

<https://doi.org/10.23913/ride.v14i28.1928>

Artículos científicos

***Validez de contenido de un instrumento de medición para medir el
éxito en la implementación del TPM***

***Content validity of a measurement instrument to measure success in TPM
implementation***

***Validade de conteúdo de um instrumento de medição para medir o sucesso na
implementação do TPM***

Paola Fernanda Castillo Salcido

Universidad Autónoma de Juárez, México

al209505@alumnos.uacj.mx

https://orcid.org/0000_0001_5481_186X

Rafael García Martínez

Tecnológico Nacional de México, Instituto Tecnológico de Hermosillo, México

rafael.garciam@hermosillo.tecnm.mx

<https://orcid.org/0000-0001-7175-5361>

Jesús Andrés Hernández Gómez

Universidad Autónoma de Ciudad Juárez, México

jhernand@uacj.mx

https://orcid.org/0000_0003_2325_2051

Eduardo Rafael Poblano Ojinaga

Tecnológico Nacional de México/ I T de Ciudad Juárez, México

e_poblano@yahoo.com

<https://orcid.org/0000-0003-3482-7252>

Salvador A. Noriega Morales

Universidad Autónoma de Ciudad Juárez, México

snoriega@uacj.mx

<https://orcid.org/0000-0001-7813-5835>



Resumen

La validación de contenido constituye una etapa crítica y compleja en el proceso de desarrollo de instrumentos de medición comúnmente empleados para evaluar constructos complejos en la investigación social y administrativa, ya que debe asegurarse que los ítems o variables medibles de cada constructo sean representativos. Por eso, el objetivo de esta investigación fue reportar y discutir la validación de contenido de manera práctica de un instrumento de medición que evalúa el éxito en la implementación del mantenimiento productivo total (TPM, por sus siglas en inglés) a partir de cinco constructos: participación y compromiso de la gerencia (PCG), educación y capacitación (EC), involucramiento del personal (IP), cambio de cultura organizacional (CCO) y éxito en la implementación del TPM (ETPM). Para llevar a cabo el procedimiento de validación se empleó el método de juicio de expertos y el índice de contenido de Lynn (I-CVI), así como una revisión de literatura, y recopilación y formulación de ítems que conforman los constructos, los cuales fueron validados mediante el juicio de expertos, quienes examinaron las dimensiones de relevancia, coherencia y claridad de los ítems. Las respuestas de los expertos fueron analizadas utilizando el método del índice de validez de contenido de Lynn (I-ICV, S-CVI/ave, S-CVI/UA) para cada ítem en cada una de las dimensiones. Los resultados muestran que los ítems son representativos del constructo correspondiente, lo que indica que el instrumento posee una validez de contenido muy alta con un nivel de significancia del 5%. Como resultado, se presenta un instrumento con validez de contenido para medir el éxito de la implementación del TPM a través de los cuatro constructos propuestos.

Palabras claves: validez de contenido, juicio de expertos, mantenimiento total productivo (TPM), factores críticos de éxito (FCE).

Abstract

Content validation is a critical and intricate stage in the process of developing measurement instruments often used to assess complex constructs in social and administrative research. Measurement instruments must exhibit content validity to ensure that the measurable items or variables for each construct are representative of those constructs. This study aims to report and discuss the practical content validation of a measurement instrument that assesses success in the implementation of Total Productive Maintenance (TPM) considering five constructs: Management Participation and Commitment (MPC), Education and Training (ET), Employee Involvement (EI), Organizational Cultural Change (OCC), and TPM implementation success (TPMS).



For the validation procedure, expert judgment method and Lynn's Content Validity Index (I-CVI) were utilized, involving a literature review, item collection, and formulation for the constructs. These were validated by expert judgment, where they assessed item dimensions: relevance, coherence, and clarity. Expert responses were analyzed using Lynn's Content Validity Index (I-CVI, S-CVI/ave, S-CVI/UA) method for each item within the dimensions. Results show that the items are representative of their corresponding construct, thus establishing the instrument's high-level content validity with a significance level of 5%. The outcome is a content valid instrument to measure TPM implementation success through the proposed four constructs.

Key words: Content validity, Expert judgment, Total productive maintenance (TPM), Critical Success Factors (CSF).

Resumo

A validação de conteúdo constitui uma etapa crítica e complexa no processo de desenvolvimento de instrumentos de medida comumente utilizados para avaliar construtos complexos em pesquisas sociais e administrativas, pois deve-se garantir que os itens ou variáveis mensuráveis de cada construto sejam representativos. Portanto, o objetivo desta pesquisa foi relatar e discutir de forma prática a validação de conteúdo de um instrumento de medição que avalia o sucesso na implementação da manutenção produtiva total (TPM) com base em cinco construtos: participação e comprometimento da gestão (PCG), educação e formação (CE), envolvimento do pessoal (PI), mudança de cultura organizacional (CCO) e sucesso na implementação do TPM (ETPM). Para a realização do procedimento de validação utilizou-se o método de julgamento de especialistas e o índice de conteúdo de Lynn (I-CVI), além de revisão de literatura e compilação e formulação dos itens que compõem os construtos, os quais foram validados por meio do julgamento de especialistas, que examinaram as dimensões de relevância, coerência e clareza dos itens. As respostas dos especialistas foram analisadas pelo método do índice de validade de conteúdo de Lynn (I-ICV, S-CVI/ave, S-CVI/UA) para cada item de cada uma das dimensões. Os resultados mostram que os itens são representativos do construto correspondente, o que indica que o instrumento possui validade de conteúdo muito elevada com nível de significância de 5%. Como resultado, é apresentado um instrumento com validade de conteúdo para medir o sucesso da implementação do TPM através dos quatro construtos propostos.

Palavras-chave: validade de conteúdo, julgamento de especialistas, manutenção produtiva total (TPM), fatores críticos de sucesso (FCS).

