

Improving User Stories: A Case Study in the Chilean Banking Industry

Jorge Gómez Opazo
Departamento de Informática
Universidad Técnica Federico Santa María (UTFSM)
Santiago, Chile
jorgegom@gmail.com

Claudia López
Departamento de Informática
Universidad Técnica Federico Santa María (UTFSM)
Santiago, Chile
claudia@inf.utfsm.cl

Abstract—Although a variety of user story refinement techniques have been proposed, there are still few empirical studies that assess their effectiveness in the industry. This paper reports the results of a mixed methods approach to evaluate the combined use of techniques for refining user stories in a Chilean banking organization. A combination of INVEST, 3 Cs and Specification by Example (SbE) techniques was proposed to improve the quality of user stories that were being generated in the organization during its transition from traditional development methods to agility. To validate the proposal, a comparison of case studies was carried out. A group of user stories that were created and refined with the method previously used in the organization was contrasted with a second and third group of stories for which the proposed technique was used. The results show that the combined use of INVEST, 3 Cs and SbE improved the user stories quality and was related to positive changes in the development team motivation.

Keywords— Agile, Scrum, User stories, 3Cs, Refinement, INVEST.

I. INTRODUCCIÓN

Muchas organizaciones están hoy haciendo una transición desde perspectivas tradicionales de gestión de desarrollo de software hacia la agilidad buscando nuevas formas de trabajo que permitan lanzar rápidamente productos innovadores al mercado, probar barato, ver cómo reaccionan los clientes y confirmar si el producto está generando los retornos esperados.

En la práctica hacer esta transición involucra una serie de decisiones respecto a la apropiación de los nuevos marcos de trabajo en cada actividad que llevan a cabo los equipos de desarrollo. Sin embargo, existe poca evidencia empírica que compare técnicas alternativas y, por lo tanto, pueda informar la toma de estas decisiones.

Para contribuir con evidencia empírica sobre la efectividad de técnicas ágiles en equipos de desarrollo de software que están haciendo una transición hacia la agilidad, este trabajo reporta un método de investigación mixto que permitió comparar dos técnicas alternativas de refinamiento de historias de usuario (HDU) en una organización bancaria de un país latinoamericano.

En la organización, las HDU eran refinadas por todo el equipo de desarrollo según las directrices de la ceremonia de refinamiento [6]. Sin embargo, el mismo equipo reportaba luego que las historias no eran de calidad ya que no las

entendían a cabalidad al momento de desarrollar, eran incompletas y no reflejaban correctamente el valor de la historia. Para resolver el problema, se propuso una técnica híbrida de refinamiento basada en INVEST [5], 3 Cs [7, 8] y *Specification by Example* (SbE) [2] que buscaba generar HDU de mejor calidad.

La hipótesis postula que aplicando la técnica híbrida es posible tener HDU de mejor calidad. Para evaluar la hipótesis, se analizó la calidad de un grupo de historias que fueron creadas y refinadas con el proceso de refinamiento inicial de la organización. Posteriormente, se realizó un entrenamiento al equipo de desarrollo en la técnica híbrida mediante sesiones de capacitación. En los siguientes dos *sprints* se crearon y refinaron las historias usando la técnica propuesta. Finalmente, se compararon indicadores de calidad de los tres grupos de HDU y las apreciaciones del equipo de desarrollo respecto a las técnicas usadas.

El resto de este documento está organizado de la siguiente forma. La segunda y tercera sección describen el contexto organizacional y el problema. La cuarta sección explica la técnica de refinamiento inicial de la organización y detalla la técnica propuesta. La quinta sección describe metodología de mixta que permitió la evaluación de las HDU generadas con ambas técnicas. La sexta sección detalla los resultados obtenidos respecto a la calidad de las historias y las apreciaciones del equipo de desarrollo. Finalmente, la última sección presenta las conclusiones que emanan del trabajo realizado y sus resultados.

II. CONTEXTO ORGANIZACIONAL

Este trabajo se encuentra inserto en una organización bancaria que se encuentra en un proceso de transformación, donde uno de sus lineamientos principales es el uso de agilidad como elemento clave para el desarrollo de proyectos de software. La organización está usando Scrum [1] como referente de agilidad. Scrum es un marco de trabajo ágil para el desarrollo y mantención de productos en forma incremental que busca entregar el máximo valor posible en cada iteración.

En la organización hay varias células o equipos de desarrollo. Cada célula está encargada de las necesidades de un producto diferente con sus propios *stakeholders*.

Si bien las células son autónomas, existe un comité de soporte a la agilidad que identifica y elimina obstáculos y

problemas que pudieran afectar a nivel general y define directrices transversales a todas las células. Además, existe un grupo de *coaches* de agilidad encargados de apoyar a las células en cuanto a la metodología.

Cada célula está compuesta por un *product owner*, un especialista del producto, un *scrum master*, el equipo de desarrollo (*devs*), un encargado del aseguramiento de la calidad del incremento terminado (QA), un desarrollador *front-end*, un encargado de la experiencia de usuario (UX), un líder técnico encargado de la arquitectura y de coordinar los desarrollos de equipos externos o *vendors*, y un encargado de automatizar pruebas funcionales.

En el proceso de desarrollo, un sprint dura dos semanas. Cada tres *sprints* se genera un *release* con todas las HDU desarrolladas y se realiza el paso a producción. Al inicio de cada sprint se lleva a cabo la reunión de planificación (*sprint planning*), y al final se realiza una demo de las historias realizadas en el sprint y una reunión de retrospectiva. Además, al inicio de cada día de trabajo se realiza una reunión.

III. PROBLEMA

Dentro del proceso de transformación hacia la agilidad, las HDU cumplen un rol primordial reflejando la necesidad del cliente. Ellas son una forma de describir una funcionalidad deseada desde la perspectiva de quien lo necesita. Una historia de usuario es trabajada por el *product owner* y luego es refinada en la ceremonia de refinamiento a mediados del sprint por el equipo de desarrollo.

Sin embargo, se ha notado a lo largo de los *sprints* que se han estado creando muchas historias de mala calidad en los siguientes sentidos: 1) El equipo de desarrollo no la entiende a cabalidad, 2) El equipo estima que se encuentra incompleta de acuerdo al *definition of ready*¹, 3) No está claro el valor que la historia aporta, 4) La historia es demasiado grande para ser abordada en un sprint, o 5) La historia tiene dependencias que impiden que pueda ser completada. Como consecuencia, en la ceremonia de planificación, el equipo no puede estimar correctamente las historias que el *product owner* ha priorizado y, por lo tanto, éstas no pueden ser abordadas.

Este tema es recurrente en las reuniones de retrospectivas, generando roce entre representantes del negocio y el equipo de desarrollo. En ocasiones se ha expuesto el tema de que los *product owners* llegan a la reunión de planificación del sprint con la historia “poco refinada.” En la práctica, no hay un indicador claro que permita a los representantes del negocio conocer el estado de una historia de usuario en cuanto a su grado de refinamiento.

