso explícito y libre de apego emocional aún madura en muchos equipos.

10.1 Las opciones reales ante la evidencia

Ante los resultados de un experimento las opciones no son binarias. Perseverar: la evidencia es lo bastante positiva para continuar en la dirección actual. Iterar: ajustar elementos concretos manteniendo el rumbo general. Pivotar: cambiar un aspecto fundamental —segmento, propuesta de valor, canal— conservando los aprendizajes. Parar: abandonar porque la evidencia indica que no es viable. La mayoría de las decisiones reales caen en los casos intermedios —iterar y pivotar—, no en los extremos.

10.2 La falacia del coste hundido

El sesgo más peligroso en estas decisiones es justificar perseverar por el tiempo o los recursos ya invertidos: «llevamos seis meses y el equipo está muy comprometido; además, los dos usuarios que completaron fueron entusiastas». Lo ya invertido es irrelevante para la decisión futura; lo relevante es cuál es la mejor inversión de los próximos recursos dada la evidencia actual. Las señales de alerta son los argumentos sobre «el esfuerzo invertido», el foco en los pocos casos positivos y la resistencia emocional a considerar el pivote.

Capturar el aprendizaje antes de pivotar. Un pivote responsable no es un giro impulsivo: documenta qué se aprendió y por qué no funcionó, formula la nueva hipótesis basándose en ello, la

valida antes de invertir y comunica el razonamiento. La frase guía: «aprendimos que [hipótesis] era falsa porque [evidencia]; la nueva hipótesis es X y la validaremos con [experimento]».

10.3 Un marco para la evidencia mixta

Cuando la evidencia no es claramente positiva ni negativa, conviene ponderar varios factores: la calidad de la evidencia —¿es robusta?, ¿hay sesgos en la muestra?—, la magnitud de las señales —¿son significativas o podrían ser ruido?—, el coste de oportunidad —¿qué dejamos de hacer?—, la reversibilidad —¿podemos iterar o la decisión es definitiva?— y la incertidumbre residual. Si esta última es alta y el coste de un experimento adicional es bajo, lo sensato es diseñar un test más específico antes de decidir. Así el discovery enlaza con el ritmo de descubrimiento continuo del equipo, donde decidir es un acto recurrente, no un hito final.

Lo que debes saber hacer

Al dominar este tema, un profesional es capaz de:

- Situar una decisión en el espectro perseverar / iterar / pivotar / parar, en lugar de plantearla como binaria.
- Reconocer la falacia del coste hundido en un argumento y reorientar la decisión hacia la mejor inversión futura.
- Estructurar un pivote responsable: documentar el aprendizaje, formular la nueva hipótesis y validarla antes de invertir.
- Aplicar el marco de evidencia mixta (calidad, magnitud, coste de oportunidad, reversibilidad, incertidumbre residual) a una decisión real.
- Decidir si la incertidumbre residual justifica un experimento adicional antes de comprometer recursos.

Errores y antipatrones frecuentes

Plantear la decisión como binaria. Perseverar o abandonar ignora iterar y segmentar, que es donde caen la mayoría de los casos reales.

Ceder a la falacia del coste hundido. Lo invertido no debe pesar en la decisión futura; pregunta por la mejor inversión de los próximos recursos.

Pivotar sin capturar aprendizajes. Cambiar de rumbo sin documentar por qué arriesga repetir los mismos errores en la nueva dirección.

Decidir con incertidumbre alta y test barato disponible. Si un experimento adicional asequible aclararía la situación, decidir sin él es precipitado.

Para profundizar

· Teresa Torres — Product discovery

· Scrum Manager — Mapa de producto