Fecha Recepción: Noviembre, 2023

Fecha Aceptación: Mayo, 2024

Introducción

El mantenimiento productivo total (TPM), una estrategia concebida por Nakajima en 1988, representa una herramienta valiosa para mejorar la calidad de los procesos productivos, ya que su enfoque hacia la mejora continua tiene como objetivo principal maximizar la productividad y prolongar el ciclo de vida del equipo industrial (Agustiady y Cudney, 2018; Pascal *et al.*, 2019). En efecto, en comparación con otras técnicas de mantenimiento —como el preventivo y correctivo—, el TPM presenta diversas ventajas, entre las que se incluyen la participación y cooperación de todos los involucrados, el desarrollo de la autonomía y habilidades técnicas de los operadores, la mejora en la comunicación entre operadores y técnicos de mantenimiento, así como la mejora en la efectividad operativa de los equipos, las capacidades del proceso y la productividad general, lo que resulta en la minimización de pérdidas de producción y costos de producción (Gupta y Vardhan, 2016; Pinto *et al.*, 2020).

Ahora bien, para evaluar la eficacia del TPM se emplean tres indicadores fundamentales: la efectividad total del equipo (OEE), el tiempo medio entre fallos y el tiempo medio de reparación (Prabowo y Adesta, 2019; San, 2021). Sin embargo, el nivel de éxito o eficacia en la implementación del TPM varía según la empresa, ya que depende de un conjunto de variables no directamente medibles, conocidas como *variables latentes* o *constructos*. En tal sentido, se ha demostrado que diversas variables latentes o constructos influyen en el éxito de la implementación del TPM, como la participación y compromiso de la gerencia (PCG) (Hooi y Leong, 2017; Mishra *et al.*, 2021; Rathi *et al.*, 2021), la educación y capacitación (EC) (Ahmad *et al.*, 2017; Ngoy y Israel, 2021), el involucramiento del personal (IP) (Agung Prabowo *et al.*, 2020; Rathi *et al.*, 2021) y el cambio en la cultura organizacional (CCO) (Farné, 2020; Mishra *et al.*, 2021).

Por lo tanto, resulta esencial abordar el desafío de identificar y medir estas variables, las cuales no siempre pueden ser cuantificadas directamente en el entorno industrial. Debido a lo anterior, surge la necesidad de desarrollar un instrumento de medición (IM) que asegure la validez de contenido, es decir, garantizar que el instrumento contenga los ítems relevantes y adecuados para medir los constructos que se pretenden evaluar. Hasta la fecha, la literatura revisada no ha reportado algún modelo que relacione los constructos PCG, EC, IP, CCO y TPM, ni tampoco algún instrumento que mida la relación entre estos constructos.

Por tal motivo, el objetivo de este trabajo de investigación es evaluar, mediante el método de juicios de expertos, la validez de contenido del IM propuesto para construir un modelo de

regresión lineal multivariante con el fin de establecer las relaciones causales entre el éxito en la implementación del TPM, como variable endógena, y los cuatro constructos PCG, EC, IP y CCO, como variables exógenas. Al respecto, cabe destacar que la validación de contenido desempeña una función crucial en el proceso de elaboración de un instrumento de medición, ya que proporciona evidencia sobre la validez de los constructos de un IM y sienta una base efectiva para la construcción de formularios (Ding y Hershberger, 2002).

Evaluar la validez de contenido del IM tiene como objetivo determinar si las preguntas propuestas o ítems del IM reflejan adecuadamente el dominio de contenido que se pretende medir, teniendo en cuenta los conocimientos, habilidades y destrezas pertinentes. El método de juicios de expertos empleado en esta evaluación consiste en determinar si existe concordancia en las calificaciones proporcionadas por los jueces respecto a cada uno de los ítems y al conjunto global de ítems, es decir, examinar si estas concordancias son atribuibles al azar desde un enfoque estadístico. Para verificar estas hipótesis se utilizan los índices de razón de validez de contenido (CVR) (Lawshe, 1975) y el índice de validez de contenido (CVI) (Lynn, 1986).

Al respecto, Almanasreh *et al.* (2019) y Wilson *et al.* (2012) señalan la preferencia por el índice CVR para probar la hipótesis de que las concordancias entre los jueces son atribuibles al azar, debido a su simplicidad en comparación con otros criterios alternativos, su facilidad de cuantificación y la disponibilidad de una tabla con los valores críticos. Sin embargo, surge un problema de monotonía con los valores críticos para el criterio de validez de estos valores críticos para el CVR. Específicamente, cuando el número de jueces expertos es igual a ocho, este valor crítico rompe la monotonía observada en el resto de estos valores. Para resolver este problema, Wilson *et al.* (2012) proponen calcular este valor crítico utilizando la ecuación (1).

$$CVR_{\alpha} = \frac{Z_{\alpha}}{\sqrt{N}} (1)$$

donde N representa el total de jueces participantes en la evaluación, y α corresponde al nivel de significancia con la que se realiza la prueba de concordancia.

Los métodos del índice de validez de contenido (CVI, por sus siglas en inglés) (Lynn, 1986) representan una alternativa al método CVR debido a su capacidad de discriminación (Romero *et al.*, 2023). Almanasreh *et al.* (2019) y Roebianto *et al.* (2023), entre otros, eligen este índice para determinar la validez de contenido del IM en sus investigaciones. En comparación con otros índices alternativos, se ha concluido que el CVI presenta ventajas en cuanto a la facilidad de cálculo, la

comprensibilidad, el enfoque en el acuerdo de relevancia en lugar del acuerdo en sí mismo, y el enfoque en el consenso en lugar de la coherencia.

Este método evalúa la concordancia entre los jueces para cada ítem del IM mediante el índice de validez de contenido individual (I-CVI), donde se sugiere que los ítems con un valor numérico del I-CVI igual o superior a 0.78, con tres o más expertos, se consideran evidencia de una buena validez de contenido (Polit *et al.*, 2007).