Este problema de la calidad de las HDU conduce a las siguientes consecuencias:

- La reunión de refinamiento se extiende en torno a conversaciones de HDU poco definidas, por lo que no se alcanzan a revisar todas las historias que el negocio ha priorizado.

¹ Acuerdo del equipo en que se definen las condiciones para que una historia pueda ser comprometida.

- La reunión de refinamiento no es suficiente como única instancia para clarificar las historias.
- Las HDU mal refinadas no se pueden estimar en la reunión de planificación del *sprint*.
- Al no poder estimar todas las historias priorizadas, en ocasiones la cantidad de historias comprometidas por el equipo ha sido menor a lo que el equipo podría desarrollar dada su velocidad de trabajo.
- Debido al poco grado de refinación de historias, hay criterios de aceptación ambiguos que no son correctamente interpretados por el equipo de desarrollo. Esto puede derivar en que se creen historias solo para corregir el comportamiento erróneo de la aplicación, generando así re-trabajo.
- Al no tener claridad acerca del valor de negocio de cada historia, se terminan priorizando historias de poco valor para que sean desarrolladas en un *sprint*.

Este trabajo propone una técnica para mejorar la calidad de las HDU por medio de la aplicación de una técnica de refinamiento que apunta a la colaboración y conversación en torno a las necesidades del negocio, mejorando el entendimiento de la historia tanto por el lado del negocio como por el lado del equipo de desarrollo, y alineando las HDU a los objetivos de negocio. Además, este trabajo busca establecer un método de investigación mixto que permita evaluar las técnicas de refinamiento alternativas para que el equipo de desarrollo pueda tomar decisiones basadas en evidencia sobre qué técnica usar en el futuro.

IV. REFINAMIENTO DE HISTORIAS DE USUARIO

A. Técnica de refinamiento inicial

En la técnica de refinamiento usada por la organización inicialmente, el *product owner* obtiene las HDU directamente de los stakeholders y las agrupa en épicas. Las HDU priorizadas se revisan en la reunión de refinamiento que se realiza una vez por *sprint*, entre el séptimo u octavo día. El resultado de esta reunión de refinamiento son las historias que se revisan en la siguiente reunión de planificación del *sprint*.

La reunión de refinamiento (*backlog grooming*) [6] es una práctica útil para aclarar dudas que tenga el equipo de desarrollo y priorizar HDU. En ella se revisan las historias que se han priorizado para el siguiente *sprint*. Si es necesario se dividen las historias en otras más pequeñas de modo de asegurar que el *product backlog* contiene los elementos correctos que se han priorizado.

Algunas actividades de esta ceremonia son las siguientes:

- Remover HDU que ya no son relevantes.
- Crear nuevas HDU en respuesta a nuevas necesidades.
- Re-assignar prioridad de historias.
- *Estimar* historias que aún no tienen estimación.
- Corregir estimaciones dada nueva información recibida.
- Dividir HDU con alta prioridad pero que son muy grandes para ser abordadas en un *sprint*.

La reunión de refinamiento es la instancia donde el equipo levanta aprehensiones que tiene sobre la historia. El equipo no

participa directamente en la creación de la historia y su aporte, solo se remite a realizar ciertas aclaraciones en la reunión de refinamiento, por lo que el valor que la historia aporta solo lo conoce el *product owner*. Se puede presentar dudas técnicas, de negocio o de interacción de usuario. Cualquier duda presentada, será revisada posteriormente en la reunión de planificación del sprint, ya que no hay más instancias para volver a revisar la historia.

Si una historia no puede ser abordada en un sprint debido a su gran tamaño, la historia es rechazada por el equipo. Lo mismo ocurre si una historia no es aclarada completamente o hay aspectos técnicos que aún no se conocen. Si aún hay dudas técnicas o dudas de interacción de usuario respecto a una historia, el equipo se ve obligado a crear HDU de investigación o de UX, desplazando el desarrollo a un sprint posterior.

B. Técnica de refinamiento propuesta

Para resolver los problemas identificados en el uso de la técnica de refinamiento de la organización, la técnica de refinamiento propuesta busca apoyar el proceso de creación de HDU desde que el *product owner* obtiene las necesidades hasta que esta historia se considera lista para ser desarrollada por el equipo. El objetivo de la técnica es permitir tener HDU de mejor calidad con un claro objetivo de negocio considerando alta colaboración entre el equipo y el *product owner* durante todo el proceso de refinamiento de cada historia de usuario.

La Figura 1 presenta la técnica de refinamiento propuesta, que incluye cinco pasos para refinar una historia de usuario maximizando el valor de negocio. Estos pasos están inspirados en las técnicas INVEST, SbE, y 3Cs, que se describen a continuación.

INVEST

INVEST [5] es una de las formas de medir la calidad de las HDU. Propone los siguientes principios para escribir una buena historia de usuario:

- **I** – Independiente: Una historia debe ser independiente de otras para una mejor planificación y priorización.
- **N** - Negociable: Una historia es una descripción corta de una necesidad que no incluye detalles. El resto se añade mediante la conversación.
- **V** - Valiosa: Debe aportar valor para el cliente.
- **E** - Estimable: Debe poder ser estimada por el equipo. Historias muy grandes o incompletas son difíciles de estimar.
- **S** - *Small* (pequeña): Debe permitir ser completada en poco tiempo. Comúnmente en menos de una semana.
- **T** - Testeable: Debe ser capaz de ser probada para verificar que se completado con éxito. Para esto se plantean criterios de aceptación dentro de la historia que nacen de la conversación.

Specification by Example

Specification by Example [2] (SbE) es un conjunto de patrones de procesos que facilitan el cambio en productos de software para asegurar que el producto correcto es entregado eficientemente. Considera el planteamiento colaborativo de ejemplos realistas en lugar de enunciados abstractos,

disminuyendo la brecha de comunicación entre los *stakeholders* del negocio y los equipos de desarrollo. Este enfoque mejora la comunicación entre el *product owner* y los integrantes del equipo. Además, permite alinear la especificación de un sistema con las pruebas de aceptación.

Los patrones clave del proceso de SbE son: derivar el alcance de los objetivos en conjunto, especificar colaborativamente, ilustrar especificaciones usando ejemplos, refinar las especificaciones, automatizar la validación sin cambiar la especificación, validar el sistema frecuentemente, y evolucionar el sistema de documentación.

En SbE, los requisitos funcionales, las especificaciones y las pruebas de aceptación se consideran lo mismo. El resultado final de una aplicación completa de SbE es un sistema de documentación viva que explica lo que el sistema hace y qué es tan relevante y confiable como el código del sistema. Documentación viva se refiere a que cualquier cambio en el código es reflejado en la documentación del sistema, manteniendo las especificaciones siempre actualizadas.

