

UNIVERSIDAD DE VALPARAÍSO

**PROPUESTA DE EVALUACIÓN DEL DESEMPEÑO DE TÉCNICOS Y
ASISTENTES TÉCNICOS EN UNA EMPRESA TABACALERA**

TESIS PRESENTADA A LA FACULTAD DE MEDICINA

**PARA OPTAR AL TÍTULO DE
PSICÓLOGO**

**Y AL GRADO DE
LICENCIADO EN PSICOLOGÍA**

ESCUELA DE PSICOLOGÍA

**POR
KARIN KIRSINGER POZO**

**PROFESOR PATROCINANTE
BIANCA DAPELO PELLERANO**

**VALPARAÍSO, CHILE
JUNIO, 2004**

A mi familia y profesores

TABLA DE CONTENIDOS

RECONOCIMIENTOS	ii
TABLA DE CONTENIDOS	iii
LISTA DE FIGURAS	v
LISTA DE TABLAS	vi
RESUMEN	vii
	Pág.
INTRODUCCIÓN	1
PROBLEMA	4
Formulación del Problema	4
Objetivos de investigación	13
MARCO TEÓRICO	14
Gestión de Recursos Humanos y Evaluación de Desempeño	14
Desempeño Laboral	18
Evaluación de Desempeño	30
Evaluación y Gestión del Desempeño	76
MÉTODO	81
Tipo de Estudio	81
Diseño de la Investigación	81
Universo y Muestra	81
Definición de Variables	85
Técnicas y Procedimientos de Recolección de Datos	91
Técnicas de Análisis de Datos	101
RESULTADOS	103
Necesidades de Evaluación	103
Propiedades Psicométricas del Instrumento	114
Evaluación del Usuario: Aceptabilidad	122
DISCUSIÓN	129
Resultados	129
Recomendaciones	134
Líneas de Investigación	137
Limitaciones y Fortalezas de la Investigación	137

REFERENCIAS.....	139
APÉNDICES.....	143
Apéndice A. Descripción de la Empresa y de su Gestión de Recursos Humano	144
Apéndice B. Descripción de Puestos	154
Apéndice C. Pauta de Entrevistas	161
Apéndice D. Escala de Observación de Comportamientos	168
Apéndice E. Dimensiones, Indicadores y Reactivos	173
Apéndice F. Síntesis Propuesta del Sistema de Evaluación	181
Apéndice G: Sistema de Evaluación: Manual del Evaluador	184
Apéndice H: Cuadernillo del Instrumento y Otros Formularios	200
Apéndice I: Resultados Complementarios del Diagnóstico de Necesidades	213
Apéndice J: Juicio de Experto	220
Apéndice K. Puntajes Brutos, por Dimensión y por Ítem: Resultados y Conclusiones	234
Apéndice L. Otros Resultados del Instrumento	275
Apéndice J. Definiciones de Categorías de Análisis	277

LISTA DE FIGURAS

N°.		Pág.
1.	Proceso de la Acción de Hacker	20
2.	Teoría de la Conducta Organizacional de Naylor, Pritchard e Ilgen	23
3.	Modelo de Desempeño de Campbell	26
4.	Comparación entre la Evaluación de Desempeño Evaluativa y Evolutiva	32
5.	Determinación de Medidas para Problemas de Desempeño	80

LISTA DE TABLAS

N°.		Pág.
1.	Composición de la Muestra del Personal Entrevistado	82
2.	Composición de la Muestra del Personal Evaluado	85
3.	Preguntas y Contenidos de la Entrevista de Diagnóstico de Necesidades	92
4.	Indicadores y Reactivos por Dimensión	95
5.	Determinación de Categorías del Método de Distribución Forzada	99
6.	Contenido y Preguntas de la Entrevista de Aceptabilidad	100
7.	Objetivos Generales	103
8.	Objetivos Específicos	103
9.	Resultados a Medir	104
10.	Dimensiones de Desempeño a Evaluar	105
11.	Ponderación de Criterios	106
12.	Métodos de Evaluación	107
13.	Evaluadores Potenciales	108
14.	Especificaciones Generales	109
15.	Especificaciones de Propiedades de las Medidas de Evaluación	109
16.	Especificaciones del Sistema de Evaluación	110
17.	Especificaciones del Proceso de Evaluación	111
18.	Especificaciones del Proceso de Diseño	112
19.	Ítems Eliminados, Agregados y Revisados por Jueces Externos	115
20.	Ítems Mantenidos y Eliminados por Juez Interno	116
21.	Correlación Ítem - Test	117
22.	Correlación Ítem – Dimensión	119
23.	Correlación Dimensión – Test	120
24.	Correlación del Instrumento con Métodos de Evaluación de Grupos	121
25.	Confiabilidad por Dimensión	121
26.	Percepción General de Utilidad	122
27.	Percepción General de Practicalidad	123
28.	Percepción de General de Confiabilidad	124
29.	Percepción General de Validez	125
30.	Percepción General de Discriminabilidad	125
31.	Percepción General de Libertad de Influencia de Sesgos	126

RESUMEN

El siguiente estudio exploratorio-descriptivo aborda el diseño de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Compañía Chilena de Tabacos, la cual se orienta a satisfacer los requisitos psicométricos de validez y confiabilidad así como las necesidades de los usuarios potenciales. Se expone detalladamente el proceso de diseño que abarca desde la fase de diagnóstico de necesidades de evaluación hasta la fase de evaluación que el usuario hace del instrumento. Cabe acotar que los resultados señalan las necesidades de evaluación en el nivel de gerencia, jefes y técnicos/asistentes técnicos e indican que el instrumento desarrollado resulta válido, confiable para la población determinada así como aceptable para el evaluador. Es importante destacar que la discusión se centra principalmente en estos últimos aspectos así como contiene recomendaciones que orientan la aplicación y validación futura de este instrumento así como el desarrollo de otras medidas de desempeño.

INTRODUCCIÓN

La importancia de la industria en el desarrollo de nuestro país es un tema indiscutible. En lo que va corrido del año, la rama de actividad económica que más aportó al Producto Interno Bruto (PIB) fue la Industria Manufacturera (INE, 2004). Del mismo modo, en el trimestre enero - marzo 2004, la Industria Manufacturera fue uno de los sectores más dinámicos en cuanto a la generación de empleo en el ámbito nacional, presentando un incremento de 1,4% (10.910 plazas) respecto a igual trimestre anterior (INE, 2004). Si bien nuestra región presentó, en el trimestre octubre- diciembre del 2003, una variación positiva del 2,1% en el índice de actividad económica regional (INACER) y una menor tasa de desocupación, su Industria Manufacturera presentó un decrecimiento en 1,2% en la actividad y en la tasa de ocupación respecto a igual trimestre anterior. No obstante, en la V Región, ésta sigue siendo la rama de actividad que más aporta a la actividad económica así como concentra la mayor parte de la inversión extranjera. De hecho, el sector Alimentos, Bebida y Tabaco capitaliza el 98% de la inversión en el trimestre octubre-diciembre 2003 (INE, 2003)

La importancia de la gestión del recurso humano, en la actualidad, también es una aseveración incuestionable para las todas las empresas. Podríamos decir, por ende, que la cuestión ya no es concientizar a los directivos sobre este tema sino diseñar, implementar y evaluar políticas y prácticas de recursos humanos que logren equilibrar las necesidades de la empresa y las de sus miembros. Hoy más que nunca, en este sentido, la Psicología del Trabajo y de las Organizaciones tiene mucho que aportar.

La evaluación de desempeño es una de las prácticas más tradicionales dentro de la gestión de recursos humanos, En el siglo III d.c., el filósofo Sun Yi ya documentaba la existencia de formatos de evaluación basados en méritos; en la Edad Media, los jesuitas acostumbraban evaluar a sus misioneros con informes y notas de actividades; y en el siglo XIX, Robert Owen, un pionero industrial de Escocia en relación con la educación de los trabajadores, horas laborales y manejo de la conducta de los empleados;

diseñó un registro en el cual cada supervisor debía hacer anotaciones diarias (Chon y Davico, 1989, Donkin, 2002). No obstante, a pesar de la larga data de la evaluación de desempeño y su universalidad, el diseño y la aplicación de sistemas de calidad aún es un desafío, ya que aún cuando existen en la mayoría de las organizaciones, estos no satisfacen las necesidades de evaluadores y evaluados. En la misma línea, Schvarstein en Páez (2001) plantea que: “Dado el carácter paradójico de los procesos de evaluación de desempeño en las organizaciones: es imposible hacerla bien y por otro lado, no se puede dejar de hacer”. Las empresas deben evaluar el desempeño de sus empleados en forma sistemática para mejorar su propio desempeño mediante la toma de decisiones administrativas o delimitando acciones para mejorar el desempeño individual.

La presente investigación se centra en una propuesta de evaluación de desempeño de técnicos y asistentes técnicos de una empresa manufacturera transnacional que se ubica en la V Región. Los grandes apartados de este proyecto se describen a continuación:

El primer apartado aborda el problema de investigación. En este, se plantea que las organizaciones necesitan evaluar el desempeño de los empleados en todos los niveles jerárquicos así como se describe el caso de Compañía Chilena de Tabacos (Chiletabacos), una empresa dedicada a la elaboración de cigarrillos que necesita evaluar el desempeño de su personal de línea. Posteriormente, se presentan los objetivos relativos a las propiedades que presenta la propuesta de evaluación.

El segundo acápite, Marco Teórico, desarrolla los antecedentes teóricos y empíricos en el área de la evaluación de desempeño. Para favorecer la comprensión, en primera instancia, se entrega una revisión muy acotada del rol de la evaluación de desempeño dentro de la Gestión de Recursos Humanos y la Psicología de las Organizaciones. A continuación, se presentan algunas definiciones de desempeño y teorías asociadas para posteriormente caracterizar la evaluación de desempeño en cuanto a sus objetivos, métodos, evaluadores, etc. Finalmente, se detallan las propiedades que debe tener todo sistema de evaluación y el papel de la evaluación dentro de una gestión del desempeño.

La tercera parte desarrolla la metodología que se emplea en este estudio, el cual se constituye como una investigación exploratoria-descriptiva, de diseño no experimental-transversal y que utiliza una muestra no probabilística. Cabe resaltar que el diseño de la propuesta de evaluación del desempeño aborda desde el diagnóstico de necesidades de evaluación hasta la evaluación que el usuario hace del instrumento mismo. Dado lo anterior, se recurren a variadas técnicas de recolección y análisis de datos, haciendo especial énfasis en el proceso asumido para desarrollar la propuesta de evaluación.

El apartado “Resultados” presenta los resultados del diagnóstico de necesidades, de la aplicación del instrumento en cuanto a las propiedades psicométricas y a la evaluación del usuario.

En el acápite “Discusión”, se explica tentativamente algunos resultados y se entregan sugerencias prácticas para las organizaciones e investigadores que necesiten implementar un sistema de evaluación así como se describen las posibles proyecciones, limitaciones y fortalezas de esta investigación.

En el último apartado, se presenta una serie de apéndices que incluyen una caracterización de la empresa y su gestión de recursos humanos, la descripción de cargo de técnicos y asistentes técnicos del Área de Operaciones, las pautas de entrevistas empleadas, los resultados y análisis complementarios, el cuadernillo de respuestas, instrumento y manual de evaluación, entre otros.

PROBLEMA

Formulación del Problema

El panorama actual presenta grandes desafíos a las empresas chilenas. La productividad, el control de gastos y la recuperación de márgenes de utilidad, la generación de valor y el desarrollo de la calidad en el servicio así como la fidelización de clientes representan desafíos que todas nuestras empresas deben afrontar en el marco de una economía globalizada. En este panorama, los individuos como recurso organizacional han cobrado mayor importancia. Encontramos así una preocupación creciente sobre el desempeño laboral en las organizaciones y ciencias afines.

Para la Psicología del Trabajo y de las Organizaciones, el desempeño laboral es uno de los principales focos de interés. Ahora bien, si la psicología como ciencia se encarga del estudio científico de la conducta humana en cuanto análisis, explicación, predicción y control, la Psicología del Trabajo y de las Organizaciones se encarga de estudiar la conducta en el contexto de trabajo (Rodríguez, 1998). El desempeño laboral, por tanto, es parte de su objeto de estudio al ser conceptualizado como el conjunto de conductas que presentan los individuos en el contexto de trabajo que contribuyen al cumplimiento de las metas organizacionales así como al buen funcionamiento de la organización (Ramos, Gracia y Peiró, 1996).

Para las organizaciones, por su parte, el desempeño laboral es el resultado del procesamiento y conversión que se efectúa en el recurso humano de modo de producir un bien o servicio (Cummings y Schawb, 1985). El desempeño de los individuos, a su vez, determina parcialmente el desempeño organizacional, por tanto, la supervivencia y desarrollo de las organizaciones mismas. Si se considera que un trabajador en un contexto productivo puede ser hasta tres veces más productivo que otro empleado, el desempeño laboral representa también para las organizaciones una gran oportunidad de mejora, ya que éstas pueden actuar así sobre factores internos que afectan su productividad (Ramos et. al., 1996).