El CVI determina el grado de concordancia entre los jueces en dos niveles: a nivel de ítems (I-CVI) y a nivel de escala (S-CVI), que mide la concordancia entre los jueces sobre todo el instrumento de medición y se representa mediante dos índices, S-CVI/ave y S-CVI/UA.

- I-CVI: Este se emplea para evaluar la validez de contenido de manera individual para cada ítem en el instrumento. Se determina mediante la opinión de un panel de expertos que califican cada ítem en función de su relevancia, claridad y coherencia con el constructo al que pertenece. Cada ítem recibe una calificación por parte de los expertos, y el I-CVI se mide como la proporción de expertos que consideran que el ítem es relevante y adecuado. Generalmente, se establece un umbral de aceptación para el I-CVI, y los ítems que no alcanzan ese umbral se revisan o eliminan.
- S-CVI/ave (índice de validez de contenido escala promedio): Esta medida evalúa la validez de contenido de todo el IM en su conjunto. Se determina promediando los I-CVI de todos los ítems del instrumento, lo cual ofrece una evaluación integral de la calidad de los ítems en términos de su representatividad y relevancia con respecto a los constructos. Un S-CVI/ave alto indica que, en promedio, los ítems del instrumento son relevantes y adecuados para medir los constructos.
- S-CVI/UA (índice de validez de contenido universal/acumulado): Esta medida también evalúa la validez de contenido del IM en su totalidad. A diferencia del S-CVI/ave, el S-CVI/UA se calcula teniendo en cuenta la concordancia unánime entre los expertos. Es decir, un ítem se considera válido si todos los expertos están de acuerdo en que es relevante y adecuado, lo cual es útil para identificar ítems que tienen un alto nivel de consenso entre los expertos. Un valor alto del S-CVI/UA indica que todos los ítems del instrumento son considerados esenciales y apropiados por unanimidad por los expertos.

En el proceso de cálculo de la validez de contenido, se asignan valores de 1 y 0 basados en las respuestas de los expertos, utilizando una escala de Likert de 4 posiciones. Además, se asigna el valor de 1 a las respuestas que obtuvieron valores de 3 y 4, mientras que se otorga el valor de 0

a aquellas respuestas que obtuvieron 1 y 2 por parte de los expertos. Una vez que los datos han sido recodificados, se procede al cálculo de los indicadores, como se muestra en la tabla 1.

Tabla 1. Fórmulas para el cálculo del CVI

Indicador	Nombre	Definición	Formula
I-CVI	Índice de validez de contenido por ítem	Proporción de jueces que declaran esencial al ítem.	$I - CVI = \frac{k}{N}$ <p>k= # de ítems esenciales; N= #de jueces</p>
S-CVI/ave	Índice de validez de contenido para la escala, basado en el método del promedio	Media aritmética de los I-CVI de los ítems; o de la proporción de ítems calificados como esenciales por cada juez (j_i).	$S - CVI/ave = \frac{\sum_{i=1}^n (I - VCI_i)}{n}$ <p>($I - CVI_i$)= I-CVI del ítem i n= # de ítems. o</p> $S - CVI/ave = \frac{\sum_{i=1}^N j_i}{N}$
S-CVI/UA	Índice de validez de contenido para la escala, basado en el método del acuerdo universal	Proporción de ítems que son considerados relevantes.	$S - CVI/UA = \frac{\sum_{i=1}^n I_i}{n}$ <p>$I_i = 1$, si el ítem i es esencial $I_i = 0$, si el ítem i no es esencial</p>

Fuente: Yusoff (2019)

Para evaluar el I-CVI, que analiza cada ítem de manera individual, se determina la proporción de expertos que evaluaron el ítem como fundamental, es decir, aquellos que le asignaron una puntuación de 1 en su valoración. Para que el I-CVI se considere excelente debe obtener un valor igual o mayor a 0.78, según las pautas establecidas por Polit *et al.* (2007). En cuanto al cálculo del S-CVI/ave, este se obtiene como la media aritmética de los I-CVI de todos los ítems evaluados. Para que se considere aceptable, el S-CVI/ave debe ser igual o mayor a 0.90. Por último, en el proceso de evaluación de la validez de contenido, es importante reportar el índice de validez de contenido escalar/universal acumulado (S-CVI/UA), indicador que refleja la proporción de elementos que han sido evaluados como esenciales por consenso entre los expertos.

Almanasreh *et al.* (2019) establecen que los índices como Cohen Kappa's K, Fleiss Kappa (Kappa de múltiples evaluadores), Kappa ponderado, índice T, Gwet's AC1, y otros, deben interpretarse con cuidado si se utilizan con fines de evaluar la validez de contenido, dado que están diseñados y desarrollados para evaluar el acuerdo general entre un conjunto de evaluadores, pero no con el propósito de cuantificar la validez de contenido. Estos índices también pueden medir el acuerdo completo entre evaluadores independientemente del tipo de concordancia, acuerdo o desacuerdo, y ajustan el riesgo del acuerdo al azar.

Metodología

Materiales y métodos

La evaluación de la validez de contenido del IM se lleva a cabo en cuatro fases. La primera, denominada “definición”, implica la selección de los constructos y la delimitación de sus dominios para definir los ítems en función de la literatura disponible. La segunda fase, llamada “juicio de expertos”, consiste en la selección de especialistas que ayudarán a evaluar los ítems y los constructos para validar estadísticamente las evaluaciones realizadas. La tercera fase, llamada “cuantificación”, y la cuarta fase, “reestructuración”, están orientadas hacia la toma de decisiones fundamentadas en los indicadores adquiridos para determinar si se deben modificar o eliminar ciertos ítems. Finalmente, en caso de que sea necesario reestructurar el IM después de considerar el juicio de expertos, se evalúa la validez de contenido del nuevo IM resultante.