Las 3 C's

Las 3Cs [7], [8] es una técnica que define que las HDU tienen tres aspectos críticos:

- "**Card**" (tarjeta): Las HDU son escritas sobre tarjetas. La tarjeta no contiene toda la información que el requerimiento busca. Solo tiene el texto suficiente para identificar el requerimiento y recordar a todos de que se trata. Es usada en la reunión de planificación y refleja una prioridad y costo.
- "**Conversation**" (conversación): El requerimiento es comunicado desde el cliente a los desarrolladores por medio de la conversación, es decir, intercambiando pensamientos, opiniones y sensaciones. La conversación es mayormente verbal, pero puede ser complementada con documentación.
- "**Confirmation**" (confirmación): Finalmente, los criterios de aceptación confirman que la historia cumple con lo esperado. Cuando la conversación acerca de una tarjeta llega a los detalles de la prueba de aceptación de la historia, el cliente y los desarrolladores establecen los detalles finales de lo que hay que hacer. Al terminar el *sprint*, los desarrolladores demuestran que estas pruebas se ejecutan correctamente. Con esto, el cliente verifica que el equipo entregará lo que se necesita.

Cada una de estas técnicas cubre cierto aspecto asociado al mejoramiento de la calidad de las HDU. INVEST contribuye en la medición de las historias, lo cual se puede tener presente en todo momento durante los refinamientos. SbE busca obtener las necesidades de negocio desde los stakeholders y plantear las HDU de alto nivel con sus correspondientes ejemplos claves. Las 3C agregan detalle, orden y prioridad a las historias por medio de la conversación y confirmación.

En base a las tres técnicas revisadas, la técnica híbrida propuesta incluye los siguientes pasos:

1) Derivación de las necesidades de negocio

El primer paso consiste en conocer las necesidades de negocio directamente de los *stakeholders*. El *Product owner* en

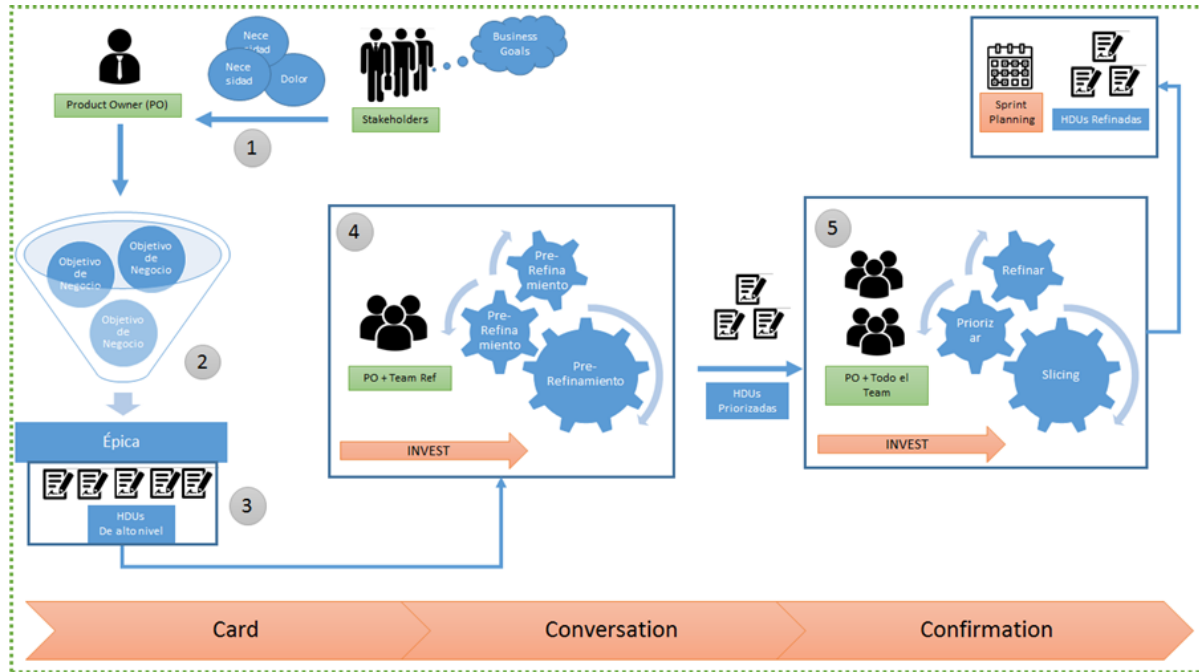


Fig. 1. Técnica de refinamiento de historias de usuario propuesta.

conjunto con el especialista del producto se reúnen constantemente con cada uno de los representantes del negocio para conocer sus problemas, sus necesidades y su visión a futuro de lo que quieren para el producto.

2) Derivación de las necesidades de negocio

El primer paso consiste en conocer las necesidades de negocio directamente de los *stakeholders*. El *Product owner* en conjunto con el especialista del producto se reúnen constantemente con cada uno de los representantes del negocio para conocer sus problemas, sus necesidades y su visión a futuro de lo que quieren para el producto.

3) Creación de épicas

Las grandes necesidades, problemas y visión del negocio se representan por medio de épicas. Una épica está compuesta por un conjunto de HDU y comúnmente involucra un volumen de trabajo que no puede ser abordado en un solo sprint. El valor de negocio aportado por una épica debe estar claramente identificado y ser medible. Además, cada épica debe tener definido el objetivo de negocio que busca. Cada historia de usuario que forma parte de una épica aportará a lograr el objetivo de negocio deseado y sumar el valor buscado.

4) Creación de historias de usuario:

Una vez que se ha identificado la épica, con su objetivo de negocio y valor aportado claro, el *product owner* en conjunto con el especialista del producto crean las HDU de alto nivel que permitirán llegar al objetivo planteado por la épica. La historia de usuario es creada en una tarjeta (“*Card*”) que define lo siguiente:

- Título de la historia

- Contexto
- Objetivo de la historia y valor aportado

Además, el *product owner* realiza la primera priorización de las HDU dentro de la épica. Esta priorización irá cambiando al ir mejorando el refinamiento de las historias.

5) Pre-refinamiento

Para cada historia definida y priorizada en el paso anterior, se realiza un pre-refinamiento diario con el objetivo de:

- Clarificar y conversar sobre el contexto y objetivos de la historia de usuario (“*Conversation*”)
- Definir los criterios de aceptación de la historia (“*Confirmation*”)
- Definir el UX si la historia considera interacciones de usuario.
- Validar técnicamente si el UX puede ser implementado dentro de un sprint. Si no es posible, se itera para llegar a algo que alcance a ser desarrollado en un sprint y que agregue el valor esperado por el negocio.
- Re-priorizar si fuese necesario. Esto puede suceder porque el equipo considera que otras historias son más prioritarias en cuanto al valor generado.
- Dividir la historia si es que no es abordable en un *sprint*. Todas las historias que surjan de este paso deben aportar valor y ser medibles.

El responsable del pre-refinamiento diario es un sub-equipo llamado equipo refinador que incluye como mínimo al líder técnico, UX, especialista del producto y *product owner*. Eventualmente se van sumando colaboradores de otras áreas en la medida que se necesita.