Es así como el estudio del desempeño es especialmente relevante para las organizaciones, la Psicología del Trabajo y de las Organizaciones no sólo por su valor teórico sino también social. Las implicancias sociales son claras: al estudiar el desempeño laboral, podemos mejorarlo. En la misma línea, Ramos et. al. (1996) establece que en el quehacer profesional del psicólogo laboral: “...uno de los objetivos debe ser la mejora del desempeño; y por ello uno de los retos de la Psicología del Trabajo es demostrar qué contribuye a un mayor y/o mejor desempeño” (pp. 283).

La Psicología del Trabajo y de las Organizaciones busca mejorar el desempeño que presentan los individuos con el objeto de mejorar el desempeño organizacional. De esta manera, la efectividad de las intervenciones y técnicas que surgen desde la Psicología es medida en función de la capacidad para determinar, predecir o contribuir a la mejora del desempeño. (Ramos et. al., 1996). A su vez, la Psicología del Trabajo y de las Organizaciones, mediante la mejora de la productividad, contribuye también a generar una mayor calidad de vida y bienestar para la población toda, lo cual constituye el objetivo último y común denominador de las distintas áreas de la Psicología. (Rodríguez, 1996).

“Esto representa una mejora en la calidad de los bienes básicos y en la variedad y cantidad de bienes de que dispone una persona para poder optar entre diversas posibilidades, por ejemplo, en relación con la vivienda, el vestido o los alimentos. Las aspiraciones humanas también abarcan el deseo de un entorno más sano y limpio, actividades culturales, la capacidad de disponer de tiempo libre y de utilizarlo de una manera agradable y unos ingresos que hagan posible sufragar estas necesidades. A fin de que una sociedad o nación pueda elevar el nivel de vida de su población, tendrá que aumentar al máximo el rendimiento de sus recursos o mejorar la productividad para que la economía crezca y sea capaz de sostener una mejor calidad de vida” (Kanawaty, 1996, p. 3)

La evaluación del desempeño laboral, para las organizaciones, representa una de las herramientas que deben usar para optimizar el desempeño organizacional. En esta línea, Deloitte y Touche (2002) concluyeron que las prácticas de evaluación del desempeño determinan sustancialmente el valor de mercado de las empresas¹. La evaluación de desempeño tiene múltiples recursos para lograr este objetivo por cuanto permite:

¹ El valor de mercado de las empresas como indicador representa la porción del valor de una empresa que no es explicada por activos físicos o financieros. La evaluación de desempeño junto con la fijación de salarios,

1. Como proceso de medición y control del desempeño: en primer lugar, mejorar el desempeño de individuos en sus cargos al: (a) detectar fortalezas y debilidades en éste, lo cual facilita la determinación de las acciones de mejora más pertinentes, (b) proveer información al propio empleado sobre lo que se espera de él y su desempeño de modo de mantener o modificar éste, (c) si las recompensas organizacionales son fijadas en función del desempeño, éstas son atractivas para el empleado y éste percibe que puede mejorar su desempeño y que la evaluación detectará las mejoras en el mismo, incrementar los niveles de motivación (Chiavenato, 2002) . En segundo lugar, permite mejorar la toma de decisiones administrativas de colocación al destinar individuos a nuevos cargos que aseguren un mayor ajuste del individuo al puesto de trabajo.

2. Como proceso de medición y control de prácticas de recursos humanos, permite mejorar la efectividad de los procesos de selección, compensación y capacitación, entre otros así como orientar las decisiones de la dirección sobre estructuras organizativas.

Para el miembro de una organización, por su parte, el evaluar el desempeño permite conocer las reglas del juego por cuanto clarifica las expectativas, las recompensas y castigos que recibirá así como la visión que el directivo tiene acerca de su desempeño. Se satisface así la necesidad de feedback. El ambiente laboral se percibe en forma ambigua y como una fuente de tensión y angustia si no se recibe feedback (Cummings et al., 1985). Además, según Schvarstein en Páez (2001): "... ser justamente evaluado es una expectativa ligada a una necesidad básica del ser humano, cual es la necesidad del reconocimiento". No hay duda alguna que la evaluación de desempeño positiva para un empleado, contribuye a la satisfacción de sus necesidades de éxito y competencia (Barnett, 1997).

Pese a la importancia de la evaluación de desempeño para la optimización de la gestión del recurso humano, no todas las empresas manejan sistemas formales de evaluación para todos sus miembros. Desde la década del setenta, a nivel internacional, se ha podido observar una tendencia descendiente a evaluar el desempeño de empleados de menor nivel jerárquico (Quijano, 1994). En nuestro país, la evaluación de desempeño ha sido usada principalmente por las grandes empresas para evaluar a los empleados que ocupan cargos directivos o que tienen un perfil gerencial, postergándose así la gestión del desempeño de los empleados que ocupan cargos de menor nivel dentro de la escala jerárquica (Novick en Aguirre, Toro & Ruz, 1998)

En los últimos años, no obstante, esta tendencia se ha ido revertiendo paulatinamente. Las empresas han tomado conciencia de la importancia de evaluar al personal de menor jerarquía como una manera de asegurar y optimizar sus resultados operacionales. De esta manera, el nuevo centro de la gestión de recurso humano es: "... el personal de línea que está en contacto directo con el cliente o en cadena primaria del negocio..." (Kanawaty, 1996)

Cabe considerar que, en nuestro país, los profesionales, técnicos y afines representan el segundo grupo de ocupación con mayor crecimiento después de Gerentes, Administradores y Directivos. En el presente trimestre del año pasado, este grupo acumuló un crecimiento del 6,9% (39.080 personas) (INE, 2003).

Cualquier organización de nuestro país que actualmente no posee un sistema de evaluación de desempeño para el personal de línea, pero que muestre interés en evaluar puede optar por utilizar un sistema ya existente o desarrollar un sistema de evaluación.

Las organizaciones que optan por la primera alternativa, se encuentran con la ausencia de sistemas de evaluación de desempeño para personal de línea en empresas productivas. Con frecuencia, sólo se puede acceder a formatos de evaluación (instrumentos) que son extranjeros o a formatos de evaluación diseñados para otros puestos de trabajo, especialmente de tipo administrativo o directivo. De todas maneras, rara vez se llega a disponer de información sobre la confiabilidad, validez, aceptabilidad, etc. de

los instrumentos mismos, siendo altamente probable que la organización se apliquen sin llegar a efectuar un estudio de validación.

En cuanto al desarrollo de sistemas de evaluación, las organizaciones usualmente diseñan sus propios sistemas o externalizan esta tarea en consultoras. Sin importar la alternativa seleccionada, una primera dificultad para desarrollar sistemas de evaluación estriba en que la evaluación de desempeño sigue siendo un área de estudio incipiente en la cual no existe acuerdo sobre los elementos principales que deben orientar el diseño de sistemas de evaluación. En efecto, se plantea que: “La evaluación de desempeño está lejos de ser un área en que, por un lado, desde la teoría, exista unanimidad en los autores respecto de cuáles son sus reales sentido y alcance, y por otro, desde el campo de la práctica administrativa, no sea objeto de diferencias y conflictos al interior de las empresas” (Barnett, 1997, pp.3). En el plano internacional, respondiendo a esta dinámica según Hansen (1994), se puede observar actualmente grandes diferencias en sistemas de evaluación de ingenieros, técnicos y personal de línea. Se suma a esto, la falta de investigación nacional sobre las condiciones necesarias, tanto en el diseño como la en la implementación, que asegure el cumplimiento de los requisitos psicométricos y de utilización. De hecho, la mayor parte de la investigación se ha desarrollado en Estados Unidos, ya que la legislación laboral de este país obliga a las organizaciones a aplicar sistemas de evaluación confiables y válidos.

Si las ciencias de la conducta no proveen de herramientas de evaluación que sean confiables y válidas, aceptables para determinar el desempeño de los empleados, las organizaciones simplemente no evaluarán el desempeño formalmente o lo harán inadecuadamente. Esto resulta inexcusable en un contexto de toma de decisiones que determinan la vida laboral de las personas y el desempeño organizacional (Páez, 2001). Así la importancia de los miembros de la organización como seres humanos y recursos organizacionales quedará sólo en el nivel discursivo.

Por todo lo anterior, la presente investigación se orientará a responder la siguiente pregunta:

¿Cuál es la confiabilidad, validez y aceptabilidad de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos² de la Fábrica de la Compañía Chilena de Tabaco de la V Región?

Estado Actual del Problema

La necesidad de evaluar el desempeño en Compañía Chilena de Tabaco de la V Región se enmarca, en primer lugar, en un proceso de mejoramiento de la gestión del recurso humano como una forma para enfrentar los nuevos desafíos que presenta el entorno social, político y económico de la empresa. Dentro de los desafíos más actuales, cabe mencionar especialmente: el “Convenio Marco para el Control del Tabaco” que fue suscrito el 21 de enero del año pasado por la Organización Mundial de la Salud; la adopción de las Normas ISO 9000; la introducción en el mercado nacional de Philips Morris, principal competidor en el ámbito internacional, y los crecientes índices de contrabando en el mercado nacional.

En correspondencia con las tendencias internacionales de la gestión de recursos humanos, la empresa identificó las variables críticas del desempeño organizacional, a saber: los Procesos, los Sistemas, la Gerencia y las Personas. En cuanto a la gestión del recurso humano, si bien la empresa no cuenta con una planificación estratégica propiamente tal, busca potenciar la capacidad de liderazgo en todos los niveles de organización de modo de dirigir la creatividad y entusiasmo de los empleados hacia un desempeño superior, optimizar el aprendizaje en un nivel personal y organizacional, desarrollar el talento así como reclutar y retener a los empleados más competentes, y fomentar una cultura abierta que se orienta a satisfacer las necesidades de clientes y consumidores (ver Apéndice A).

² Los técnicos, según la Clasificación Internacional de Ocupaciones (CIU-88), son técnicos y profesionales de nivel medio de las ciencias físicas y químicas, la ingeniería y afines. Según su área de especialización podemos encontrar en la empresa técnicos en electrónica y telecomunicaciones (Código 3114) y técnicos en mecánica y construcción mecánica (3115). Por su parte, los asistentes técnicos de la Compañía Chilena de Tabaco pertenecen al grupo de operadores de máquina y montadores. En la empresa, podemos encontrar operadores de máquinas para elaborar productos del Tabaco (código 8279) y operadores de imprenta (código 8251). Cabe destacar que la categoría de técnico y de asistente técnico no establece directamente el nivel educacional de los empleados que ocupan dichos cargos. Para más información sobre estos puestos de trabajo, ver Apéndice B.

Una de las primeras iniciativas en el Área de Operaciones, orientada por estos principios, fue el proceso de reingeniería del proceso productivo que se efectuó el año 2000. Este proyecto contempló un cambio en la distribución de la maquinaria, la estandarización de los procesos productivos, nuevas descripciones de cargo, cursos de comunicación, liderazgo y coaching, entre otras acciones. Una de las consecuencias más relevantes fue el cambio de la estructura organizacional en el Área de Operaciones. Se instauró una estructura horizontal más propia de equipos autoguidados, eliminándose así los mandos medios. Actualmente, hay dos cargos al interior de cada célula³: técnico electrónico/mecánico y asistente técnico. Con anterioridad al proceso de reingeniería, existían más de 10.

La empresa así, con anterioridad al proceso de reingeniería, empleaba el método de Escalas Gráficas para evaluar el desempeño del personal de línea. Según lo establecido en el Informe de Evaluación de Desempeño (Chiletabacos, 1994), el instrumento referido presentaba tres grandes debilidades: (a) una débil relación entre los aspectos críticos de los puestos de trabajo y variables de evaluación, (b) la tendencia de los evaluadores (jefe - directivo) a otorgar calificaciones más altas a los empleados con mayor jerarquía y (c) la retroalimentación resultaba ser poco significativa. Como estas evaluaciones de desempeño estaban asociadas a las antiguas descripciones de cargo, no se siguieron aplicando.

En la actualidad, los técnicos y asistentes técnicos son los únicos cargos en que su personal no es evaluado formalmente. Entre el año 2000 a la fecha, se ha estudiado implementar dos sistemas de evaluación de desempeño para técnicos y asistentes técnicos, un sistema de Escala Gráfica que es usado por Motorola Italia y un sistema basado en el Modelo de Competencias que fue especialmente diseñado para la empresa, pero ambas iniciativas no fructificaron. En relación con estos sistemas de evaluación, cabe destacar el sistema de evaluación de competencias elaborado por la Universidad Diego Portales, una Escala Anclada en el Comportamiento (BARS), presentó deficiencias de validez por cuanto evaluaba sólo competencias técnicas y obviaba las competencias generales consideradas relevantes por la empresa así

³ Unidad de trabajo autónoma

como también resultó tener baja consistencia en los puntajes otorgados por diferentes evaluadores. Cabe destacar además que los evaluados no participaron en el proceso de diseño de este sistema.