Metodología

La metodología de esta investigación se basó en un enfoque cuantitativo, con un alcance exploratorio. El diseño de la investigación fue de tipo observacional y transeccional o transversal de tipo inferencial. Se seleccionó una muestra de seis expertos mediante el método de conveniencia, los cuales poseen las siguientes características: estudios de ingeniería, al menos 30 años de experiencia laboral, tres de ellos con estudios a nivel de doctorado en el área de ingeniería industrial, dos con experiencia en la implementación de la estrategia del TPM en la industria y dos actualmente trabajando en la industria manufacturera. Estos expertos evaluaron los ítems del instrumento utilizando una escala de Likert de 4 puntos para precisar las siguientes tres características: relevancia, claridad y coherencia de los ítems del IM.

Definición

En esta fase, se determinan los constructos a través de una exhaustiva revisión de la literatura. Para ello, se buscó establecer una base teórica sólida para definir operativamente el instrumento que medirá la variable latente o constructo. Asimismo, se utilizaron diversos buscadores como ResearchGate, Sciencedirect, Emerald, Google Scholar, IEEE, entre otros, con el objetivo de identificar artículos relacionados con la implementación de la filosofía TPM, los factores críticos de éxito del TPM y los modelos estructurales del TPM (en este caso, en el periodo de 2017 a 2023). Posteriormente, se identificaron las variables latentes para definir los constructos y desarrollar los ítems correspondientes. En total, se establecieron cinco constructos: participación y compromiso de la gerencia (PCG), educación y capacitación (EC), involucramiento del personal (IP), cambio de cultura organizacional (CC), y éxito en la implementación del TPM (ETPM).

Análisis

Los datos obtenidos en el método de juicio de expertos se analizaron utilizando el índice de validez de contenido (CVI), propuesto por Lynn (1986) y respaldado por Yusoff (2019). Este método establece rangos según la cantidad de expertos y evalúa la concordancia entre los jueces con respecto a cada uno de los ítems a través del índice de validez de contenido individual (I-CVI). Además, se mide la calidad de la escala de los ítems mediante el índice de validez de contenido escalar promedio (S-CVI/ave). También se representa el índice de validez de contenido universal acumulado (S-CVI/UA), que indica el porcentaje de ítems que fueron declarados como esenciales

por decisión unánime de los expertos. Estos parámetros fueron utilizados siguiendo las recomendaciones de Almanasreh *et al.* (2019) y Saputra *et al.* (2023).

Características por evaluar

Las categorías evaluadas por los expertos para cada ítem perteneciente al instrumento de medición fueron las siguientes:

- **Relevancia:** Se refiere a la importancia del ítem y determina si es necesario eliminarlo del IM.
- **Coherencia:** Evalúa la correspondencia del ítem con la dimensión a la que pertenece.
- **Claridad:** Mide qué tan adecuada es la redacción para describir el ítem.

Criterios de aceptación

Se sume que existe concordancia entre los jueces: para cada ítem, si el valor del indicador I-CV1 es mayor o igual que 0.78 (Polit *et al.*, 2007); para el IM, si el valor del indicador S-CV1/ave, es mayor o igual que 0.90 (Naye, *et al.*, 2022; Polit *et al.*, 2007).

Evaluación del IM

Las categorías se evaluaron a través de la escala de Likert de 4 puntos en cada una de ellas, donde se considera que las calificaciones 1 y 2 indican baja importancia, mientras que las calificaciones 3 y 4 que los ítems son esenciales. A continuación, se establecen las descripciones de las ponderaciones de acuerdo a los siguientes criterios:

Relevancia

- 1. No importante: El ítem es prescindible y no causa impacto significativo en la dimensión.
- 2. Poco importante: El ítem posee cierta importancia, pero podría ser medido por otro elemento.
- 3. Importante: El ítem es considerado relevante.
- 4. Muy importante: El ítem es altamente relevante y se recomienda incluirlo.

Coherencia

- 1. No importante: El ítem carece de coherencia con la dimensión evaluada.
- 2. Poco importante: El ítem está parcialmente relacionado con la dimensión.
- 3. Importante: El ítem muestra una conexión intermedia con la dimensión.

- 4. Muy importante: El ítem está íntimamente vinculado a la dimensión que se está evaluando.

Claridad

- 1. No importante: El ítem carece de claridad y no es comprensible.
- 2. Poco importante: El ítem requiere bastantes modificaciones para mejorar su comprensión.
- 3. Importante: El ítem requiere algunas modificaciones para mejorar su claridad.
- 4. Muy importante: El ítem es claro y adecuado en su redacción.

Reestructuración del IM

Después de evaluar la concordancia entre los expertos, se efectuó un análisis para determinar si es necesario ajustar o eliminar los ítems que no se consideran apropiados. En caso de requerir modificaciones, el instrumento se somete nuevamente a una revisión por parte de los expertos, quienes pueden ser los mismos que participaron en la primera evaluación o un nuevo grupo de expertos, siguiendo el enfoque sugerido por Almanasreh *et al.* (2019).

Resultados

Definición de constructos

Se identifican cinco factores o constructos que influyen en el resultado obtenido en el proceso de implementación del TPM: participación y compromiso de la gerencia (PCG), educación y capacitación (EC), involucramiento del empleado (IP), cambio de cultura organizacional (CCO) y éxito en la implementación del TPM (ETPM). Para cada constructo, se desarrollan diversos ítems para ser evaluados, lo que resulta en un instrumento de medición con 48 ítems correspondientes a los diferentes constructos: PCG cuenta con 15 ítems, EC con 10 ítems, IP con 8 ítems, CCO con 5 ítems y ETPM con 10 ítems.

El constructo PCG constituye la variable exógena, definida como el principal sustento crítico para alcanzar el éxito en la implementación del TPM. Este incluye el compromiso y el involucramiento de la gerencia, que se traduce en apoyo para contribuir al logro de los objetivos, gestionar la implementación en un ambiente de cooperación y confianza, y dotar a los trabajadores de las herramientas necesarias, lo cual se logra a través de la comunicación y el liderazgo, que

transmiten la comprensión del enfoque, la supervisión, la capacidad de cambio y la motivación necesaria para los trabajadores (Agung Prabowo *et al.*, 2020; Hooi y Leong, 2017).