La dinámica es la siguiente: mediante reuniones diarias, de a lo más una hora, se expone una historia de usuario que el negocio ha priorizado para el siguiente *sprint* (o el subsiguiente) y se realizan las siguientes preguntas clave:

- ¿Qué valor agrega la historia y por qué es necesaria?
- ¿Hay alguna alternativa de solución más sencilla?
- ¿Es más prioritaria esta historia que otras historias del *backlog* que agregan más valor?
- ¿Cuántos clientes o usuarios la usarán?
- ¿Hay que generar algún proceso manual para soportar la parte no desarrollada?
- ¿Hay algún riesgo operacional asociado?
- ¿Cómo se medirá el impacto de la historia para el objetivo que se busca?
- ¿Se puede acortar el alcance de la historia o dividirla y seguir agregando valor?
- Además, se preguntará por ejemplos de alto nivel de lo que busca la historia. Esto ayuda a aterrizar los objetivos de la historia y permite escribir buenos criterios de aceptación (según *SbE*).

6) Refinamiento

Se debe agendar al menos una reunión de refinamiento en donde esté presente todo el equipo. Idealmente debieran ser dos reuniones por *sprint* (una semanal), las que se realizan el tercer día y octavo día del *sprint*. Los objetivos de estas reuniones son:

- Revisar las historias que se priorizaron para el *sprint* y que se verán en la próxima reunión de planificación con ayuda de una presentación gráfica, realizada por el *product owner*, y que le da contexto a las historias.
- Generar conversación sobre la historia y obtener visiones de todo el equipo (ya que es posible que algún integrante del equipo no haya participado de los pre-refinamientos). Es importante, por ejemplo, la visión de QA y desarrolladores para mejorar los criterios de aceptación o conversar acerca de cómo se medirá el valor aportado.
- Aclarar dudas, si es necesario, y agregar detalles a la historia de usuario de manera acorde.
- Realizar *slicing*² si el equipo considera que una historia es muy grande aún para ser abordada en un *sprint*.
- Realizar una pre-estimación a alto nivel del costo de realizar cada historia de usuario revisada. Esto ayudará a visualizar si es abordable en un *sprint* o es necesario realizar más *slicing*.
- Re-priorizar nuevamente, si es pertinente, y reordenar el *product backlog* según corresponda.
- Eliminar elementos innecesarios del *product backlog*.
- El *product owner* debe presentar al equipo como va cambiando la visión del producto con las historias que se están refinando.

V. METODO DE INVESTIGACION

Para evaluar la efectividad de ambas técnicas, se mide si los grupos de HDU refinados con la técnica propuesta son mejores

en cuanto a su calidad que el grupo de historias refinadas con la técnica inicial (que es considerada como grupo de control). Para esta evaluación empírica se siguen los siguientes pasos:

- 1) En un *sprint*, se crea y refina un grupo de HDU como se hace actualmente (grupo de control).
- 2) Se registran mediciones acerca de la calidad de HDU, y de las apreciaciones de los integrantes del equipo. El detalle de las mediciones consideradas se explican más adelante.
- 3) Se entrena al equipo en la técnica propuesta mediante dos sesiones de una hora, de modo que puedan aplicar la técnica propuesta en futuros *sprints*. En la primera sesión se realiza una presentación a todo el equipo de la técnica (representada en la Figura 1) junto con la explicación de los pasos que la técnica propone. En la segunda sesión se concretizan los pasos con un ejemplo de una historia de usuario que se había desarrollado anteriormente.
- 4) En el *sprint* siguiente, se crea y refina un nuevo grupo de HDU con el mismo equipo, pero esta vez se aplica la técnica propuesta.
- 5) Se registran las mediciones mencionadas en el punto 2 para este *sprint*.
- 6) En el *sprint* sub-siguiente, se repiten los pasos 4 y 5 para un nuevo grupo de HDU.
- 7) Se comparan las mediciones de los tres grupos de historias, aquellas desarrolladas con la técnica inicial versus aquellas desarrolladas con la técnica híbrida propuesta.

Más específicamente, para las mediciones se utiliza el método INVEST para evaluar la calidad de las HDU. Además, se miden conceptos asociados al equipo de desarrollo como la motivación y satisfacción con el fin de evaluar las técnicas en relación a estos aspectos.

A. Mediciones de calidad de las historias de usuario.

Las métricas para medir calidad están basadas en los criterios de INVEST y son los siguientes:

a) Valor de negocio (*Business Value*).

Asociado al criterio *Valuable*, el valor de negocio se refiere al valor que aporta una historia de usuario según estimación de los *product owners*.

Una buena historia debe aportar valor y este debe ser medible. Para medir el valor de negocio de una historia, se definen cuatro indicadores que reflejan las dimensiones asociadas al producto que los *product owner* de la organización consideran cuando deben asociar valor a las historias. Los indicadores son: aumento en ventas, reducción del costo operacional, disminución del riesgo, y fidelización del cliente, como se muestra en la Tabla 1.

TABLE I. INDICADORES DEFINIDOS POR EL *PRODUCT OWNER*

Indicadores	Relevancia
Aumento de ventas (\$)	Alta
Reducción del costo operacional	Media
Disminución del riesgo	Media
Fidelización del cliente	Baja

² Dividir una historia de usuario en otras más pequeñas

Por cada historia de usuario, se establece su relevancia en función de su impacto en cada indicador usando una escala ordinal de tres niveles: baja, media, y alta. Considerando la relevancia, el *product owner* estima el valor de negocio de cada historia con una escala Likert de 0 a 10, donde:

- 0: Historia no aporta valor al negocio.
- 10: Historia aporta muchísimo valor al negocio

b) Tamaño de las historias de usuario.

Asociado al criterio *Small*, se espera que el tamaño de una historia debe ser lo suficientemente pequeña para ser abordada en un sprint. El tamaño refleja cuanto esfuerzo el equipo estima que involucra desarrollar la historia de usuario.

El tamaño de una historia de usuario se mide en “puntos por historia.” Una historia se considera de mayor tamaño en la medida que aumenta la cantidad de puntos que el equipo estima que la historia tiene.

Esta métrica es obtenida desde CA Agile Central³ (antes conocida como Rally) que es la herramienta de gestión de proyectos ágiles con la que el equipo trabaja. CA Agile Central es un producto de CA Technologies reconocido como uno de los líderes para herramientas de planificación ágil [9].

c) Historias de usuario testeables.

Asociado al criterio *Testable*, una buena historia de usuario debe tener criterios de aceptación testeables para validar que se hizo lo esperado.

Para medir este aspecto se usa el porcentaje de HDU que el equipo indica que se pueden testear. En nuestro caso, esta métrica es obtenida desde CA Agile Central.

d) Independencia de historias de usuario.

Asociada al criterio *Independent*, se refiere a que una historia no depende de otra para ser completada. Con esto, se facilita el cambio de orden de las historias si el *product owner* lo estima conveniente. Dos historias dependientes entre ellas no pueden ser desarrolladas en el mismo sprint sin resolver antes las dependencias.