La falta de un sistema de evaluación de desempeño es percibida negativamente por técnicos y asistentes técnicos, así como por los jefes de producción (jefes directos de estos últimos). En los estudios orientados a evaluar la gestión del recurso humano como la evaluación de la efectividad de la capacitación, estudios de satisfacción laboral, estudios de clima arrojan resultados similares respecto de este punto. En general, podemos afirmar que no existe relación clara y directa entre la toma de decisiones administrativas y de recompensas organizacionales y el desempeño individual. Al respecto, se plantea que:

1. Producto del proceso de reingeniería, una creciente automatización del proceso productivo y de una mayor polifuncionalidad de técnicos y asistentes técnicos, se ha reducido, continua y sistemáticamente, la planta de técnicos y asistentes técnicos. La toma de decisiones de desvinculación se efectúa sobre la base de criterios externos al desempeño del empleado como antigüedad/edad, salario, situación habitacional. A su vez, los jefes de producción, basándose en observaciones informales y en opiniones de terceros, también consideran el compromiso con la empresa, la actitud hacia el trabajo, entre otros. Al no existir indicadores claros para predecir quiénes serán despedidos, en técnicos y asistentes técnicos es posible observar una fuerte percepción de inestabilidad laboral. Por su parte, como los jefes de producción no disponen de datos válidos y confiables para tomar las decisiones de desvinculación experimentan una mayor tensión asociada al proceso.

2. Del mismo modo, las decisiones de promoción no son validadas por el personal por la ausencia de criterios claros que permitan sustentarla. Específicamente, según el Estudio de Efectividad Organizacional (International Survey Research Offices Worldwide, 2001), sólo el 31% de los técnicos y asistentes técnicos considera que hay acceso igualitario para las oportunidades basadas en el desempeño y el 43% cree que la empresa hace un buen trabajo en promover a las personas más competentes.

3. Los técnicos y asistentes técnicos no perciben además una relación clara entre el desempeño presentado y las recompensas organizacionales. Si bien existe un sistema de recompensa para el equipo en función del logro de los objetivos de la célula, las contribuciones individuales no son observadas ni recompensadas. Así, sólo el 38% de los técnicos y asistentes técnicos cree que la empresa usa bien el reconocimiento no financiero y las recompensas para estimular el buen desempeño (International Survey Research Offices Worldwide, 2001), lo cual se suma a que sólo el 20 % de los mismos considera que la empresa relaciona bien a muy bien el salario de la persona con su desempeño.

4. Las acciones tendientes a mejorar el desempeño del personal se dan sólo en el nivel grupal, por lo tanto, no es posible priorizar los recursos que siempre son escasos. Hoy la evaluación del desempeño individual cobra mayor importancia dada la reducción del presupuesto de capacitación. De esta manera, una evaluación que detecte las fortalezas y debilidades en el desempeño del individuo sería útil para determinar prioridades de capacitación o establecer otras acciones de mejora.

5. Finalmente, es relevante para los jefes de producción y los empleados dar/recibir feedback sobre el desempeño. En cuanto a estos últimos, sólo el 44% señaló que recibe feedback valioso de su supervisor/directivo de línea (International Survey Research Offices Worldwide, 2001). En cuanto a los primeros, entregar feedback constituye un indicador formal de la evaluación de desempeño en la dimensión que refiere a la gestión del talento de los subordinados.

Objetivos de Investigación

Objetivo General

Explorar las propiedades psicométricas y la aceptabilidad de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Fábrica de la Compañía Chilena de Tabaco

Objetivos Específicos

Detectar las necesidades de evaluación en el nivel de gerencia, evaluadores y evaluados a considerar en el diseño de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Fábrica de la Compañía Chilena de Tabaco

Determinar la validez de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Fábrica de la Compañía Chilena de Tabaco

Determinar la confiabilidad de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Fábrica de la Compañía Chilena de Tabaco

Detectar la aceptabilidad de una propuesta de evaluación del desempeño de técnicos y asistentes técnicos de la Fábrica de la Compañía Chilena de Tabaco

MARCO TEÓRICO

Gestión de Recursos Humanos y Evaluación del Desempeño

En la actualidad, se reconoce indudablemente que los recursos humanos son el principal activo y fuente para lograr o mantener la ventaja competitiva de una organización. Al respecto, Deloitte y Touche (2002) plantea que el activo relacionado con las prácticas de recursos humanos explica el 43% de la variación en el valor de mercado de las empresas⁴.

Hoy más que nunca, la gestión de recursos humanos se ve enfrentada a un conjunto de desafíos, provengan ya del entorno, de las organizaciones mismas o de los miembros que la componen. Ante esta multiplicidad de demandas, las organizaciones pueden adoptar una gestión estratégica de recursos humanos o dedicarse a diseñar e implementar, en forma reactiva, una serie de prácticas y políticas que no presentan mayor coherencia entre sí.

Una gestión sistémica implica una evaluación de desempeño integrada funcionalmente a los demás sistemas de recursos humanos. Al respecto, Quijano (1994) afirma: “La evaluación de rendimiento en el trabajo se sitúa, como parte del proceso de gestión de recursos humanos, en el nivel de los individuos. Entre las acciones directivas que integren dichas gestión, cabe situar las siguientes: reclutamiento, selección, formación, evaluación, administración de recompensas y sanciones, colocación, diseño de puestos de trabajo, etc. Todas ellas tienen una vertiente personal, referidas a individuos particulares, y una vertiente institucional, referida a sistemas generales de proceder en la Organización” (pp. 264 y 265).

Esta integración, sin embargo, es un desafío más que una realidad en la gran mayoría de las organizaciones. Al respecto, Barnett afirma que: “La supuesta integración sistémica de la función de evaluación del desempeño es descrita, la mayoría de las veces, de forma más bien mecánica, ya sea como

⁴ Indicador de la capacidad de las empresas para generar valor más allá de los activos financieros y físicos.

un elemento de una sucesión temporal de tareas a realizar, o bien, como una fuente genérica de información para los demás subsistemas de recursos humanos.” (1997, pp. 6).

Una integración de la evaluación de desempeño y gestión de recursos humanos, en primer lugar, requiere de la formulación de una estrategia de recursos humanos que se alinee con la estrategia general del negocio (Alles, 2000). Recientemente, el estudio de Deloitte y Touche (2002) encontró que una selección de variables de evaluación que considera la estrategia y orientación del negocio se correlaciona positivamente con el valor de mercado de las empresas. Específicamente, concluyó que las empresas orientadas a la excelencia operacional (industria manufactura y transporte - comunicaciones)⁵ deben evaluar el desarrollo de nuevas habilidades; así como las empresas de servicio (Comercio, Banca, Hotelería y Restaurantes), el conocimiento de los estándares de productividad.

La gestión estratégica de recursos humanos, no obstante, sigue siendo hoy un desafío más que una realidad. Hax y Majluf plantean que: “Sin embargo, a pesar del creciente interés, el manejo estratégico de los recursos humanos está lejos de ser una realidad en la mayor parte de las empresas norteamericanas” (1996, pp. 329). En Chile, esta realidad no es muy diferente porque una gestión estratégica del recurso humano implica una gestión formal de los recursos humanos que, según Ramírez (2000), posea una orientación proactiva y una concepción del personal como recurso clave a optimizar; diagnostique las amenazas y oportunidades del ambiente, fortalezas y debilidades de la organización; formule objetivos congruentes con este diagnóstico e implemente prácticas y medidas para lograr dichos objetivos.

Las categorías de prácticas que deben considerarse en la formulación de una estrategia de recursos humanos son: (a) el manejo de la selección, promoción y asignación de puestos de trabajo, (b) la fijación y distribución de recompensas y beneficios complementarios así como el apoyo motivacional, (c) el desarrollo ejecutivo, (d) el manejo de las relaciones y canales de comunicación y (e) la evaluación del desempeño. Cada categoría encierra una serie de alternativas estratégicas. Es vital que exista congruencia

⁵ La estrategia de las empresas que buscan excelencia operacional, según los autores, enfatiza la eficiencia y mejoramiento de los procesos productivos, centraliza la toma de decisiones para potenciar la estandarización y el manejo de costos así como enfatizan el trabajo en equipo, el mejoramiento continuo y la calidad.

en la selección las alternativas de cada categoría y entre las categorías mismas. Con relación a lo expuesto,

Hax y Majluf (1996) plantean que:

“Es fácil ver cómo están vinculadas las cinco categorías de decisión estratégica de recursos humanos. Selección, promoción y destinación identifican al mejor talento disponible para llevar a cabo los trabajos críticos de la empresa. Por sí solas, éstas no serían suficientes en un mundo en el que las habilidades y conocimientos tienen que ser puestos al día para que los recursos humanos continúen siendo competentes. De modo que el desarrollo de los ejecutivos es una actividad estratégica clave. Más aún, es necesario que una fuerza de trabajo adecuadamente desarrollada y hábil esté comprometida de todo corazón con la organización. Esto hace resaltar la importancia de las actividades motivacionales y recompensas, así como un cuidadoso manejo de las relaciones y canales de comunicación con obreros y empleados. Y por último, ninguna de estas tareas se podría efectuar de buena forma en ausencia de un efectivo sistema de evaluación, que permita hacer el mejor uso de los recursos humanos disponibles ante las tareas necesarias en la organización, la planificación y ejecución de los esfuerzos de capacitación de ejecutivos, el diseño y funcionamiento de un sistema inteligente de recompensas y la identificación de las actividades necesarias para establecer el tipo de medio ambiente participativo que se necesita para el éxito de la empresa.” (pp. 335 y 334).

En relación con los sistemas de evaluación de desempeño, cabe mencionar que las decisiones estratégicas son:

1. Sistema orientado al proceso v/s a resultado: El primer tipo de sistema se orienta a medir el logro de indicadores de rendimiento; obviando los factores del individuo y del entorno que influyen en estos. Por el contrario, el segundo considera dichos factores, ya que se aboca a evaluar el proceso por el cual se obtienen determinados resultados.

2. Identificación de necesidades de capacitación v/s de contratación de personal nuevo: Un sistema destinado a identificar las necesidades de capacitación se preocupa por el desarrollo de los empleados en cuando a habilidades y conocimientos. Por su parte, un sistema orientado a determinar las necesidades de contratación establece si se deberán reclutar personal para el puesto en que los empleados son evaluados o para otras vacantes.

3. Evaluación de rendimiento individual-grupal v/s divisional-corporativo. La evaluación del primer tipo se da en el nivel de las personas mientras que la segunda refiere al logro de las metas de la organización.

4. Comparación de series cronológicas v/s corte transversal. Cuando se comparan series cronológicas, se evalúa el desempeño histórico que se presenta en un marco (evaluación del desempeño presentado en un periodo de tiempo).

La evaluación de desempeño además se ubica dentro de la administración del desempeño, la cual puede ser definida como un conjunto de prácticas que aseguran que las metas y las actividades individuales se relacionen y potencien el logro de los objetivos de la organización. En un sentido más amplio, implica especificar, revisar y mejorar de manera continua, en la actividad cotidiana, los desempeños organizacionales, grupales e individuales conducentes al logro de la misión empresarial (Barnett, 1997). Según el autor, aparte de la evaluación, el análisis y descripción de cargos con la posterior fijación de objetivos o expectativas, la retroalimentación, el análisis del desempeño y la planeación de las mejoras del mismo componen la administración o gestión del desempeño.

Como ya se ha expuesto, la evaluación de desempeño se vincula más directamente con las prácticas de reclutamiento y selección, compensación, capacitación, desarrollo. Algunas de las relaciones son:

1. La evaluación de desempeño también permite planificar la sucesión y colocación de personas a corto plazo sobre la base de las motivaciones y destrezas de los empleados.

2. La evaluación de desempeño permite evaluar la efectividad de la capacitación, diseñar programas en función de necesidades específicas así como determinar que empleados se verán más beneficiados con un programa de capacitación determinado.

3. La evaluación de desempeño permite fijar las recompensas y castigos organizacionales.

4. Los resultados de las evaluaciones de desempeño permiten evaluar el proceso de reclutamiento/selección en cuanto a la efectividad de los predictores que éste último emplea. También se puede mejorar el desempeño del personal al contratar personal más calificado en un área de desempeño que un análisis de los resultados de la evaluación muestra como deficitario o focalizar el reclutamiento en

un mercado de trabajo determinado. Por último, los criterios y variables de evaluación pueden ser usados para orientar el proceso de selección y reclutamiento.

5. La evaluación de desempeño, cuando es más que completar un formulario de evaluación, promueve el análisis de variables que intervienen en el desempeño, las cuales deben identificarse y, en lo posible, ser manipuladas de modo de optimizar el mismo.

6. La evaluación de desempeño entrega importante información al sistema de control. Es empleada así para enfrentar eventuales juicios laborales⁶, tomar decisiones, y evaluar la eficacia de la gestión de recursos humanos.

Desempeño Laboral

Cómo definimos desempeño constituye un elemento central de la práctica de evaluación de desempeño así como en la administración del mismo. Las definiciones se orientan invariablemente al nivel individual del fenómeno. De hecho, según Ramos et. al. (1996): “... se ha dedicado más atención a estudiar a los indicadores y la medición, con un notorio sesgo psicométrico, que a definir claramente el constructo” (pp. 291).