Los constructos siguientes también son definidos del siguiente modo:

- EC: Capacidad para desarrollar el conocimiento adecuado del TPM en los trabajadores mediante formación especializada según los puestos, lo cual les permite adquirir habilidades necesarias para el éxito en la etapa inicial de ejecución de la filosofía del TPM (Mishra *et al.*, 2021; Ngoy e Israel, 2021).
- IP: Colaboración del personal en todos los niveles, desde la gerencia hasta los operadores, para involucrarse en la toma de decisiones y trabajando en equipo (Pascal *et al.*, 2019; Rathi *et al.*, 2021; Zennaro *et al.*, 2019).
- CCO: Disposición para cambiar de prácticas tradicionales a prácticas nuevas en la actividad diaria, considerando el cambio de cultura organizacional (Badiea *et al.*, 2023; Rathi *et al.*, 2021).
- ETPM: Mejora en la eficiencia, reducción de costos, mejora de calidad, participación del personal, autonomía en el mantenimiento, seguridad laboral y capacidad de mantener y mejorar a lo largo del tiempo (San, 2021; Singh *et al.*, 2023; Sukma *et al.*, 2022).

Juicio de expertos y cuantificación

Se conformó un grupo de seis expertos que evaluaron los 48 ítems del IM utilizando una escala de Likert de 4 puntos para evaluar la relevancia, claridad y coherencia de cada ítem. Asimismo, se calcularon el I-CVI, el S-CVI/Ave y S-CVI/UA siguiendo el procedimiento establecido por Polit *et al.* (2007). En la primera ronda de evaluación, los expertos identificaron 17 ítems como insuficientes, de los cuales 2 fueron eliminados debido a su falta de relevancia, por lo que dejaron los 15 ítems restantes que requerían reestructuración. Esto condujo a una segunda ronda del método de juicio de expertos.

Una vez que se redefinieron los 15 ítems calificados como insuficientes, se integraron en la posición correspondiente en el IM para someterlo a una segunda ronda de evaluación con los mismos 6 expertos que participaron en la primera ronda. Al recabar nuevamente las calificaciones asignadas por los expertos para estos 15 ítems, se encontró que 9 de ellos obtuvieron una concordancia favorable, por lo que se eliminaron 6 ítems. Esto dio como resultado un IM con validez de contenido (tabla 2), ya que todos los I-CVI se situaron por encima del valor de 0.78, con valores que oscilan entre .833 y 1. Por ende, se concluyó que existe concordancia entre los expertos para cada uno de los ítems.

En cuanto al S-CVI/ave, se obtuvieron valores superiores a .90 en cada una de las categorías: relevancia con 0.988, claridad con 0.9841 y coherencia con 0.9801. Los valores del indicador S-CVI/UA para las categorías fueron 0.929 en relevancia, 0.9047 en claridad y 0.8809 en coherencia. Por último, el IM demostró tener validez de contenido, con 5 constructos desarrollados en 42 ítems (tabla 3), distribuidos de la siguiente manera: PCG con 14 ítems, EC con 9 ítems, IP con 8 ítems, CCO con 4 ítems y ETPM con 7 ítems.

Tabla 2. Evaluaciones de las características *relevancia*, *claridad* y *coherencia* del instrumento de medición

Ítem	Relevancia			Claridad			Coherencia		
	I-CVI	$\frac{I - CVI}{ave}$	$\frac{I - CVI}{UA}$	I-CVI	$\frac{I - CVI}{ave}$	$\frac{I - CVI}{UA}$	I-CVI	$\frac{I - CVI}{ave}$	$\frac{I - CVI}{UA}$
1	1	0.988	0.929	0.833	0.9841	0.9047	0.8333	0.9801	0.8809
2	1			1			1		
3	1			1			1		
4	1			1			1		
5	1			1			1		
6	1			1			1		
7	1			1			1		
8	1			1			0.833		
9	1			1			1		
10	1			1			1		
11	0.833			1			1		
12	1			1			1		
13	1			0.833			1		
14	1			1			1		
15	1			1			1		
16	1			1			1		
17	1			1			1		
18	1			1			1		
19	1			1			1		
20	1			1			1		
21	1			1			1		
22	1			1			1		
23	1			1			1		
24	1			1			1		

25	1			1			0.833 3		
26	1			1			1		
27	0.8333			0.833 3			0.833 3		
28	1			1			1		
29	1			1			1		
30	1			1			0.833 3		
31	1			1			1		
32	1			1			1		
33	1			1			1		
34	1			1			1		
35	1			1			1		
36	1			1			1		
37	1			1			1		
38	1			1			1		
39	1			1			1		
40	1			1			1		
41	1			0.833 3			1		
42	0.8333			1			1		

Fuente: Elaboración propia

Instrumento de medición

El IM que se obtuvo posteriormente a la evolución que realizaron los expertos se presenta en la tabla 3, donde se expone la lista de los ítems correspondientes a cada uno de los constructos o variables: PCG; EC; IP y CCO.

Tabla 3. Formato del IM aplicado en el juicio de expertos

Ítem	Variable latente: participación y compromiso de la gerencia	Característica por evaluar												Observaciones
		Coherencia				Claridad				Relevancia				
		1	2	3	4	1	2	3	4	1	2	3	4	
1	PCG1. Las responsabilidades dentro de nuestro programa de mantenimiento están claramente definidas en todas las secciones de la planta.													
2	PCG2. La formulación de la misión, estrategias y políticas de la empresa reflejan el compromiso de la alta gerencia con el mantenimiento del equipo.													
3	PCG 3. La gerencia comprende claramente el enfoque TPM.													
4	PCG 4. Las metas de la organización apoyan el desarrollo del programa TPM.													
5	PCG 5. Se percibe liderazgo de la dirección en la ejecución de los programas de TPM.													
6	PCG 6. La dirección apoya los objetivos de calidad con la implementación del TPM.													
7	PCG 7. Se tiene apoyo de gerencia en el programa de mantenimiento.													
8	PCG 8. La gerencia motiva a trabajar en equipo.													
9	PCG 9. La gestión de la gerencia es eficaz para ejecutar el TPM.													
10	PCG 10. Los gerentes de la organización brindan un ambiente de cooperación y confianza para la implementación del TPM.													
11	PCG 11. La alta gerencia asigna a un mentor/supervisor eficiente para la implementación del TPM.													
12	PCG 12. La dirección ejecutiva provee las herramientas indispensables para llevar a cabo la implementación del TPM.													
13	PCG 13. La organización tiene la capacidad de mejorar las habilidades de los empleados en el trabajo.													
14	PCG 14. Se tienen incentivos que motivan a los empleados.													