Este concepto se mide a través de la cantidad de HDU independientes por sprint. Se considera que una historia de usuario es independiente cuando no posee dependencias no resueltas que impiden que se complete en forma independiente. Una dependencia podría ser análisis técnicos no realizados, interacciones de usuario no definidas, otra historia de usuario no desarrollada que contiene parte de la funcionalidad o análisis de negocio no realizados. Esta métrica se obtiene desde CA Agile Central de la organización.

e) Claridad y definición de las historias de usuario.

Asociado al criterio *Estimable*, se refiere a que si una historia está clara y bien definida, es muy probable que sea fácilmente estimable por el equipo ya que se reduce la incertidumbre y permite visualizar si la historia puede ser abordada en un sprint o debe ser dividida.

Este indicador se obtiene por medio de una encuesta a cada integrante del equipo al final de cada sprint bajo estudio. Para evaluarlo, el equipo usa una escala Likert de 1 a 5 que indica:

- 1: Ninguna historia de usuario estaba clara y bien definida.
- 5: Todas las HDU estaban claras y bien definidas.

f) Negociación de las historias de usuario.

Asociado al criterio *Negotiable*, el que una historia sea negociable se refiere a que el detalle de la historia fue creado y acordado en conjunto entre el *product owner* y el equipo mediante la conversación.

Este indicador se obtiene por medio de una encuesta a cada integrante del equipo al final de cada sprint, usando una escala Likert de 1 a 5 que indica:

- 1: Ninguna historia de usuario se negoció antes de ser comprometida.
- 5: Todas las HDU se negociaron antes de ser comprometidas.

B. Motivación y satisfacción del equipo.

Además de las métricas de calidad, se miden aspectos de motivación y satisfacción de los integrantes del equipo con el objetivo de evaluar efectos en aspectos transversales al equipo. Las métricas que se usan son las siguientes:

a) Motivación post refinamiento y planificación del sprint.

Se mide cuán motivado declara estar cada integrante del equipo. Se utiliza una escala Likert de 1 a 5 que representa:

- 1: Me encuentro poco motivado.
- 5: Me encuentro *full* motivado.

b) Satisfacción del equipo con respecto al proceso de refinamiento del sprint.

Se mide cuán satisfecho declara encontrarse cada integrante del equipo en este aspecto. Se utilizará una escala Likert de 1 a 5 donde:

- 1: No me gustó para nada el proceso de refinamiento.
- 5: Me siento muy satisfecho con el proceso de refinamiento.

c) Apreciación del equipo sobre cómo se refinó y planificó el sprint.

Esta corresponde a una pregunta abierta con el fin de obtener retroalimentación de los integrantes del equipo acerca de la técnica aplicada.

La Figura 2 representa la metodología propuesta en este trabajo. Por cada sprint bajo estudio, se obtuvieron los datos necesarios desde CA Agile Central (Rally), se realizaron encuestas y el *product owner* estimó el valor de negocio de cada historia de usuario. Cabe notar que nuestro método es independiente de la herramienta CA Agile Central. Las mismas métricas podrían ser obtenidas por otros medios.

³ <https://www.ca.com/us/products/ca-agile-central.html>

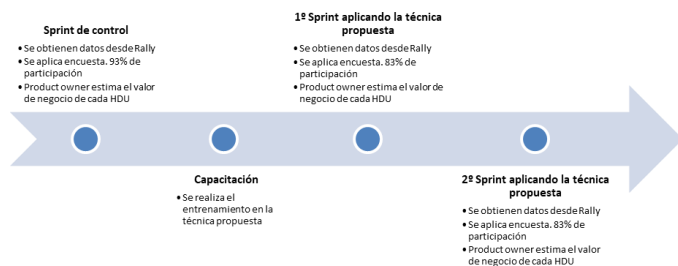


Fig. 2. Metodología

VI. RESULTADOS

Esta sección detalla los resultados obtenidos al comparar la calidad de las HDU generadas con la técnica inicial versus la técnica propuesta. Además, describe los hallazgos asociados a las apreciaciones del equipo de desarrollo respecto al uso de cada técnica.

A. Calidad de las historias de usuario.

Sprint de control con técnica inicial

En el sprint con la técnica inicial se observa que sólo una historia de usuario, de las siete que se trabajaron, fue liberada a producción y sólo dos impactaban un KPI importante para el negocio. Las otras cinco historias no aportaban valor al negocio. Dos historias tenían dependencias no resueltas y dos HDU tenían criterios de aceptación no testeables. La Tabla 2 muestra el detalle de las HDU trabajadas durante el sprint de control. La columna EST describe el estado que puede ser Aceptada (A) indicando que la HDU ha sido aceptada por el *product owner*, o Liberada (L) si es que la HDU que se encuentra liberada en producción de cara a cliente. La columna EE denota la estimación del equipo medida en puntos de historia. KPI se refieren al área clave de la organización que es impactado por la HDU. VNE es el valor de negocio estimado, DNR refleja si hay dependencias no resueltas, CA es el número de criterios de aceptación y CAT es el número de criterios de aceptación considerados testeables.

TABLE II. HDU EN SPRINT DE CONTROL

HDU	EST	EE	KPI	VNE	DNR	CA	CTA
Agregar GTM derivación SII	A	5	Fidelizar	0	1	1	0
Derivación a Empresarios	L	20	Ventas	3	1	4	4
Flujo Simulador	A	13	-	1	0	4	4
Flujo navegable IF	A	13	-	0	0	4	4
Caluga [UX]	A	8	-	0	0	3	0
Certificación - Simulador	A	13	-	0	0	0	0
Certificación - IF	A	13	-	0	0	0	0
Total		85		4	2	16	12

Primer sprint aplicando técnica propuesta

En el sprint en que se usó la técnica de refinamiento propuesta por primera vez, se observa que las dos historias del *sprint* fueron liberadas a producción y ambas historias impactaron KPIs definidos por el negocio. Además, no hubo

dependencias no resueltas ni criterios de aceptación no testeables. En la tabla 3 se muestra el detalle de este *sprint*.

TABLE III. HDU EN EL 1ER SPRINT APLICANDO LA TÉCNICA PROPUESTA

HDU	EST	EE	KPI	VNE	DNR	CA	CTA
Desplegar monto en caluga ejecutivo	L	20	Ventas Costos	10	0	4	4
Derivar a clientes sin margen a IF	L	20	Ventas Costos	10	0	4	4
Total		40		20	0	8	8

Segundo sprint aplicando técnica propuesta

En el segundo sprint en el que se usó la técnica de refinamiento propuesta. Se observa al igual que en el sprint anterior que las dos historias del sprint fueron liberadas a producción y ambas historias impactan KPIs definidos por el negocio. Además, tampoco hubo dependencias no resueltas ni criterios de aceptación no testeables (Ver Tabla 4).