El desempeño, en términos generales, es conducta (Campbell en Ramos et. al, 1996). Existen dos posturas frente al tipo de conductas que conforman el “Desempeño laboral”. La postura tradicional, representada por Chiavenato (2000), conceptualiza al desempeño como el comportamiento de rol del ocupante de un cargo. Se resalta así a la contribución que hace el individuo a las metas y objetivos de la organización al realizar las tareas propias de su cargo. La segunda postura, por su parte, considera que son relevantes las conductas que contribuyen al buen funcionamiento de la organización aún cuando no sean

⁶ La evaluación de desempeño tiene gran relevancia en Estados Unidos por cuanto agencias estatales y ONG (ejemplo: Equal Employment Opportunity Commission, Civil Rights Division del Departamento de Justicia y Office of Federal Contract Compliance Programs) velan por la igualdad de oportunidades laborales. De hecho, la investigación sobre evaluación desempeño se ha centrado en los sesgos de las evaluaciones por esta razón.

propias de un cargo o estén establecidas formalmente. De este modo, según esta perspectiva, el desempeño incluye conductas de rol y conductas extra-rol. (Ramos et. al, 1996).

Las conductas propias del rol son: “aquellas conductas requeridas y esperadas como parte de la ejecución de las obligaciones y responsabilidades del rol asignado. (Van Dyne, Cummings y Mc Lean Praks en Ramos et. al., 1996, pp. 285)

Las conductas extra rol, concepto acuñado por Katz y Kahn según Ramos et. al. (1996), son conductas que: (a) no son parte de la prescripción formal del puesto, (b) son voluntarias, (c) facilitan el buen funcionamiento de la organización y (d) benefician a los demás. Ejemplos de conductas extra-rol que podrían ser incluidas en el concepto de desempeño son las conductas prosociales, conductas de denuncia, conductas de disenso de la organización por principios, etc.

El límite conductas de rol y extra-rol es un poco difuso tanto en el nivel conceptual como en el práctico, ya que directivos y empleados no siempre concuerdan en qué conductas son propias del cargo y cuáles son externas a éste.

El desempeño, por tanto, se define como lo que las personas hacen para contribuir al cumplimiento de las metas u objetivos así como al buen funcionamiento de la organización. Cabe resaltar que el “Desempeño”, según Campbell en Ramos et. al. (1996) denota la acción y no a los resultados de la misma por lo que no se debe confundir desempeño con eficacia y eficiencia.

La eficacia, en primer lugar, se refiere al logro de objetivos. Es un resultado parcial de la contribución de los individuos, ya que existen factores externos que están en juego. La eficiencia, en segundo lugar, se refiere a la ejecución de las tareas propias del cargo en función de los criterios de costo, calidad, cantidad y oportunidad, interviniendo en ésta factores externos también (Toro, 1996, Ramos et. al., 1996).

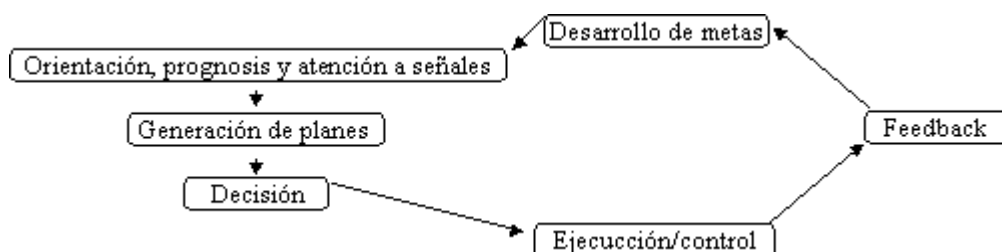
Enfoques del Desempeño Laboral

Explicar el porqué de un desempeño es uno de los focos de interés de la Psicología del Trabajo. No obstante, las teorías no alcanzan a explicar el fenómeno en toda su extensión y no son sólidas (Campbell en Ramos et. al., 1996). A continuación, se presentan dos teorías y un modelo de desempeño que si bien no son exhaustivos, permiten una primera aproximación teórica al fenómeno.

Teoría de Acción de Hacker

Esta primera teoría, formulada en la década de los 70', postula que la conducta laboral está regulada consciente o automáticamente por cogniciones que se dan a lo largo de un proceso que incluye las fases que se ilustran en la siguiente figura. La teoría de la acción se aplica desde las secuencias automatizadas de movimientos estereotípicos, acciones concernientes a la solución de problemas y análisis complejos hasta la generación de metaplanes.

Figura N° 1: Proceso de la Acción de Hacker



Fuente: Ramos
et al. 1996 p.

La figura N° 1 muestra que la acción, como conducta orientada por metas, se da a lo largo de un proceso cíclico.

La teoría postula que, en una primera fase del proceso, las metas parten siendo un mero deseo, a continuación, el deseo se transforma en un querer algo (pretensión) y, cuando es importante y oportuno, en una intención determinada, la cual actúa como verdadera meta. Las intenciones relevantes se relacionan con los resultados de la tarea que la organización delimita en términos muy genéricos. Luego, el propio individuo es quién hace suya la tarea, dividiendo ésta en tareas específicas. Esto último depende de diversos factores como la comprensión de la tarea, grado de aceptación, voluntad de llevarla a cabo, etc.

Es importante acotar que las metas: (a) son relativamente estables, (b) cuando llegan a ser patrones de acción para el individuo, dejan de ser conscientes y (c) son múltiples, es decir, la conducta de un individuo puede estar orientada por varias metas, las cuales inclusive pueden estar en conflicto entre sí.

La persona en la segunda fase, se orienta hacia la meta establecida. Esto supone la búsqueda y recogida de información, el desarrollo de analogías y el uso de esquemas. Es decir, se atiende a lo que es importante para la acción, integrando esto a los conocimientos sobre la tarea laboral.

En la tercera fase, el trabajador genera planes, o sea, determina de un modo consciente o automático el cómo va a proceder. El nivel de detalle y la orientación temporal de los planes así como la jerarquía de los subplanes pueden variar. Miller en Ramos et. al. (1996) señala que los planes representan el puente entre el pensamiento y la acción.

En cuanto a la ejecución de los mismos, cuarta fase, la coordinación entre los planes, la rapidez en su ejecución y la flexibilidad para adaptarlos sin abandonar la meta son relevantes.

La última fase del proceso de acción, feedback, permite al sujeto observar el progreso hacia la meta. Frese y Zapf, según Ramos et. al. (1996), distinguen diversos tipos de feedback: concurrente v/s terminal, extrínseco v/s intrínseco, inmediato v/s diferido. Sin importar el tipo, los autores señalan que éste se caracteriza por:

1. El grado en que responde a la realidad. El feedback debe ser lo más realista posible de modo de evitar interpretaciones sesgadas por parte del empleado y potenciar un desempeño efectivo.
2. Los componentes ante los cuales reacciona el sujeto. Responder al contenido social del feedback implica estrategias defensivas que no son generadas por el contenido de desempeño.
3. La tasa de búsqueda de feedback. Es importante, sobre todo en situaciones ambiguas y dinámicas, promover la búsqueda activa de feedback.

Cabe destacar finalmente que ésta ha potenciado la investigación en torno a los errores, la relación de la personalidad y trabajo, las competencias y las características de la tarea. En cuanto a la aplicación en las organizaciones, ha permitido además mejorar el diseño del trabajo. (Ramos et. al., 1996)

Teoría de la Conducta en las Organizaciones de Naylor, Pritchard e Ilgen

Esta es una de las teorías más formalizadas y detalladas que se centra en la interacción de un organismo con su entorno, siguiendo el esquema estímulo, organismo, respuesta y feedback. Para ésta teoría, desarrollada en 1980, la conducta es:

1. Un proceso en curso en el cual el acto es la unidad de la misma. No obstante, la teoría se centra en los productos relevantes de la misma más que en el proceso. Cabe mencionar que son productos relevantes aquellos que son de interés para la organización.

2. La conducta es un problema de asignación de recursos del sujeto, a saber: tiempo y esfuerzo. A su vez, implica la elección de una conducta determinada respecto de un conjunto de conductas alternativas. El tiempo y esfuerzo destinado representa el compromiso del individuo con la conducta seleccionada.

La presente teoría constituiría una extensión de la teoría de las expectativas, presentando una mayor complejidad respecto de ella, ya que contiene una serie de factores organizacionales e individuales así como intenta explicar una serie de procesos interdependientes. Dada esta complejidad, se expondrán por separado sus principales elementos, a saber: entorno, diferencias individuales, percepciones, resultados y afecto, valencias, utilidades y actos. Sin perjuicio de lo anterior, la figura N° 3 presenta las relaciones entre estos elementos.

Entorno

El entorno general de la organización y el específico del trabajador afectan a la conducta por cuanto los sistemas de medidas de la organización determinan qué productos (lo que hace el empleado) son relevantes. El rendimiento así varía en función de los actos relevantes que son considerados por los sistemas de evaluación y no sólo en función de lo que realmente hace el empleado. Es importante notar que los resultados (recompensas) para el empleado se derivan de los resultados de la evaluación más que de su rendimiento real.

Características de personalidad. Conjunto de creencias sobre uno mismo y sobre el entorno así como patrones de conducta que se generan para procesar y manejar los estímulos del entorno. Las características de personalidad son estables e influyen en la conducta. Cabe acotar que la teoría no especifica las características que influyen en la conducta.

En suma, estas tres diferencias influyen en la conducta por cuanto afectan: (a) la percepción de sí mismo, la conducta, productos y resultados extrínsecos (Proceso A), (b) la transformación de utilidad de los actos en actos reales (Proceso I), (c) los afectos⁷ y valencia de los resultados (Proceso E y F). A su vez, éstas son afectadas por el tiempo, la conducta y los resultados así como el mismo afecto.

Percepciones y evaluaciones percibidas

La teoría reconoce cuatro tipos de percepciones que tienen efectos diferentes en el procesamiento de la información, a saber: percepciones de características y atributos de sí mismo, la conducta, los productos y los resultados extrínsecos. La evaluación, la autoevaluación así como la evaluación percibida que los demás realizan afectan las relaciones que el individuo establece entre el acto-producto, producto-evaluación y evaluación-resultados. Todos estos elementos más las expectativas de rol del sujeto y de los otros miembros influyen en la evaluación que la persona hace de su propio desempeño

Resultados y afectos

La evaluación del desempeño conlleva a la formación de la percepción de los resultados absolutos y relativos. Los primeros se refieren al valor que el sujeto otorga a los resultados extrínsecos que obtuvo mientras que los segundos establecen una comparación entre los resultados obtenidos por el individuo así como los resultados obtenidos por los demás. De esta manera, la percepción de los resultados relativos se ve influida por las evaluaciones del desempeño que hace la organización, los demás así como el propio individuo. El afecto, por su parte, influye más que en la conducta misma, en el estado temporal de las necesidades y en la valencia que se le otorga a los resultados. En cuanto a los resultados

⁷ El término afecto es equivalente al concepto de satisfacción laboral.

extrínsecos, el afecto está dado por la comparación entre los resultados obtenidos y la satisfacción de las necesidades. Por el contrario, en el caso de los intrínsecos (resultados no controlados por el entorno), el afecto está influido por la percepción de los actos inherentemente placenteros o divertidos o por el nivel de desempeño establecido por la evaluación del mismo.

Valencias, utilidades y actos

Los resultados percibidos, el afecto asociado a estos así como el estado temporal de las necesidades determinan la valencia que el sujeto entrega a los resultados. Esta valencia junto con la contingencia evaluación-resultado y producto-evaluación determinan la utilidad de los productos de la conducta. A su vez, esta utilidad con la contingencia acto-producto determinan el valor de los actos en cuanto si llevan a alcanzar o no los productos que el individuo considera que son útiles, o sea, determinan la utilidad de los actos. Dada la utilidad de los actos, se generan o no intenciones de conducta.

La teoría señala que las diferencias individuales descritas y el entorno pueden facilitar o impedir que la intención de conducta se transforme en conducta real. Por último, se establece que sólo una parte de los actos o conductas presentadas serán observadas y, posteriormente, consideradas, evaluadas y premiadas.