Ítem	Variable latente: educación y capacitación	Característica por evaluar												Observaciones
		Coherencia				Claridad				Relevancia				
		1	2	3	4	1	2	3	4	1	2	3	4	
15	EC1. El programa de capacitación adecuado ayuda a aprender sobre los beneficios del TPM.													
16	EC2. Los empleados adquieren las habilidades adecuadas para implementar el TPM después de la capacitación.													
17	EC3. Los empleados conocen las nuevas tecnologías debido a una capacitación adecuada.													
18	EC4. Antes de la implementación del TPM, se ofrece capacitación personalizada según los distintos puestos con el objetivo de adquirir los conocimientos y habilidades necesarios para un cumplimiento efectivo.													
19	EC5. La gerencia domina la metodología del TPM.													
20	EC6. La capacitación previa a la implementación del TPM muestra por qué y para qué de dicha capacitación.													
21	EC7. La capacitación previa a la implementación del TPM apoya a iniciar un cambio de mentalidad cultural en los trabajadores.													
22	EC8. La capacitación previa a la implementación del TPM apoya a mejorar el ambiente de trabajo.													
23	EC9. La capacitación previa a la implementación del TPM fomenta mayor interés de los trabajadores en la eficiencia de los equipos.													
Ítem	Variable latente: involucramiento del personal	Característica por evaluar												Observaciones
		Coherencia				Claridad				Relevancia				
		1	2	3	4	1	2	3	4	1	2	3	4	
24	IP1. Se trabaja en equipos para resolver problemas de producción y fomentar la participación de los empleados.													

25	IP2. Todos los puestos desde gerencia hasta operación se involucran en las actividades para alcanzar la implementación efectiva del TPM.																
26	IP3. Los empleados participan en la toma de decisiones.																
27	IP4. Los operadores son responsables del mantenimiento de sus máquinas.																
28	IP5. Los operadores sean responsables de inspeccionar su propio trabajo.																
29	IP6. Existe un programa que asegura la evaluación y aplicación regular de las sugerencias presentadas por los empleados.																
30	IP7. Existe un programa que garantice que la administración comunique porque las sugerencias fueron o no implementadas.																
31	IP8. Se asignan reuniones periódicas que mantienen la comunicación en los diferentes niveles.																
Ítem	Variable latente: cambio cultura organizacional	Característica por evaluar															
		Coherencia				Claridad				Relevancia				Observaciones			
		1	2	3	4	1	2	3	4	1	2	3	4				
32	CCO1. Considero que la implementación del TPM es claramente efectiva para mejorar las prácticas de mantenimiento en la industria.																
33	CCO2. El cambio cultural de los trabajadores es importante en el éxito de la implementación del TPM.																
34	CCO3. El éxito de la implementación del TPM depende en gran parte de la cultura organización de cada empresa.																
35	CCO4. La implementación del TPM contempla más herramientas de mejoramiento que otros programas.																
	Ítems por dimensión o constructo a evaluar	Característica por evaluar															
	Éxito en la implementación del TPM Qué tan de acuerdo estoy con que...	Coherencia				Claridad				Relevancia				Observaciones			

		1	2	3	4	1	2	3	4	1	2	3	4	
36	ETPM1. El TPM desarrolla habilidades laborales competitivas													
37	ETPM2. El TPM mejora la calidad del producto final.													
38	ETPM3. La cantidad de rechazos y re-trabajos se reducen con la implementación del TPM.													
39	ETPM4. La implementación del TPM reduce costos.													
40	ETPM5. La implementación del TPM incrementa la efectividad operativa de los equipos de planta.													
41	ETPM6. La implementación del TPM incrementa la mejora continua.													
42	ETPM7. La implementación del TPM reduce la cantidad de accidentes en el área de trabajo.													

Fuente: Elaboración propia

Discusión

El instrumento de medición (tabla 3) demuestra validez de contenido en las tres características evaluadas: relevancia, claridad y coherencia. Esto se evidencia en el análisis de los datos obtenidos de la evaluación de los expertos, cuyos resultados numéricos superan los valores críticos establecidos para los indicadores I-CVI (≥ 0.78), S-CVI/ave (≥ 0.90) y S-CVI/UA (≥ 0.90) (Polit *et al.*, 2007). La primera ronda de evaluación no proporcionó resultados satisfactorios, lo que motivó la realización de una segunda ronda con el mismo panel de expertos para alcanzar los indicadores correctos.

Conclusiones

El instrumento de medición presentado en este estudio de investigación ha demostrado validez de contenido, lo que quiere decir que los ítems asignados a cada constructo lo representan de manera adecuada y completa. Estos ítems son altamente relevantes, claros en su formulación y coherentes en su relación con los conceptos que representan dentro del ámbito de definición de cada constructo. Por lo tanto, este instrumento es apto y válido para la recopilación de datos necesarios para realizar análisis estadísticos destinados a evaluar la validez estadística de las hipótesis relacionadas entre los cinco constructos considerados. Al respecto, cabe recordar que la validación de contenido se basa en el uso de criterios confiables, como los indicadores I-CVI, S-CVI/ave y S-CVI/UA, que son criterios sólidos para estimar la validez de contenido.