TABLE IV. HDU EN EL 2DO SPRINT APLICANDO LA TÉCNICA PROPUESTA

HDU	EST	EE	KPI	VNE	DNR	CA	CTA
Lógica Caluga (Mandatos DD)	L	20	Ventas Costos	10	0	4	4
Permitir curse con excepción	L	20	Ventas Costos	9	0	6	6
Total		40		19	0	10	10

a) Valor de negocio por historia de usuario.

Con respecto al valor de negocio aportado por historia, los resultados muestran que, usando la técnica de refinamiento propuesta, hubo un mayor valor de negocio estimado por el *product owner* por sprint (ver Figura 3). En el *sprint* de control se registró un aporte de valor de negocio de cuatro, mientras que en los *sprints* donde se usó la técnica propuesta hubo un aporte de valor de negocio de 20 y 19, respectivamente. Esto se relaciona a un más alto promedio de valor de negocio aportado por historia de usuario por cada sprint. En el *sprint* de control se registró un promedio de valor de negocio aportado por historia de 0.6, mientras que en los *sprints* donde se hizo uso de la técnica propuesta los promedios fueron 10 y 9.5, respectivamente.

Es factible que esto suceda porque el *product owner* esté refinando y priorizando las historias que generan más valor de cara a cliente con la nueva técnica. Es parte de la técnica de refinamiento propuesta considerar el valor de negocio como un elemento significativo al refinar una historia y privilegiar las historias que generan más valor y que impactan los KPIs de negocio más importantes.

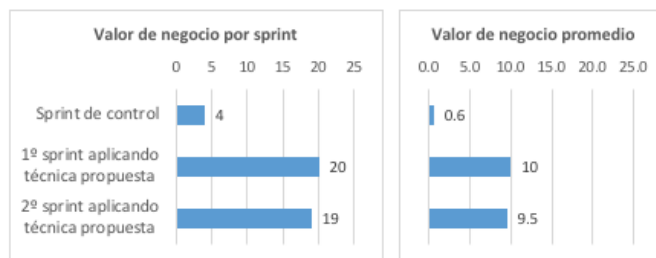


Fig. 3. Valor de negocio en cada *sprint* absoluto y promedio

b) *Tamaño de las historias de usuario.*

En cuanto al tamaño de las HDU, se evidencia un tamaño promedio por historia de usuario más alto en los *sprints* en que se usó la técnica propuesta (ver Figura 4). Mientras el tamaño promedio en el *sprint* de control fue 12.14 puntos por historia, el tamaño promedio fue 20 puntos por historia en cada *sprint* donde se aplicó la técnica propuesta. Por otro lado, en los *sprints* en que se usó la nueva técnica, se observó un número menor de la cantidad de puntos totales por historia que el equipo estaba comprometiendo por *sprint*. En el *sprint* de control el equipo se comprometió con 85 puntos, mientras que en los *sprints* donde se aplicó la nueva técnica, el equipo se comprometió con 40 puntos en cada *sprint*.

Una historia pequeña permite que esta pueda ser desarrollada en dentro de un *sprint*. Es parte del refinamiento de la historia realizar la división necesaria para que esta pueda ser abordada en un *sprint*. Si una historia es muy grande, se debe dividir en historias más pequeñas o simplemente no podrá ser tomada por equipo. Además, siempre se debe considerar que la división de la historia debe suponer que las historias más pequeñas que se crearán agregan valor al negocio. Los resultados muestran que las historias siguen siendo relativamente pequeñas, pero con el uso de la técnica de refinamiento el equipo se compromete con una cantidad menor de puntos sin incidir negativamente en el valor de negocio aportado por *sprint* focalizando de mejor forma los esfuerzos del equipo.

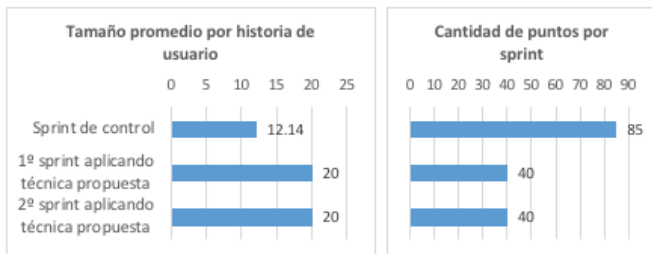


Fig. 4. Tamaño promedio por historia de usuario y puntos por sprint

c) *Historias de usuario testeables.*

También se observó una mayor cantidad de criterios de aceptación testeables en los *sprints* con la técnica propuesta comparado al total de criterios de aceptación del *sprint* con la técnica inicial (ver Figura 5). En el *sprint* de control había 16 criterios de aceptación, pero solo 12 eran testeables. En los *sprints* donde se aplicó la técnica, todos los criterios de aceptación definidos eran testeables. Los criterios de aceptación no testeables en este caso eran criterios del tipo “generar propuesta”. No había claridad, en este caso, sobre cómo y cuándo el *product owner* validaría que tales propuestas corresponden a lo que se quería. Un criterio de aceptación ambiguo es difícilmente testeable. Asegurar que esto no suceda es parte de la fase de confirmación de la historia de la técnica propuesta.

Además, en el *sprint* de control hubo dos HDU no tenían criterios de aceptación. En estos casos, no había claridad sobre cómo se validaría la historia. El verificar que esto no suceda es también parte de la fase de confirmación de la historia de la nueva técnica.

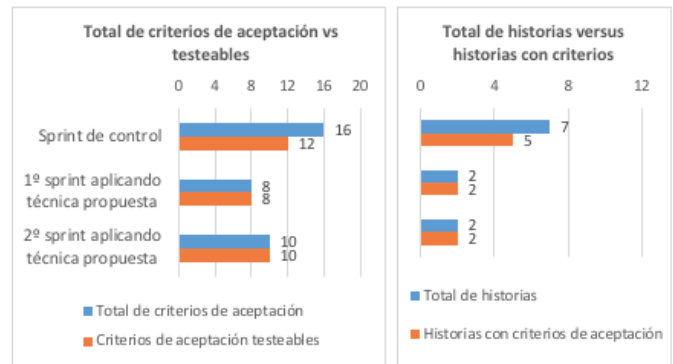


Fig. 5. Criterios de aceptación y criterios de aceptación testeables por sprint

d) *Independencia de historias de usuario.*

En cuanto a la independencia de las HDU, se observó una menor cantidad de dependencias no resueltas en los *sprints* en que se usó la nueva técnica, comparado al *sprint* con la técnica inicial. En el *sprint* de control se registraron dos dependencias no resueltas mientras que en los *sprints* en los que se aplicó la técnica propuesta no se registraron dependencias no resueltas.