Modelo de Desempeño de Campbell

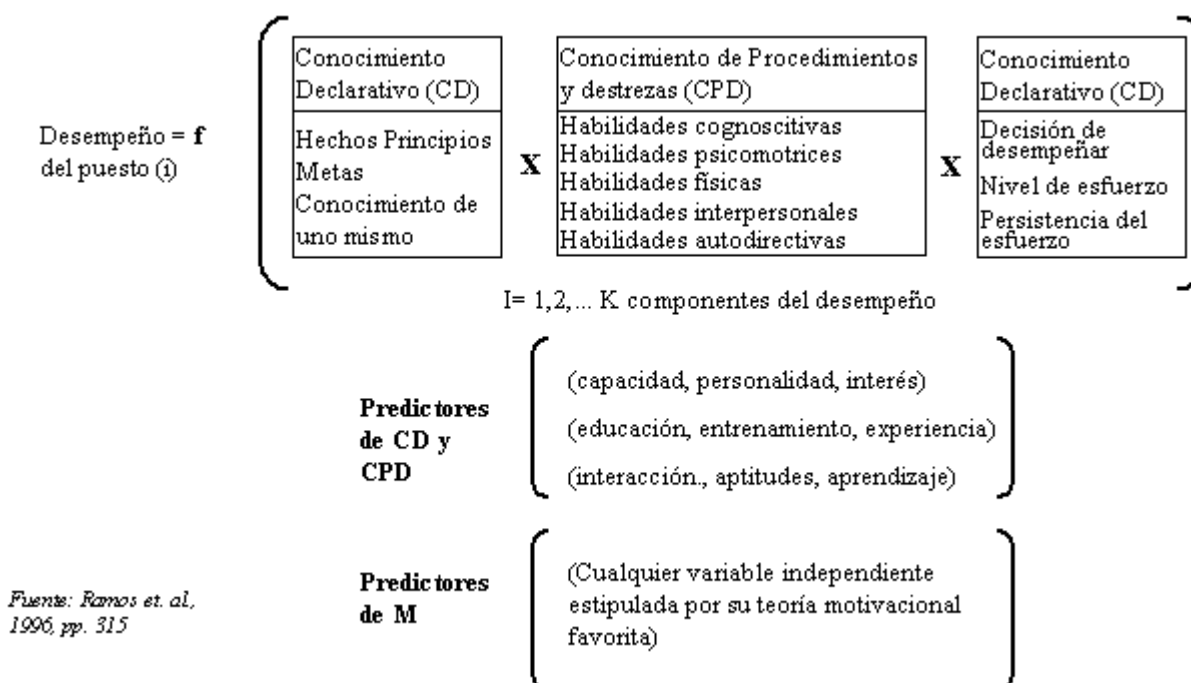
Este modelo tentativo, desarrollado en la década del noventa, se centra en los determinantes del desempeño. Para el modelo, el desempeño es acción y no los resultados de la misma. Está determinado por el conocimiento declarativo, el conocimiento de los procedimientos – destrezas y la motivación (Ramos et. al., 1996). En primera instancia, el conocimiento declarativo refiere al conocimiento del qué debe hacerse (hechos, cosas, principios, metas, conocimiento de uno mismo), o sea, alude a la comprensión de los requerimientos de la tarea. Por su parte, el conocimiento de los procedimientos refiere al conocimiento del cómo debe hacerse así como las destrezas a las habilidades físicas, cognitivas,

psicomotrices, perceptivas, interpersonales y de autodirección. La motivación, a su vez, implica la dirección, amplitud y persistencia del esfuerzo.

El autor diferencia determinantes de predictores y componentes del desempeño. Los segundos son propios de cada determinante, los cuales se señalan en la próxima figura. Al respecto, Campbell establece que para mejorar el desempeño se debe actuar sobre los determinantes del desempeño (Ramos et. al. 1996)

El autor del modelo, según Ramos et. al. (1996), entrega también una taxonomía general de componentes, los cuales son: (a) pericia en tareas específicas del puesto, o sea, en las tareas que distinguen un puesto de otro; (b) pericia en tareas no específicas, las cuales son propias de varios puestos, (c) tareas de comunicación oral y escrita, (d) mantenimiento de la disciplina personal, (e) facilitar el desempeño del equipo y compañeros, (f) supervisión y (g) dirección. Es importante mencionar que los puestos enfatizan algunos componentes sobre otros.

Figura N° 3: Modelo de Desempeño de Campbell



Determinantes del Desempeño

Los determinantes del desempeño han sido considerados como factores del ambiente que afectan positiva o negativamente el mismo y/o como un conjunto de variables del individuo. Si bien las teorías presentadas y especialmente el Modelo de Campbell enfatizan algunos determinantes, a continuación, estos serán revisados desde una óptica más integral.

La postura más tradicional plantea que el desempeño está determinado sólo por variables del individuo, de tipo psicológicas, siendo la motivación y habilidad las más destacadas. Así, para Lawler en Toro (1996), el desempeño es: "... una función del conocimiento, habilidad y motivación. Cada uno de estos determinantes interactúa con los demás y entre todos regulan y determinan las actuaciones laborales" (pp. 98). Habilidad, motivación y conocimientos son agentes causales. Las condiciones sociales, culturales, demográficas, las condiciones del puesto de trabajo y de la organización así como las características de personalidad influyen en la habilidad, motivación y conocimientos, pero no determinan directamente el desempeño (Toro, 1996). Cummings y Schwab (1985), van más allá, proponen que el desempeño es un: "Fenómeno individual que es determinado por las habilidades y motivación. Los factores ambientales, incluyendo las actividades llevadas a cabo por los gerentes para mejorar el desempeño del empleado, tienen un impacto, si acaso, en las habilidades y motivación." (pp. 14). Dentro de los factores ambientales podemos encontrar: diseño de trabajo, supervisión, compañeros de trabajo, compensaciones, condiciones de trabajo, capacitación y evaluación. La postura de Cummings y Schwab puede graficarse mediante la siguiente fórmula: desempeño es igual a motivación por habilidad. Se intenta representar aquí que cuanto más motivado esté un trabajador, más efectiva será su capacidad, por lo tanto, mejor será su desempeño (Rodríguez, 1998).

Chivenato (2000), por su parte, expone que el desempeño como fenómeno que varía de persona a persona depende de múltiples variables como el valor de las recompensas y la percepción de este valor, el esfuerzo individual, las capacidades del individuo y la percepción acerca del papel desempeñado. Recientemente, este autor ha acuñado el concepto oportunidad de desempeño. El desempeño así está

determinado por la existencia o ausencia de recursos de apoyo. De este modo, el desempeño laboral estaría definido por factores internos y externos al individuo.

Antes de abordar los factores externos, es necesario destacar que la aproximación desde el modelo de competencias ha establecido, de manera muy similar al modelo de Campbell, que las competencias específicas se relacionan estrechamente con el desempeño laboral. Warr y Conner, en Ramos et. al. (1996), definen competencia como “...el conjunto de conductas, conocimientos, procesos de pensamiento y/o actitudes que se ven reflejadas en el desempeño.” (pp. 334), a su vez, establece que las competencias específicas son “... tipos específicos de conductas competentes en la situación de trabajo...” (pp. 333). Según el autor, las competencias específicas constituirían el desempeño laboral en la medida en que las conductas son eficientes.

La línea entre el desempeño y los determinantes de éste se vuelve difusa. Así, el autor también señala que: “Se asume que uno de los principales determinantes de la conducta laboral reside en las competencias concretas para desarrollar conductas eficientes en el marco de un trabajo concreto. En este sentido, la mejora del desempeño va relacionada con el progreso y el aumento de competencias específicas, y permite identificar las condiciones ambientales que dificultan la adquisición de competencias” (1996, pp.334)

La investigación se ha centrado mayormente en las competencias específicas, las cuales siguen teniendo un carácter situacional, se aplican a un reducido número de puestos de trabajo. Por ello, Tordera en Ramos et. al., (1996), explica que la mayoría de las investigaciones ha estudiado las competencias específicas de directivos.

En relación con el papel que juegan los determinantes del individuo en la evaluación del desempeño, cabe destacar finalmente que éstos no constituyen el objeto de evaluación. Es decir, la evaluación no debe juzgar el perfil de personalidad, motivación o competencias que lleva al empleado a conseguir un rendimiento determinado, sino sólo establecer el rendimiento que éste presenta (Pontifes, 2000).

A pesar de lo expuesto, no todos los determinantes del individuo son de índole psicológica. La investigación ha probado que factores como la alimentación, la presencia de enfermedades, el consumo de sustancias, etc. afectan el desempeño por cuanto influyen en las capacidades del individuo (ej, atención, concentración, activación, motricidad). La edad es otro antecedente relevante. Cabe destacar que no existe una relación lineal entre la edad y el desempeño, ya que hay múltiples factores contextuales (feedback, creencias generacionales, motivación, etc.) que modulan dicha relación (Ramos et. al., 1996; Waldman y Avolio, 1986)

La línea de investigación centrada en la mejora del desempeño a partir de factores externos, por su parte, habla de facilitadores e inhibidores, todos factores propios del ambiente que permiten o impiden que un individuo logre el nivel máximo de desempeño que puede alcanzar. Facilitadores o inhibidores no son siempre los mismos. Así, por ejemplo, en cuanto a los recursos materiales, si estos faltan, el empleado no presentará el máximo desempeño, pero no hay garantía que, dados los recursos materiales, el trabajador presente el nivel máximo de desempeño que puede aspirar.

En cuanto a los inhibidores, Peter y O'Connor, en Ramos et. al. (1996), entregan una clasificación de 8 factores que representa un buen punto de partida. Estos son: (a) información relevante para realizar el trabajo, (b) herramientas y equipos, (c) materiales e insumos, (d) apoyo financiero, (e) servicio financiero, (f) servicios requeridos y ayuda de otros, (g) preparación de la tarea en relación con la educación, entrenamiento, experiencia previa, (h) disponibilidad de tiempo y (i) entorno laboral.

En cuanto a los facilitadores, el entrenamiento, el establecimiento de metas y los cambios sociotécnicos presentan una mayor capacidad para mejorar el desempeño. Los incentivos económicos, el rediseño del trabajo, el feedback así como el liderazgo influyen también positivamente en el desempeño. La dirección por objetivos y las prácticas de supervisión no afectarían el desempeño (Guzzo, Jette y Katzell en Ramos et. al., 1996).

Es importante destacar que los facilitadores e inhibidores varían según los puestos de trabajo y las condiciones propias de la organización.

Evaluación del Desempeño

La evaluación de desempeño ha sido siempre definida desde una perspectiva amplia y general. Probablemente, la diversidad de objetivos, objetos, métodos, entre otros elementos que emplea en virtud de las necesidades particulares de cada organización ha dificultado caracterizar este proceso. Sin embargo, una regularidad implícita es el carácter racional que se le atribuye. La evaluación sería entonces un proceso cognoscitivo que no estaría influenciado por afectos y emociones⁸.

A continuación, se presentan algunas definiciones de evaluación desempeño:

1. Para Chiavenato (2000), la evaluación es una técnica de dirección imprescindible para la administración. Como instrumento, busca mejorar los resultados del recurso humano mediante la apreciación sistemática del rendimiento de cada persona en el cargo o su potencial de desarrollo. Este autor postula además que: “es un medio para obtener datos e información que puede registrarse, procesarse y canalizarse para mejorar el desempeño humano. En el fondo, la evaluación de desempeño sólo es un buen sistema de comunicaciones que actúa en sentido horizontal y vertical en la empresa” (pp. 366)

2. Por su parte, Álamos (1994) postula que es un sistema usado en la administración del recurso humano en cuanto al logro de sus objetivos, específicamente a través del control y desarrollo de los empleados. Pretende analizar el desempeño del individuo en su cargo, así como su potencial de desarrollo, mediante la evaluación de sus conductas en un periodo determinado de tiempo.

3. La evaluación de desempeño también puede definirse como una interacción formal y estructurada entre el subordinado y el supervisor que usualmente toma la forma de una revisión anual o semi-anual, en la que el desempeño laboral del subordinado es examinado y discutido para identificar las debilidades y fortalezas, así como las oportunidades para desarrollar competencias (Archer y North Associates, 2003)

⁸ Cardy y Dobbins (1986) señalan que se han desarrollado sólo modelos “fríos” de evaluación que excluyen los afectos y estados de ánimo como variables relevantes a pesar de la existencia de una investigación creciente que muestra que estas afectan los procesos cognoscitivos.

Objetivos

La evaluación de desempeño puede cumplir simultáneamente una serie de objetivos, lo cual ha sido uno de los principales problemas en el desarrollo de sistemas de evaluación (Hansen, 1994). El énfasis en alguno de ellos depende de los objetivos y la política general de gestión del recurso humano de la empresa (Hodgest en Cohn y Davico, 1994). Una de las posibles deficiencias que pueden presentar los sistemas de evaluación es lograr un gran número de objetivos, diversos y no priorizados adecuadamente. Seleccionar los objetivos a alcanzar es vital, ya que esto determina el método de evaluación y los posibles resultados (Huse y Bowditch en Cohn et al., 1994) Cabe señalar que, en la evaluación de desempeño, la organización y los individuos evaluados tienen sus propios objetivos, los que no necesariamente son antagónicos.

Objetivos de la Organización

La evaluación de desempeño puede cumplir tres grandes objetivos para la organización: mejorar el desempeño del recurso humano, mantener el control organizacional y medir la eficacia con que se usan los recursos humanos.

El objetivo principal de la evaluación de desempeño, según lo postulado por Cummings et al. (1985) y Chiavenato (2000), es mejorar el desempeño del recurso humano en la empresa mediante el incremento del desempeño o potencial de desarrollo de un empleado determinado y/o la administración más efectiva del recurso humano, lo cual se relaciona con las decisiones de promociones, compensaciones, etc. que se basarán en la información suministrada por la misma evaluación. En otro sentido, Werther et al. (1995) postula que: “El objetivo de la evaluación es proporcionar una descripción exacta y confiable de la manera en que el empleado lleva a cabo su puesto” (p. 233).