Futuras líneas de investigación

Los índices Kappa modificados y la prueba de concordancia de Kendall, propuestos por Polit *et al.* (2007), evalúan la concordancia entre los jueces para las calificaciones asignadas al IM, pero no determinan esta concordancia para cada ítem individualmente. Por lo tanto, no es posible determinar si la discrepancia observada en cada ítem se debe al azar (Almanasreh *et al.*, 2019). A pesar de su incapacidad para discriminar los efectos de la aleatoriedad en la evaluación de la concordancia en el juicio de los expertos, estos métodos aún se aplican para validar el contenido de los IM (Carlton *et al.*, 2022; Feng *et al.*, 2022; Ma *et al.*, 2023; Reick *et al.*, 2023). Es decir, desde un punto de vista estadístico, los resultados obtenidos al evaluar la validez de contenido de un IM con estos métodos son cuestionables.

Por otro lado, los índices CVR y CVI se utilizan para determinar si la concordancia de los jueces para el IM y para cada ítem es atribuible al azar (Alqahtani *et al.*, 2023). Según la teoría, estos dos últimos métodos son superiores a los dos primeros. Ante estas dos alternativas disponibles para evaluar la validez de contenido del IM, se recomienda realizar un estudio de simulación para comparar la eficacia de estos cuatro indicadores en la evaluación de la concordancia para el IM y para los ítems.

Agradecimientos

Agradecimiento a la SEP-SES por el apoyo a Paola Fernanda Castillo Salcido, actual becaria, a través del Programa Apoyo de Estancia Reducida Nacional para Estudios de Doctorado.

Rafael García Martínez, recibió apoyo económico para la investigación, autoría y/o publicación de este artículo por parte del Consejo Nacional de Humanidades Ciencia y Tecnología (CONAHCYT) a través del Programa de Estancias Posdoctorales en México 22 (1).

Referencias

- Agung Prabowo, H. A., Hilmy, I. and Triblas Adesta, E. Y. (2020). Total productive maintenance implementation's barriers and enablers in Indonesian manufacturing companies. *2020 Advances in Science and Engineering Technology International Conferences, ASET 2020*, 1-6. <https://doi.org/10.1109/ASET48392.2020.9118224>
- Agustiady, T. K. and Cudney, E. A. (2018). Total productive maintenance. *Total Quality Management and Business Excellence*, 0(0), 1-8. <https://doi.org/10.1080/14783363.2018.1438843>
- Ahmad, N., Hossen, J. and Ali, S. M. (2017). Improvement of overall equipment efficiency of ring frame through total productive maintenance: a textile case. *International Journal of Advanced Manufacturing Technology*, 94(1-4), 239-256. <https://doi.org/10.1007/s00170-017-0783-2>
- Almanasreh, E., Moles, R. and Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 15(2), 214-221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Alqahtani, T. M., Yusop, F. D. and Halili, S. H. (2023). Content validity of the Constructivist Learning in Higher Education Settings (CLHES) scale in the context of the flipped classroom in higher education. *Humanities and Social Sciences Communications*, 10(1), 1-12. <https://doi.org/10.1057/s41599-023-01754-3>
- Badiae, A. M., Wael, B. A. and Adel, A. A. (2023). Implementing Total Productive Maintenance in Yemeni Company for Industry and Commerce (YCIC): Case Study. *European Journal of Applied Sciences*, 11(2). <https://doi.org/10.14738/aivp.112.12966>
- Carlton, J., Powell, P. A. and Project HERCULES Carer Group. (2022). Measuring carer quality of life in Duchenne muscular dystrophy: a systematic review of the reliability and validity of self-report instruments using COSMIN. *Health and Quality of Life Outcomes*, 20(1), 57. <https://doi.org/10.1186/s12955-022-01964-4>
- Ding, C. S. & Hershberger, S. L. (2022). Assessing Content Validity and Content Equivalence

- Using Structural Equation Modeling, *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 283-297. https://doi.org/10.1207/S15328007SEM0902_7.
- Farné, S. (2020). Application and Results of the TPM Methodology in a Factory for the Production of Agricultural and Earth-Moving Machines. *Asian Journal of Basic Science & Research*, 2(4), 1–11. <https://doi.org/DOI: 10.38177/AJBSR.2020.2401>
- Feng, D., Huang, S., Lang, X., Liu, Y. and Zhang, K. (2022). Developing an exercise attitudes and behavior intentions questionnaire for survivors of aortic dissection: an exploratory factor analysis. *Reviews in Cardiovascular Medicine*, 23(10), 337. <https://doi.org/10.31083/j.rcm2310337>
- Gupta, P. and Vardhan, S. (2016). Optimizing OEE, productivity and production cost for improving sales volume in an automobile industry through TPM: A case study. *International Journal of Production Research*, 54. 2976-2988. <https://doi.org/10.1080/00207543.2016.1145817>
- Hooi, L. W. and Leong, T. Y. (2017). Total Productive Maintenance and Manufacturing Performance Improvement. *Journal of Quality in Maintenance Engineering*, 23(1). 2-21. <http://dx.doi.org/10.1108/JQME-07-2015-0033>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- Lynn, M. R. (1986). Determination and Quantification Of Content Validity. *Nursing Research* 35(6), 382-386. <https://doi.org/10.1097/00006199-198611000-00017>
- Ma, H., Hu, Z., Wang, Z., Ma, Y., Zhai, H. and Li, J. (2023). Development and Validation of a New Comprehensive Assessment Scale of HIV-Related Psychological Distress in China: A Methodological Study. *Journal of the Association of Nurses in AIDS Care*, 34(3), 292-306. <https://doi.org/10.1097/JNC.0000000000000402>
- Mishra, R. P., Gupta, G. and Sharma, A. (2021). Development of a Model for Total Productive Maintenance Barriers to Enhance the Life Cycle of Productive Equipment. *Procedia CIRP*, 98, 241–246. <https://doi.org/10.1016/j.procir.2021.01.037>
- Nakajima, S. (1988). *Introduction to Total Productive Maintenance (TPM)*. Cambridge, Massachusetts Productivity Press.
- Naye, F., Décar, S. and Tousignant-Laflamme, Y. (2022). Development and content validity of a rating scale for the pain and disability drivers management model. *Archives of Physiotherapy*, 12(1), 1-11. <https://doi.org/10.1186/s40945-022-00137-2>
- Ngoy, K. R. and Israel, K. (2021). The Strategy of Successful Total Productive Maintenance (TPM): Implementation and Benefits of TPM (Literature Review). *Ijirmmps*, 9(6), 43–52.