Una historia con dependencias no resueltas debe esperar que esa dependencia sea solventada para poder ser completada. Una historia de usuario sin dependencias facilita la re-priorización y permite que se pueda completar sin esperar que otra lo haga antes. Si hay dependencia entre dos historias, éstas no pueden ser tomadas dentro de un mismo *sprint*. Estos temas deben estar resueltos antes de tomar una historia y es parte del refinamiento propuesto el resolverlos con anticipación.

e) *Claridad y definición de las historias de usuario.*

Los integrantes del equipo evaluaron mejor la claridad de las HDU refinadas con la nueva técnica que aquellas historias refinadas con la técnica inicial. En el rango de uno a cinco, el promedio de evaluación de claridad de las historias del *sprint* de control fue 2.54 (ver Figura 6). Por otro lado, el promedio de evaluación de claridad de las historias del primer y segundo *sprint* usando la técnica propuesta, los que fueron 4.4 y 4.6, respectivamente. Este indicador es muy importante ya que una historia clara y bien definida tendrá menos dudas y la estimación tenderá a ser más exacta, ayudando a realizar un *slicing* correcto.

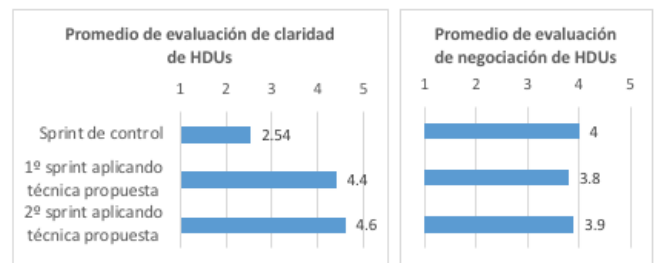


Fig. 6. Evaluación de claridad y negociación de las historias de usuario

f) *Negociación de las historias de usuario.*

En el caso de la negociación de las HDU, se evidenció que la apreciación de los integrantes del equipo fue similar entre los tres *sprints*. En el rango de uno a cinco, el promedio de

evaluación para el sprint de control fue de 4, mientras que para los *sprints* en que se aplicó la técnica propuesta el promedio fue 3.8 y 3.9, respectivamente.

Este indicador mide si existe una negociación en las HDU. Es decir, que como parte del proceso de refinamiento se realiza una conversación en torno a la historia y se negocia su contenido y detalles. La leve baja puede deberse a que parte del equipo no participa de los refinamientos diarios, aunque debe hacerlo en los refinamientos semanales.

En resumen, la evidencia muestra que la técnica de refinamiento propuesta generó HDU de mejor calidad que aquellas generadas con la técnica anterior. Los *sprints* en que se usó la técnica propuesta generaron un número más reducido de historias, pero que en promedio y en términos absolutos aportaban mayor valor al negocio. Aunque el tamaño promedio de las HDU era más grande en el *sprint* de control, la cantidad de puntos de función comprometidos por sprint fue más moderado, lo que se podría asociar a la mayor cantidad de historias liberadas a producción en los *sprints* en que se usó la técnica de refinamiento propuesta. Además, las HDU generadas con la nueva técnica fueron consideradas más testeables, con menos dependencias, más claras y casi igualmente negociables que las historias generadas con la técnica inicial.

B. Motivación y satisfacción del equipo:

Adicional al análisis de calidad, se midió la motivación y satisfacción del equipo buscando posibles relaciones del cambio de la técnica de refinamiento en estos aspectos. Los resultados se detallan a continuación.

a) Motivación post refinamiento y planificación del sprint.

Se observó una mayor motivación de los integrantes del equipo en los *sprints* en que se aplicó la técnica de refinamiento propuesta. En un rango de uno a cinco, el promedio de evaluación de la motivación del equipo fue de 3.54 en el sprint de control. En comparación, esta métrica alcanzó valores de 4.4 y 4.3 en los *sprints* donde se aplicó la técnica propuesta (ver Figura 7).

Es factible que la nueva técnica haya incidido en este resultado debido a que ahora el equipo participa más activamente del proceso de creación de las HDU y se llega a la reunión de planificación con historias más claras, mejor definidas y que aportan mayor valor al negocio. Mantener el equipo motivado es un desafío a diario se considera importante en este proceso de transición hacia un nuevo marco de trabajo.

b) Satisfacción del equipo con respecto al proceso de refinamiento del sprint.

En cuanto a la satisfacción, hubo también un mejor indicador en la evaluación que los integrantes del equipo en los últimos *sprints*. De uno a cinco, el promedio de evaluación de la satisfacción del equipo fue de 3.15 en el sprint de control. En los *sprints* donde se aplicó la técnica propuesta hubo un nivel de satisfacción de 4.3 y 4.2.

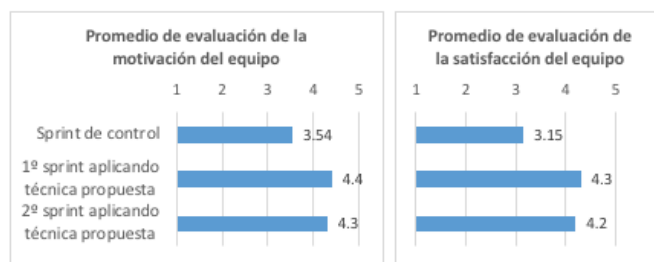


Fig. 7. Evaluación de motivación y satisfacción del equipo

c) Apreciación del equipo sobre cómo se refinaron y planificaron los sprints.

A través de una pregunta abierta, se capturaron sensaciones acerca de cómo se sentía el equipo respecto a la técnica usada en la organización y como ese proceso de refinamiento estaba afectando al equipo. La Tabla 5 muestra ejemplos representativos de la retroalimentación obtenida en esta pregunta. Para el sprint de control, los comentarios que se recibieron mencionan: “Improvisación”, “Faltó llegar con el backlog del sprint más detallado, priorizado antes de refinar”, “Sentí desordenados los últimos sprints”. Estos comentarios contrastan fuertemente con una mayoría de comentarios positivos respecto a la técnica propuesta, como por ejemplo: “Definitivamente superior a Sprints anteriores y mejorable para los próximos”, “Se encontraban bastante detalladas las historias por lo cual se pudieron planificar de buena forma”, “Bien”, “Excelente”, “Mejorando”, “Orden”, “Eficiente”. En general, se ve una mejora en la opinión del equipo.

TABLE V. FEEDBACK DEL EQUIPO

Sprint de control	1er sprint aplicando técnica propuesta	2do sprint aplicando técnica propuesta
-Poco equilibrada la carga para el equipo	-Definitivamente superior a Sprints anteriores y mejorable para los próximos.	-Mejor que sprint anterior pero aún falta disciplina en el proceso y dar más énfasis al valor de negocio de cada HDU.
-Se realizó pensando en cascada	-Empezamos tarde pero sabemos que podemos mejorar los tiempos.	-Orden
-Mejorando pero mejorable	-Se encontraban bastante detalladas las historias por lo cual se pudieron planificar de buena forma.	-El product owner sigue siendo la cara visible de los refinamientos, el negocio sigue aprendiendo.
-Faltó llegar con el backlog del sprint más detallado, priorizado antes de refinar.	-Bien	-Eficiente
-Bajo presión (poco tiempo)	-Vamos mejorando, pero aún falta aprender a priorizar	-En proceso de adaptación a la nueva forma de trabajar. Aun en rodaje.
-Aún queda mucho por mejorar.	-Mejorando continuamente	-La presentación de HDU en forma gráfica aporta bastante al entendimiento.
-Falta de coordinación	-Descripción gráfica hace entendible para todos	-Excelente
-Nos quedamos sin historias y el proceso fue muy improvisado	-Aún en rodaje	-Mejorando.
-como no-equipo	-Faltó incluir a todo el equipo en refinamientos	-Generando el hábito de refinar.
-Improvisación	-Evolución positiva de trabajo	
-Sentí desordenados los últimos sprints.		
-Poca rigurosidad en tomar historias que solo indicaban el título.		
-Improvisación		