Cummings et al. (1985), por su parte, postula que mejorar el desempeño del recurso humano implica que la evaluación de desempeño posee dos funciones, evaluativa o apreciativa y evolutiva o de desarrollo, las cuales tienen diferentes campos, objetivos, resultados potenciales, etc.. Estas funciones, según al autor, originan diferentes sistemas de evaluación y pueden generar conflicto de roles en el

manager si se intentan desempeñar simultáneamente. Además, cuando coexisten las dos funciones en un mismo sistema de evaluación, la función evaluativa podría reducir o eliminar el valor evolutivo, ya que una gran parte de los empleados no admitirían tener bajo desarrollo en algunas áreas de desempeño, si un aumento del sueldo o una promoción dependiera de ello (Archer y North Associates, 2003). Sin embargo, según la teoría de las expectativas en motivación, no relacionar el desempeño con las recompensas organizacionales afectaría los niveles de esfuerzo y las metas del individuo, lo que afectaría a su vez el nivel de desempeño. Esto se agravaría, si las empresas distribuyen arbitrariamente las recompensas organizacionales.

En síntesis, optar por la función evaluativa o de desarrollo es un tema de debate, ya que hay investigaciones que apoyan una y otra postura⁹. No obstante, existiría consenso respecto de que la evaluación de desempeño no debe orientarse simultáneamente al desarrollo y la toma de decisiones administrativas (Cummings et al., 1985; Meyer, Kay y French en Ceberblom, 1980).

Figura N° 4: Comparación entre la Evaluación de Desempeño Evaluativa y Evolutiva

	<i>Función evaluativa o apreciativa</i>	<i>Función evolutiva o de desarrollo</i>
<i>Campo</i>	El desempeño en el pasado	Progresos en el desempeño futuro
<i>Objetivos</i>	Mejorar el desempeño mediante una administración más efectiva de las áreas del personal y las compensaciones. La evaluación suministra información válida y confiable.	Mejorar el desempeño o el potencial del desempeño del empleado a través del auto-estudio y el crecimiento.
<i>Métodos</i>	Procedimientos varios de clasificación y rango	Series de pasos evolutivos, Administración Participativa por Objetivos
<i>Resultados</i>	Discriminar entre el bajo y alto desempeño en los empleados.	Identificar las áreas que requieren crecer y mejorar.
<i>Rol del evaluador</i>	Enjuiciamiento. (Usa dimensiones como bueno - malo, eficaz- ineficaz, etc.)	Consejero (relación de ayuda)
<i>Rol del evaluado</i>	Pasivo o activo, frecuentemente defensivo	Involucración activa en el aprendizaje
<i>Requisitos para el</i>	Precisión en la medición	Precisión en la evaluación y deseo del empleado de mejorar su desempeño

⁹ Una investigación reciente de Bannister y Balkin (1990) señala que los evaluadores presentan un mayor grado de aceptación del proceso de evaluación y satisfacción con el sistema de evaluación, si el proceso está directamente relacionado con la distribución de recompensas. (en Archer y North Associates, 2003).

En la práctica, por otro lado, la mayoría de las empresas combinan dichas funciones por motivos de costos, tiempo, etc. (Quijano, 1994, Ramos, et. al, 1996). Un estudio efectuado por Campbell en 1970 determinó que de las 33 grandes compañías de Estados Unidos, sólo una había desarrollado sistemas diferenciados. El resto de las empresas combina 2 a 3 objetivos de evaluación. (Quijano, 1994).

En relación con los objetivos específicos relacionados con prácticas de Recursos Humanos, encontramos los siguientes:

1. Reclutamiento y Selección: Validar los predictores de desempeño usados en el proceso de selección mediante de la evaluación del desempeño de los postulantes empleados finalmente y la comparación del desempeño del predictor/desempeño del empleado contratado. (Cummings et al., 1985, Robbins, 2000, Werther, et al., 1995)

2. Capacitación: (a) Identificar, en un nivel individual o colectivo, las habilidades claves que se deben adquirir o desarrollar en un programa de capacitación. El método de incidentes críticos ha sido usado exitosamente. (Cummings et al., 1985; Parra, Robbins, 2000); (b) identificar a los empleados que se beneficiarían más de un programa de capacitación, ya que capacitar a los empleados con un desempeño satisfactorio o superior es un desperdicio de los recursos de la organización (Cummings et al., 1985); y (c) proveer información óptima para medir el impacto de la capacitación, ya que ésta se vincula con la determinación de las contribuciones de los empleados a los objetivos organizacionales (Cummings et al., 1985; Robbins, 2000, Werther et al., 1995).

3. Desarrollo: (a) Determinar los empleados de desempeño superior que podrían tener un potencial no aprovechado por la organización (Werther et al., 1995), (b) entregar información para la elaboración de un plan de carrera (Chon et al., 1989; Werther et al., 1995), y (c) orientar las decisiones sobre posibilidades laborales específicas. (Werther et al., 1995).

4. Colocación: Determinar quiénes serán promovidos, trasladados, descendidos e incluso despidos en base a la información suministrada por la evaluación de desempeño (Chon et al., 1989; Werther, 1995).

5. Diseño de puestos de trabajo: Determinar errores en el diseño de puestos o explicitar la necesidad de rediseño de un puesto determinado por cambios en la organización misma o su ambiente. (Werther et al. 1995).

6. Descripción de puestos: Analizar las características y el entorno del puesto desempeñado y revisar los objetivos fijados de cada puesto.

7. Compensación: Determinar quiénes recibirán compensaciones (bonos, aumento de sueldo) sobre la base del desempeño presentado.

Otros objetivos específicos que pueden orientar el desarrollo de evaluaciones de desempeño son:

(a) Mejorar el clima de la empresa mediante el enfrentamiento de problemas de relación interpersonal entre los distintos niveles jerárquicos; (b) contribuir a la detección de desafíos externos que influyen en el desempeño como la salud, aspectos financieros, (Werther et al. 1995); (c) proveer información básica para la investigación de recursos humanos (Chiavenato, 2000); (d) facilitar la operación y adaptación de cambios en el puesto de trabajo; (e) mejorar y detectar problemas en la supervisión (Chiavenato, 2000); y (f) motivar a los empleados a desempeñarse más eficazmente en su cargo. (Barnett, 1997). Al respecto, se requiere que la evaluación de desempeño no sea sólo un instrumento de valoración cuantitativa.

Las empresas sólo usan algunos de los objetivos mencionados. Robbins (2000) presenta los resultados de un estudio realizado en 600 empresas de Estados Unidos que establecen que los tres objetivos más utilizados a la hora de diseñar un sistema de evaluación son distribuir recompensas, dar feedback y tomar decisiones de promoción.

El segundo objetivo organizacional mencionado, mantener el control, se relaciona con el ciclo de trabajo (captar insumos tecnológicos, humanos, físicos, etc., procesar y transformar dichos insumos en productos como bienes, servicios y conocimientos; y traspasar estos productos al medio ambiente). La evaluación del desempeño y de los productos es fundamental para definir cambios en la captación de

insumos o técnicas de procesamiento (Ej. Automatizar ciertos procesos) con el objeto de contribuir a la supervivencia de la organización. Es decir, la evaluación de desempeño del recurso humano aporta información para evaluar el desempeño organizacional (Cummins et al., 1985). En un nivel individual, la evaluación de desempeño puede medir la eficacia del uso de los recursos humanos referido al ciclo descrito. Así cada persona al ingresar a una organización aporta una serie de competencias, conocimientos, aspiraciones, etc. que serán procesadas para convertirse en un determinado desempeño. La eficiencia de la captación y conversión debe ser evaluada, para ello, es necesario fijar objetivos de gestión relacionados con las prácticas y procesos organizacionales que fueron descritos (Ej. Evaluar la eficacia de la capacitación).

Objetivos del Individuo

El objetivo clave de la evaluación para el individuo es otorgar información, cuantitativa y/o cualitativa, sobre su rendimiento. Aunque existe variabilidad entre los empleados en cuanto a sus deseos por mejorar su rendimiento, en la mayoría podemos encontrar una necesidad cognitiva de información sobre si su desempeño cumple con los objetivos organizacionales, expectativas de sus compañeros, de sus supervisores, etc. (Billikopf, 2002)

La retroalimentación permite aprender más de sí mismo y si la retroalimentación tiene valencia positiva, satisfacer las necesidades de éxito y competencia. Sin embargo, cabe considerar que si bien el individuo busca una retroalimentación válida y eficaz, intenta ser evaluado tan positivamente como sea posible de modo de obtener recompensas organizacionales (Porter, 1975 en Barnett, 1997). Por otro lado, la retroalimentación de valencia negativa deteriora la autoestima y el concepto de sí mismo. Esto junto con la posibilidad de consecuencias administrativas como despidos y amonestaciones conlleva a que los empleados presenten resistencias frente al proceso de evaluación. (Quijano, 1994)

Objeto de la Evaluación

Medir el desempeño implica una serie de decisiones sobre el objeto de evaluación, ya que existen variados criterios. Así, en primer lugar, se debe decidir medir resultados y/o desempeño. En segundo lugar, si se ha seleccionado este último, se debe optar por evaluar conductas o rasgos de personalidad. A su vez, se debe definir si se utilizarán criterios simples, múltiples o compuestos. Todas estas decisiones tendrán un fuerte impacto sobre el comportamiento de los empleados, por tanto, en el desempeño.

Evaluación de Resultados versus Evaluación de Desempeño

Los resultados son los indicadores más objetivos de los logros laborales específicos (Billikopf, 2002). La recolección de datos está dada por el objetivo específico que será medido y por el tipo de producción y tecnología utilizados para alcanzarlo. Así, por ejemplo, los principales indicadores para los operarios de empresas manufactureras son la calidad y cantidad de trabajo producido y, para los gerentes de plantas, la cantidad producida, el desperdicio generado y el costo por unidad de producción, entre otros.

Evaluar resultados puede ser útil para distribuir recompensas organizacionales cuando no existe gran influencia de variables externas e incontrolables en la producción del trabajo (Cummings et al., 1985) y/o para validar prácticas organizacionales de gestión de recursos humanos. También es recomendable usar este criterio en puestos con una alta equifinalidad¹⁰ (Kane y Lawler en Quijano, 1994). No obstante, si se emplean los resultados como único criterio, los empleados pueden presentar un conjunto de conductas que favorezcan el logro de objetivos a corto plazo por sobre los de largo plazo o que atenten con los valores de la empresa. (Quijano, 1994, Robbins, 2000)

Evaluar el desempeño, a diferencia de evaluar resultados, sirve más para explicar el cómo un trabajador obtiene unos resultados determinados y las áreas en que un empleado es mejor o peor en relación con sus compañeros (Quijano, 1994, Ramos et. al., 1996). Esto a su vez, facilita la planificación

¹⁰ La equifinalidad del puesto se refiere al grado en que los empleados usan diferentes medios para alcanzar un mismo objetivo.

de acciones de mejora que sean oportunas. Cabe considerar además que la fijación y medición del logro objetivos individuales, en muchas empresas y cargos, es inviable porque los resultados son propios de un equipo de trabajo (Robbins, 1999) o simplemente no se disponen de medidas, lo cual se da mayormente en trabajos con orientación al servicio de personas (Ej. Bomberos, profesores, etc.)

Para zanjar esta discusión, Cravino (1997) plantea que: “El desempeño de una organización, grupo o persona está definido por una integración sistémica de lo que debió lograr en el pasado, logra en el presente y podría lograr en el futuro. Entendiendo el logro como una función integrada entre el QUÉ (objetivos/resultados) y el CÓMO (competencias/comportamientos)”¹¹. Es decir, siempre que sea posible, se debe evaluar desempeño y resultados.

Evaluación del Desempeño mediante Conductas o Rasgos de Personalidad

Robbins (1999) postula que los rasgos son el conjunto más débil de criterios. Se ha demostrado que la evaluación de rasgos es muy poco válida y fiable (Quijano, 1994). No obstante, éstas todavía son muy usadas por las organizaciones.

Según Koontz y Weihrich (1998), “La evaluación de los rasgos de personalidad equivale de este modo a sustituir la consideración de lo que realmente hace un individuo por la opinión de un tercero acerca de ese individuo” (pp. 419). El diseño de una evaluación orientada a los rasgos debe contemplar un análisis previo de la relación entre el desempeño y los rasgos a evaluar. Según el autor, el riesgo de un análisis inadecuado resta valor a la evaluación. Así, por ejemplo, “Rasgos como tener buena actitud,

¹¹ En torno al debate del qué y el cómo: Jack Welch, alto ejecutivo de General Electric expone que: “ Ninguna organización tiene dudas acerca de las personas que cumplen con el QUÉ y el CÓMO simultáneamente, presentan siempre el desempeño buscado. Tampoco hay demasiadas dudas cuando alguien no cumple con ninguno de estos dos elementos; independientemente de la acción que se tome, éste es claramente el desempeño que las organizaciones no desean. Nos obstante, si alguien no cumple con los objetivos, pero “hizo las cosas bien” debería dársele una segunda oportunidad. Por otra parte, el que alguien cumpla con los objetivos, pero mediante comportamientos que se oponen al sistema de valores de la empresa es uno de los grandes desafíos para la gestión de recursos humanos. Una prueba de madurez organizacional. Si la Dirección “perdona o minimiza” los desvíos en los CÓMO a partir del éxito de los objetivos logrados, el futuro de la empresa puede verse comprometido seriamente. Las organizaciones maduras, según Welch, tienen el coraje de decirle a un gerente exitoso que no comparte los valores organizacionales: “O cambia o se va”. (Cravino, 1997)

mostrar seguridad, ser confiable o cooperativo, parecer ocupado o poseer una rica experiencia, podrían no estar correlacionados positivamente con los resultados de la tarea” (pp. 419).