www.ijrmps.org

- Pascal, V., Toufik, A., Manuel, A., Florent, D. and Frédéric, K. (2019). Improvement indicators for Total Productive Maintenance policy. *Control Engineering Practice*, 82(September 2018), 86–96. <https://doi.org/10.1016/j.conengprac.2018.09.019>
- Pinto, G., Silva, F. J. G., Baptista, A., Fernandes, N. O., Casais, R. and Carvalho, C. (2020). TPM implementation and maintenance strategic plan - A case study. *Procedia Manufacturing*, 51(2020), 1423–1430. <https://doi.org/10.1016/j.promfg.2020.10.198>
- Polit, D. F., Beck, C. T. and Owen, S. V. (2007). Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Res Nurs Health*, 30(4), 459-67. <https://doi.org/10.1002/nur>.
- Prabowo, H. A., & Adesta, E. Y. T. (2019). A study of total productive maintenance (TPM) and lean manufacturing tools and their impact on manufacturing performance. *International Journal of Recent Technology and Engineering*, 7(6), 39–43.
- Rathi, R., Singh, M., Sabique, M., Al Amin, M., Saha, S. and Hari Krishnaa, M. (2021). Identification of total productive maintenance barriers in Indian manufacturing industries. *Materials Today: Proceedings*, 50, 736-742. <https://doi.org/10.1016/j.matpr.2021.05.222>
- Reick, S., Burckhardt, M., Palm, R. and Kottner, J. (2023). Measurement instruments to evaluate diaper dermatitis in children: Systematic review of measurement properties. *Nursing Open* 10(9). <https://doi.org/10.1002/nop2.1863>
- Roebianto, A., Savitri, S. I., Aulia, I., Suciñana, A. and Mubarokah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *TPM*, 30(1), 5-18. <http://dx.doi.org/10.4473/TPM30.1.1>
- Romero, M., Díaz Costa, E. and Faouzi Nadim, T. (2023). A Review of Lawshe's Method for Calculating Content Validity in the Social Sciences. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1271335>
- San, S. (2021). A Systematic Literature Review of Total Productive Maintenance On Industries. *Performa: Media Ilmiah Teknik Industri*, 20(2), 97. <https://doi.org/10.20961/performa.20.2.50087>
- Saputra, A., Kusumawardhani, A. A. A. A., Elvira, S. D. and Wiguna, T. (2022). An item development, content validity, and feasibility study towards the Indonesian recovery scale for patients with schizophrenia. *Heliyon*, 8(11). <https://doi.org/10.1016/j.heliyon.2022.e11826>
- Singh, S., Kumar, M. andamp; Kumar, S. (2023). Implementation of Total Productive Maintenance to Improve Productivity of Rolling Mill. *Indian Journal of Engineering & Materials Sciences*,

30(december), 882–890. <https://doi.org/10.56042/ijems.v30i6.3158>

- Sukma, D. I., Prabowo, H. A., Setiawan, I., Kurnia, H. andamp; Fahturizal, I. M. (2022). Implementation of Total Productive Maintenance to Improve Overall Equipment Effectiveness of Linear Accelerator Synergy Platform Cancer Therapy. *International Journal of Engineering, Transactions A Basics*, 35(7), 1246–1256. <https://doi.org/10.5829/ije.2022.35.07a.04>
- Wilson, F. R., Wei, P. and Donald, A. S. (2012). Recalculation of the Critical Values for Lawshe's Content Validity Ratio. *Measurement and Evaluation in Counseling and Development*, 45(3). <https://doi.org/10.1177/0748175612440286>
- Yusoff, M. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal*, 11(2), 49–54. <https://doi.org/10.21315/eimj2019.11.2.6>
- Zennaro, I., Battini, D., Sgarbossa, F. andamp; Persona, A. (2019). *TPM Implementation in automated flow line manufacturing systems : criticalities and key factors to support a faster implementation*. Conference: Summer school Francesco Turco.

Rol de contribución	Autor (es)
Conceptualización	Paola Fernanda Castillo Salcido
Metodología	Paola Fernanda Castillo Salcido
Software	Rafael García Martínez
Validación	Paola Fernanda Castillo Salcido, «igual» Rafael García Martínez
Análisis formal	Paola Fernanda Castillo Salcido
Investigación	«Principal» Paola Fernanda Castillo Salcido, «igual» Rafael García Martínez, «que apoya» Jesús Andrés Hernández Gómez, «que apoya » Eduardo Rafael Poblano Ojinaga, «que apoya» Salvador A. Noriega Morales.
Recursos	Paola Fernanda Castillo Salcido
Curación de datos	Paola Fernanda Castillo Salcido
Escritura - Preparación del borrador original	Paola Fernanda Castillo Salcido
Escritura - Revisión y edición	«principal» Paola Fernanda Castillo Salcido, «igual» Salvador A. Noriega Morales.
Visualización	«principal» Paola Fernanda Castillo Salcido, «igual» Rafael García Martínez, «igual» Jesús Andrés Hernández Gómez, «igual» Eduardo Rafael Poblano Ojinaga, «igual» Salvador A. Noriega Morales.
Supervisión	Paola Fernanda Castillo Salcido
Administración de Proyectos	«igual» Paola Fernanda Castillo Salcido, «igual» Rafael García Martínez, «igual» Jesús Andrés Hernández Gómez,
Adquisición de fondos	Paola Fernanda Castillo Salcido