Hay temas que se repiten en varios comentarios del equipo como la improvisación o el concepto de ir aprendiendo. Con respecto a la improvisación, la sensación de que no había un rumbo claro estaba instalada antes de aplicar la técnica. Se estaban priorizando y planificando historias que el equipo consideraba que no aportaban valor o se llegaba con historias incompletas y confusas. En cuanto al concepto de aprendizaje, generar el hábito de refinar y adaptarse a la nueva técnica es un proceso que se debe trabajar día a día y tener al equipo motivado ayuda.

Mencionar además que comentarios como “Faltó llegar con el Backlog del Sprint más detallado, priorizado antes de refinar”, “Se encontraban bastante detalladas las historias por lo cual se pudieron planificar de buena forma”, “La presentación de HDU en forma gráfica aporta bastante al entendimiento” o “Descripción gráfica hace entendible para todos” se relacionan directamente con los criterios de calidad de las HDU que se abordaron en el paso anterior del método de investigación, como la claridad y definición de las historias y el valor de negocio aportado.

VII. CONCLUSIONES

La aplicación de la técnica de refinamiento propuesta tuvo mejores resultados respecto al valor de negocio que una historia aporta. Se observó que en el sprint en que se usó la técnica inicial se consideraban muchos puntos de historia, pero con poco valor de cara a cliente. Del sprint de control, sólo una historia fue liberada a producción y la cantidad de puntos comprometidos en este sprint fueron más del doble que los de los sprints en los que se aplicó la técnica. Esta reducción en la velocidad del equipo, entendida la velocidad como la cantidad de puntos por sprint que el equipo realiza, se explica en el hecho de que había muchas historias que implicaban tareas que ahora se realizan antes de comprometer la historia. Se perdía tiempo en crear y definir UX, hacer análisis durante el transcurso del desarrollo de la misma historia o eliminar dependencias. Estas tareas ahora forman parte del refinamiento de la historia, por lo que el equipo se puede enfocar completamente en las tareas de desarrollo.

En cuanto a las dependencias, se puede mencionar que cambios de UX, análisis con resultados negativos o dependencias difíciles de resolver pueden llevar a que la historia se alargue y no se termine en el sprint como estaba propuesto. Esto además provoca que sea más difícil realizar las estimaciones considerando el grado de incertidumbre de temas o dependencias que deberían estar resueltas antes de que el equipo la tome. En el sprint de control esto quedó de manifiesto, ya que gran parte del esfuerzo del sprint fue dedicado a estas tareas.

Con respecto a la claridad de las HDU, el equipo evaluó mejor las historias en las que se aplicó la técnica propuesta versus las del sprint de control. Aquí se puede comentar que las presentaciones gráficas mediante ejemplos que acompañan a cada historia (según la técnica propuesta), clarifican la historia y ayudan a mejorar su entendimiento según la retroalimentación recibida desde el equipo con comentarios como “Descripción gráfica hace entendible para todos” o “La

presentación de HDU en forma gráfica aporta bastante al entendimiento”. Estos ejemplos gráficos son parte de la técnica propuesta de refinamiento, donde se da prioridad a la conversación por sobre la escritura de la historia en un formato extenso (como el que se utilizaba antes).

Adicionalmente, se notó que es difícil crear el hábito en el equipo de refinar a diario. Se ha visto que lo vertiginoso del sprint consume el tiempo del equipo, pero que el sacar integrantes del equipo del sprint para refinar es ganancia para la visión del producto. Esto ayudó a un mejor valor de negocio aportado por cada historia en los *sprints* en los que se aplicó la técnica propuesta, ya que se refinó sólo lo que realmente aportaba valor.

Por último, cabe hacer notar que la metodología permitió proveer evidencia asociada a la hipótesis de que la técnica propuesta, que se basa en el uso conjunto de las técnicas SbE, 3C e INVEST, mejoraría la calidad de las HDU. Aunque la metodología no permitió distinguir el efecto de cada una de las técnicas base por separado, este trabajo contribuye con evidencia empírica en un contexto organizacional sobre cómo comparar la efectividad de dos técnicas alternativas de refinamiento de HDU, justo en el momento en que la organización está en una transición hacia la agilidad en la gestión de proyectos de software. La literatura provee escasa evidencia respecto a este tipo de comparación. Existen pocos casos de estudio llevados a cabo sobre el impacto positivo de SbE en algunas organizaciones [2]. En cuanto a las otras técnicas base, no se encontró evidencia empírica asociada. Nuestros resultados dejan preguntas de investigación aún abiertas respecto a evaluar la efectividad de cada técnica por separado. Sin embargo, informan sobre la efectividad de su uso conjunto en un contexto organizacional latinoamericano.

REFERENCIAS

- [1] Ken Schwaber and Jeff Sutherland, “The Scrum Guide™. The Definitive Guide to Scrum: The Rules of the Game”, Scrum.Org and Scrum Inc., Julio 2016.
- [2] Gajko Adzic, “Specification by Example: How Successful Teams Deliver the Right Software”, Manning Publications, 2011.
- [3] Angel Lozano. Requisitos vs Casos de uso vs Historias de Usuario [En línea]. Disponible en: <http://www.angelozano.com/requisitos-del-sistema-vs-casos-uso-vs-historias-usuario/>
- [4] Mountain Goat Software. *User Stories* [En Línea]. Disponible en: <https://www.mountaingoatsoftware.com/agile/user-stories>
- [5] Alexander Menzinsky. ¿Cómo medir la calidad de las historias de usuario? [En línea]. Disponible en: <http://scrum.menzinsky.com/2016/06/como-medir-la-calidad-de-las-historias.html>
- [6] Agile Alliance. *Backlog Grooming* [En línea]. Disponible en: <https://www.agilealliance.org/glossary/backlog-grooming/>
- [7] Ron Jeffries. *Essential XP: Card, Conversation, Confirmation* [En línea]. Disponible en: <http://xprogramming.com/articles/expcardconversationconfirmation/>
- [8] Agile Alliance. *The Three C's* [En línea]. Disponible en: <https://www.agilealliance.org/glossary/three-cs>
- [9] CA Technologies. *Gartner Magic Quadrant for Enterprise Agile Planning Tools* [En línea]. Disponible en: <https://www.ca.com/us/collateral/industry-analyst-report/gartner-magic-quadrant-for-enterprise-agile-planning-tools.html>