El diseño de evaluaciones, por lo tanto, debe evitar emplear rasgos como criterio. Es importante destacar que es posible traducir una característica personal en resultados específicos o conductas. Por ejemplo, “En lugar de referirse a responsabilidad (característica personal) de un trabajador, puede tomarse en cuenta la calidad de sus informes sobre las tareas terminadas (productividad)” (Billikopf, 2002, pp. 6).

Dado lo anterior, las organizaciones optan por evaluar comportamientos porque, en primer lugar, estas medidas presentan una mayor fiabilidad y validez así como un alto potencial de feedback; y en segundo lugar, las conductas son indicadores directos y fácilmente observables así como los evaluadores se encuentran familiarizados con ellas, estableciendo así un marco de referencia común para todos los evaluadores. Al respecto, Quijano (1994) expone que: “Resulta más fácil y seguro observar conductas concretas –sobre todo si sabemos de antemano cuáles son las que hay que observar-, y evaluar a un empleado de acuerdo con su ejecución de las mismas en su trabajo, que intentar hacerlo sobre sus cualidades o rasgos, cuyo contenido y manifestación empírica no resultan sencillos ni precisos” (pp. 279).

Criterio Simple, Múltiple o Compuesto

Un criterio simple representa una medida única del desempeño global de un empleado, la cual se deriva del juicio del evaluador sin que éste emplee ningún indicador formal. Un criterio compuesto, por otro lado, también representa una medida única, la cual se obtiene de la evaluación del desempeño del sujeto en diferentes dimensiones del trabajo. Por su parte, un criterio múltiple es aquél que agrupa diferentes medidas del desempeño de un empleado en varias dimensiones del trabajo. Este último es más recomendable para evaluaciones que estén orientadas a validar y controlar sistemas de gestión o a desarrollar al personal así como para evaluaciones de potencial. Cabe destacar que las ventajas de los dos primeros son las mismas. Específicamente, resultan muy útiles para tomar decisiones administrativas

(Quijano, 1994). No obstante, como el criterio simple es muy vulnerable a errores de medida, la verdadera de decisión es optar entre el criterio compuesto o el múltiple.

Ponderación de Criterios. Para ponderar los subcriterios se usan diferentes procedimientos. Dentro de los cuales, la combinación del análisis factorial con el juicio de experto es la estrategia más adecuada. También la predictibilidad y confiabilidad de los mismos pueden ser útiles. Al respecto, es importante destacar que no ponderar es una opción válida. Efectivamente, una parte de los investigadores sostiene que es imposible conocer el peso específico de los subcriterios, por tanto, la ponderación subcriterios introduciría sesgos y errores en la evaluación. (Quijano, 1994)

Evaluadores

El objetivo de la evaluación, las dimensiones de desempeño a evaluar (Cummings et al., 1985), así como la política de recursos humanos y ciertas características de la empresa como su cultura determinan la responsabilidad de evaluación, la que puede recaer en el supervisor o gerente, el empleado y/o el equipo de trabajo, etc. (Chiavenato, 2000).

Gerente o Supervisor

En la mayoría de las organizaciones, el 95% de los supervisores poseen la responsabilidad de evaluar a sus subordinados de los niveles jerárquicos bajo o medio, en base a los medios y criterios establecidos por el Departamento de Recursos Humanos (Robbins, 1999). Este último ejerce una función de staff relacionada con el control del sistema mientras que el supervisor conserva la autoridad de línea. (Chiavenato, 2000). Cascio, en Ramos et. al. (1996), plantea que “Se trata de la persona que probablemente mejor puede observar el desempeño real del empleado, y relacionarlo con los objetivos de la organización y del departamento” (pp. 338). Sin embargo, la evaluación del supervisor o directivo presenta también ciertos riesgos a considerar. Cumming et. al. (1985) señala: “En primer lugar, el ser evaluado por quien controla las compensaciones y los castigos, puede atemorizar al subordinado. Más aún,

el proceso de comunicación de la evaluación a menudo siente que debe defenderse y justificar sus acciones” (pp. 120). En segundo lugar, según Cumming et. al. (1985) y Chiavenato (2000), los supervisores pueden mostrarse reticentes a evaluar porque sienten que no están calificados para evaluar las contribuciones de sus empleados y objetan también el papel de “jugar a ser Dios” con las carreras de sus empleados. Según Chiavenato (2000), cabe considerar también que las nuevas formas de trabajo (equipos autodirigidos, trabajo a distancia, etc.) disminuyen la frecuencia de interacción entre jefes y subordinados, por lo tanto, también la oportunidad de observar el desempeño, lo que podría distorsionar los resultados de la evaluación. La evidencia, no obstante, indica que si la evaluación tiene fines administrativos, sólo directivos o supervisores pueden desempeñar el rol de juez y determinar recompensas y castigos, promociones y despidos. (Quijano, 1994)

Empleado

Cuando el empleado mismo es el que mejor que puede observar sus métodos de trabajo, resultados y comportamientos (King, 1984; Cummings et al., 1985), y cuando estamos ante organizaciones más democráticas donde se considera clave el autocontrol, la autodirección, etc.; se deben desarrollar autoevaluaciones (Chiavenato, 2000; Robbins, 1999). En éstas es el empleado es quien evalúa su desempeño, eficacia y/o eficiencia sobre la base de los criterios establecidos por la organización y que generalmente se orientan al desarrollo.

Las autoevaluaciones son positivas porque favorecen la participación del empleado en el proceso de evaluación y disminuye la defensividad ante la entrega de feedback en las reuniones de evaluación. Se pueden ocupar además para planear el mejoramiento del desempeño (Billikopf, 2002).

Los resultados de la autoevaluación, sin embargo, frecuentemente no son congruentes con las evaluaciones de los superiores (presentan puntuaciones más altas), lo que dificulta el proceso de retroalimentación. (Robbins, 1999) y genera tensiones entre subordinados y superiores. Según Quijano (1994), algunas condiciones que asegurarían la validez y fiabilidad de las autoevaluaciones son: (a) un

alto nivel de inteligencia y necesidad de logro así como un locus de control interno en los evaluadores, (b) entrega de feedback a los evaluadores sobre la validez y fiabilidad de la autoevaluación, (c) expectativa del evaluado que se comparará la autoevaluación con medidas objetivas de rendimiento, (d) uso de métodos de evaluación de grupos, (e) capacitación y (g) confidencialidad de la autoevaluación.

Compañeros de Trabajo

La evaluación de pares, en el pasado, ha tenido una baja aceptación de los usuarios, aún cuando, se ha demostrado que son válidas y confiables (Buller y McEvoy, 1987; Cederblom y Lounsbury, 1980, Quijano, 1994). Pocos estudios han explorado esta problemática, empero los resultados de los estudios son consistentes. Las evaluaciones de pares que se orientan a dar feedback y planificar el desarrollo de carrera tienen mayor aceptación que las evaluaciones cuyo objetivo es tomar decisiones administrativas.

Estas evaluaciones presentan una alta validez y confiabilidad porque: (a) los miembros del equipo se encuentran en interacción directa y constante con el evaluado (Borman y Lawler en Cederblom y Lounsbury, 1980; Robbins, 1990), (b) ofrecen la posibilidad de realizar múltiples evaluaciones, (c) los compañeros de trabajo observan comportamientos que usualmente el supervisor no puede observar (Borman y Lawler en Cederblom y Lounsbury, 1980), (d) proveen información más útil que otras fuentes (Borman y Lawler en Cederblom y Lounsbury, 1980), (e) son más estables en el tiempo (Smith en Buller y McEvoy, 1987) y (f) son más capaces de diferenciar el esfuerzo del desempeño y se focalizan en las competencias y habilidades relevantes para la tarea. Weslwy y Klimosky, Buller et al. (1987), concluyen que: "... potencialmente son los juicios más exactos sobre del comportamiento del empleado" (pp. 60).

Es posible, empero, encontrar una alta resistencia a implementar evaluaciones de pares tanto en managers como en los mismos evaluados. Las causas de esta resistencia se relacionan con: (a) la percepción generalizada que son concursos de popularidad (Buller et al. , 1987; Cederblom et al., 1980) por lo que las evaluaciones serían distorsionadas fuertemente por la amistad o animadversión entre los

miembros del equipo (Robbins, 1999)¹², (b) se anticipa o percibe efectos negativos sobre el grupo como conflictos entre los miembros, un clima de desconfianza, etc. (Medlan en Cederblom et al. 1980), (c) los administradores perciben que éstas disminuyen el control sobre las evaluaciones (Smith en Buller et al., 1987) y (d) se percibe además que un empleado que obtiene bajas puntuaciones puede calificar con bajas puntuaciones a sus compañeros por venganza (Denisi y Mitchell en Buller et al., 1987). Por esto último, la mayoría de las evaluaciones de pares son múltiples y anónimas. (Billikopf, 2002).

Su aceptación, por el contrario, se relaciona con el valor de la retroalimentación individual para los programas de entrenamiento como para el desarrollo de carrera (Rodman en Cederblom et al., 1980) . La aceptación se incrementa si las evaluaciones por los pares se orientan al desarrollo por sobre la toma de decisiones administrativas (Cederblom et al., 1980; Buller y McEvoy, 1987) y si los ratings previos son satisfactorios para el empleado (Buller y McEvoy, 1987). Además, los últimos investigadores encontraron que la aceptación del usuario se correlaciona significativa y negativamente con la percepción de distorsión por amistad.

Cabe acotar que la aceptación varía con la antigüedad de los usuarios, ya que los empleados de mayor antigüedad perciben negativamente la posibilidad de ser evaluados por miembros más jóvenes e inexperimentados (Buller et al., 1987).

Órgano de Gestión del Personal y Comité de evaluación

Las organizaciones que ocupan este tipo de evaluadores se caracterizan por altos niveles de centralización y burocracia (Chiavenato, 2000). En el primero, el departamento de recursos humanos evalúa apoyándose de la información suministrada por cada gerente. Como no se trabaja con el desempeño individual, los programas de acción son generales para las unidades o departamentos de la organización. Por su parte, los comités de evaluación se componen de un grupo determinado de empleados

¹² Cederblom et. al. (1980) señala que varias investigaciones indican que la distorsión cuya fuente es la amistad no invalidaría la valides en las evaluaciones de pares.

de diferentes áreas de la organización, tanto permanentes (gerente general, director o gerente de gestión de personal, etc.) como transitorios (gerente o supervisor del evaluado) (Chiavenato, 2000).

Evaluación de 360°

Es el enfoque más moderno, siendo usado en el 26% de las compañías norteamericanas (Robbins, 1999). Los evaluadores son los individuos que mantienen cierto grado de interacción con el sujeto a evaluar y su trabajo. Así, se incluye el gerente o supervisor, clientes internos, subordinados, equipo de trabajo y/o proveedores por lo cual el número de evaluadores varía, siendo el promedio de 5 a 10 evaluadores (Robbins, 1999). Su valor radica en el potencial de retroalimentación, su confiabilidad y facilitación de la participación de todos los actores de la organización en el proceso.

Se debe destacar que algunas investigaciones han dado cuenta de la alta varianza entre los resultados de las evaluaciones efectuadas por supervisores, subordinados, pares y las autoevaluaciones. Esto se ha explicado en relación con: (a) el aumento del número de evaluadores, (b) los evaluadores con diferentes posiciones organizacionales consideran, en forma diferente, cuáles aspectos del desempeño son relevantes¹³ y (c) los evaluadores usan en forma diferente las escalas de un mismo instrumento (London, Rowland y Zammuto, 1982). En relación con el segundo punto, las conclusiones del estudio efectuado por los autores (1982) establece que: un buen diseño de un método de evaluación debe considerar que cada tipo de evaluador emplearía diferentes modelos de desempeño, o sea, asigna diferentes peso a los criterios de un mismo instrumento para determinar el desempeño global. Los supervisores y compañeros evalúan el desempeño del subordinado o colega respecto de cómo éste afecta a su propio desempeño. Los supervisores evalúan así la cantidad y calidad de trabajo, así como la obediencia de la normativa; los pares principalmente evalúan los aspectos relacionales del trabajo como habilidades de comunicación y trabajo

¹³ El estudio de estos autores descubrió que para evaluar el desempeño, compañeros y supervisores se centran principalmente en la calidad y cantidad, organización del trabajo y apariencia personal mientras que las enfermas que realizaron autoevaluaciones en variables internas (esfuerzo, habilidad, aceptación de responsabilidad). Esto permite explicar por qué en general los resultados de las evaluaciones entre compañeros y supervisores son más congruentes entre sí que comparados con los de las autoevaluaciones. Estas observaban puntuaciones más altas precisamente en calidad y cantidad de trabajo. Así como efectividad total.

en equipo. Por su parte, cuando uno autoevalúa su desempeño se basa principalmente en variables internas y controlables como el esfuerzo, las competencias individuales y el compromiso.

Métodos

La evaluación de desempeño es una función clave en toda empresa moderna que puede ser realizada mediante diferentes métodos. De este modo, en una misma empresa se pueden usar métodos diferentes para diferentes cargos o combinar dos o más métodos para evaluar el desempeño de un empleado determinado

La adecuación entre la organización y el método es de importancia para la obtención de resultados. Cabe recordar que la selección de un método depende de los objetivos de la evaluación de desempeño que fueron seleccionados por la empresa, el tipo de trabajo, el número de empleados a evaluar, cultura evaluativa, entre otros.

Resulta útil diferenciar entre los métodos basados en el desempeño pasado y los métodos orientados a evaluar el desempeño futuro. Los primeros son considerados los métodos tradicionales para evaluar el desempeño, incluyen el método de Escalas gráficas, Elección forzada, Investigación de campo, Incidentes críticos y mixtos. Los segundos, en cambio, se centran en la evaluación del potencial del empleado o en el establecimiento de objetivos. Dentro de estos, se encuentran el método Administración por Objetivos y las evaluaciones psicológicas (Werther, 1995).

Escala Gráfica de Calificación

Las escalas gráficas de calificación de puntuación son el método más antiguo y usado (Chiavenato, 2000; Robbins, 1999; Werther, 1995). En éste el evaluador entrega una apreciación subjetiva del desempeño, lo que se traduce en que por cada factor, se asigna un valor en una escala continua, semicontinua o discontinua. Estos valores se traducen usualmente en incrementos salariales proporcionales a la puntuación total.

Algunas de las ventajas y desventajas que ofrece éste método son:

1. Facilita el análisis y las comparaciones entre empleados mediante los resultados cuantitativos. (Chiavenato, 2000; Robbins, 1999)

2. Fácil desarrollo y comprensión. Esto último disminuye el tiempo para su puesta en práctica porque los evaluadores requerirían poca capacitación (Werther, 1995)

3. Se puede administrar a un gran número de empleados (Werther, 1995). Esto se vincula con que el registro de la evaluación demanda poco tiempo al evaluador (Chiavenato, 2000)

4. Según Chiavenato (2000): “Posibilita una visión integrada y resumida de los factores de evaluación, es decir, de las características de desempeño más destacadas por la empresa y la situación de cada empleado” (pp. 371)

5. El evaluador puede distorsionar los resultados en relación con el error de indulgencia, efecto halo, prejuicios personales, etc. Además, según Chiavenato (2000): “No permite mucha flexibilidad al evaluador; en consecuencia, el evaluador debe ajustarse al instrumento, y no éste a las características del evaluado” (pp. 367). Esto podría potenciar que el evaluador forzará los resultados de la evaluación para que estos se adaptasen a otros criterios que el evaluador considera como importantes.

5. Tiende a rutinizar y generalizar los resultados de las evaluaciones (Chiavenato, 2000), ya que: “... reduce los resultados a expresiones numéricas mediante la aplicación de procedimientos matemáticos y estadísticos para corregir las distorsiones personales introducidas por los evaluadores” (pp. 367). Además, como estos resultados son muy generales, la utilidad de la retroalimentación para mejorar el desempeño del empleado es baja (Werther, 1995).

Método de Selección Forzada

El evaluador debe seleccionar la oración que mejor describa el desempeño del empleado. Las oraciones descriptivas se presentan en grupos de dos, tres o cuatro oraciones, los que contienen oraciones con valencia positiva o negativa (Werther, 1995). Posteriormente, se suman los puntajes alcanzados en

cada factor de desempeño. Cabe mencionar que la selección de oraciones se realiza mediante procedimientos estadísticos para observar su adecuación a los criterios existentes de la empresa y su capacidad de discriminación. (Chiavenato, 2000)

Las ventajas de este método son: (a) los resultados son menos influenciados por los sesgos del evaluador, (b) es fácil administrarlo y estandarizarlo, (c) requiere poca capacitación de los evaluadores, y (d) se puede adaptar a una gran variedad de puestos. (Werther, 1995)

Las desventajas, por otro lado, están dadas por una elaboración e implementación muy complejas y unos resultados muy genéricos. Respecto a esto último, la evaluación sólo discrimina entre el desempeño insatisfactorio, satisfactorio y sobresaliente sin aportar más información por lo que éste método posee un reducido potencial para entregar retroalimentación al evaluado y al evaluador, y necesita de evaluaciones complementarias para detectar necesidades de capacitación y desarrollo laboral (Chiavenato, 2000). Además, cuando los grupos obligan a seleccionar entre dos o más oraciones de valencia positiva, el empleado puede percibir como injusta la selección de una frase por sobre otra (Werther, 1995)

Lista de Verificación

Un método muy similar es la *Lista de verificación*. En ésta el evaluador debe escoger las oraciones que mejor describen el desempeño del empleado. La diferencia es que éstas no son presentadas en conjuntos de oraciones descriptivas con valencia positiva y negativa, o sea, las oraciones no compiten entre sí.

La lista de verificación con valores puede asignar un valor diferencial a los ítems para obtener un valor del desempeño global, pero generalmente, el evaluador habitualmente desconoce la ponderación de los ítems. Si hay suficientes puntos puede lograrse una descripción detallada del desempeño del empleado, pero si las afirmaciones son amplias o vagas se reduce el grado de relación con el puesto específico.

Las ventajas son muy similares al método anterior, no obstante, el diseño de éste método es más fácil y económico. Dentro de las posibles desventajas podemos encontrar que es muy susceptible a las distorsiones del evaluador, el cual además puede interpretar equivocadamente algunas frases, y no permite asignar puntuaciones relativas.

Método de Registro de Acontecimientos Críticos

En este método, el evaluador describe, tan pronto como es posible, los comportamientos claves del empleado que hayan sido especialmente ineficaces o eficaces. Si existe influencia de variables externas, no deben incluirse como un acontecimiento propiamente tal. Determinar que comportamiento depende o escapa al campo de acción del empleado es una tarea difícil para el evaluador (Werther, 1995). Si esta técnica se utiliza sola, los supervisores y empleados pueden tener dificultades para traducir la información del registro en acciones tendientes a mejorar el desempeño. Cabe acotar además que como sólo se describen los comportamientos excepcionales, no se podrá realizar una evaluación óptima durante un determinado periodo de tiempo si sólo hubo comportamientos normales (Billipokf, 2002)

Las fortalezas que ofrece el registro de incidentes críticos son: (a) disminuir el efecto de acontecimientos recientes, (b) poseer un alto potencial de retroalimentación, ya que proporciona ejemplos de los comportamientos deseables, y (c) detectar necesidades de capacitación (Werther, 1995)

Dentro de las falencias que posee, se observa que, en primer lugar, los evaluadores tienden, con el paso del tiempo, a disminuir su registro de los acontecimientos críticos, pero cuando se aproxima la evaluación, lo aumentan. Esto aumenta el efecto de los acontecimientos recientes. En segunda instancia, los evaluados considerarían que el efecto de eventos negativos registrados se mantiene durante todo el periodo de evaluación y afecta exageradamente su evaluación de desempeño (Werther, 1995).

Escalas Basadas en la Conducta

Las escalas conductuales se centran en conductas por sobre rasgos, lo que se significa, según Quijano (1994): "... centrarse en acciones físicas, mentales o interpersonales, realizadas por los trabajadores" (pp. 307). Este tipo de escalas no sólo sirve para evaluar el desempeño sino para realizar análisis del puesto de trabajo así como detectar necesidades de capacitación

Este método combina elementos del registro de incidentes críticos y las escalas gráficas de puntuación. Para disminuir la distorsión y la subjetividad, se compara el desempeño del empleado con parámetros conductuales específicos y objetivos que indican en que grado el desempeño es aceptable o inaceptable. Es decir, para cada factor o dimensión de desempeño, se presentan comportamientos reales (equivale a los puntos de las escalas gráficas) en una escala discontinua. Los parámetros conductuales se obtienen en el diseño de puestos o mediante entrevistas a empleados y supervisores, lo que proporciona ejemplos de los comportamientos a considerar. En esto radica su alto potencial de retroalimentación, ya que "Las escalas de comportamiento también ofrecen a los empleados ejemplos específicos del tipo de comportamiento a realizar (o evitar) si quieren progresar en la organización. Además, las escalas de comportamiento animan a los supervisores a que sean precisos cuando informan posteriormente a los empleados" (Gómez- Mejía, Balkin y Crady, 2001, pp. 251).

Las escalas basadas en conductas pueden ser de dos tipos, Escalas de Observación de Comportamiento (B.O.S.) y las Escalas Ancladas en la Conducta (B.E.S.).

Escalas de Observación de Comportamiento (B.O.S.)

Las primeras se centran fuertemente en: "... identificar y describir aquellos comportamientos específicos, que deciden la consecución eficaz de los objetivos del puesto, y componen el conjunto de tareas consideradas como esenciales en el mismo" (Quijano, 1994, pp. 308). En cuanto al diseño de estas escalas, las siguientes son las fases a seguir:

Definición de objetivos. Se delimitan los objetivos del puesto en función a los objetivos generales de la organización. Hay que considerar que este paso se obvia, si los objetivos ya están definidos.

Identificación y especificación de los indicadores. Se determinan las conductas claves que se dan en el ejercicio del puesto. Esta fase, implica disponer o elaborar un cierto tipo de análisis de puesto. Éste puede requerir la aplicación de entrevistas personales, grupos de discusión, cuestionarios o planillas de observación. Al respecto, la entrevista de incidentes críticos resulta especialmente útil.

Análisis de datos. El simple análisis de los incidentes críticos revela las categorías o dimensiones del desempeño, los cuales contiene conductas positivas y/o negativas para la consecución de los objetivos. No todas las dimensiones parecen tener la misma importancia, sin embargo, la ponderación y selección de factores debe ser fijada sobre la base de criterios estadísticos (Quijano, 1994). Al respecto, el autor plantea “Por otra parte, y como es obvio, las técnicas estadísticas de que disponemos nos permiten analizar y corroborar empíricamente las dimensiones y factores hipotetizados mediante la intuición. El análisis factorial confirmatorio puede resultar, en la práctica, de gran utilidad para este propósito” (pp.313)

Elaboración del cuestionario. Se elaboran los ítems por factor o dimensión, los cuales pueden referirse a conductas positivas o negativas. Cabe destacar que la reconversión de los puntajes de los ítems con conductas negativas puede conllevar a errores en el registro así como en el cálculo de los puntajes por lo que debe hacerse con especial cuidado. Posteriormente, los ítems deben ser distribuidos aleatoriamente.

En relación con las alternativas de respuestas, se pueden utilizar las escalas tipo Likert o procedimientos psicofísicos. En cuanto a las primeras, se puede usar una escala que muestre el grado de acuerdo o desacuerdo con lo enunciado por un ítem o la frecuencia de los comportamientos dentro de una escala de 5 o más grados como nunca, casi nunca, regularmente, casi siempre o siempre (Llorens, 2000). Es importante notar que Dipboye, en Llorens (2000), propone una escala de 1 a 5, en la cual el 1 representa que la conducta se presenta el 64% de las veces; el 2, 65 a 74%; el 3, 75 a 84%; el 4, 84 a 94% y el 5, el 99% de éstas.

Escalas Ancladas en el Comportamiento

Este tipo de escala se orienta a evaluar las conductas que cabe esperar de un empleado en determinadas situaciones. Es decir, no emplea como criterio el desempeño presentado por el empleado, sino el desempeño futuro. Es más, sus ítems indican sólo algunas de las conductas que se podrían presentar. Para ser efectivo y práctico sólo se pueden usar un reducido número de parámetros (Wherther, 1995; Robbins, 1999). Las fases del diseño de estos cuestionarios son diferentes a las de las escalas de observación de comportamientos, ya que cada fase requiere de la participación de una grupo diferente de empleados, jefes y directivos que conozcan el trabajo (Quijano, 1994). El proceso que se da del siguiente modo:

1. Un primer grupo identifica y define las dimensiones relevantes a evaluar así como define los niveles de desempeño (alto, medio, bajo) por dimensión, entregando además ejemplos concretos.
2. Un segundo grupo, reclasifica los ejemplos de conductas, sin conocer a cuál dimensión pertenecían. Se seleccionan los ejemplos en los que hay un porcentaje considerable de acuerdo entre el primer y segundo grupo.
3. Un tercer grupo, reclasifica las conductas según nivel de desempeño, los cuales son más numerosos que los descritos en el primer paso. Se revisa también el consenso en la clasificación.
4. Se escogen los ítems y se formulan en términos de conductas esperadas para luego efectuar un estudio piloto y hacer las modificaciones en el cuestionario.

Método de Investigación de Campo

El método exige que un especialista de recursos humanos solicita información al superior directo sobre el desempeño del sujeto a evaluar. Tras procesar dicha información, se elabora un informe para ser discutido con el superior directo, el que a su vez lo discutirá con su subordinado. El representante del área de personal registra las puntuaciones y conclusiones.

Las ventajas y desventajas que encierra son las siguientes:

1. Aumenta la confiabilidad y comparabilidad, pero esto se traduce en mayores costos y disminución de la viabilidad dado que se requieren de múltiples entrevistas supervisor-especialista.

2. Según Chiavenato (2000), “Permite ver las causas, orígenes y motivos del desempeño. Es un 3. método de evaluación más amplio que permite, además de emitir un diagnóstico del desempeño del empleado, planear junto con el superior inmediato su desarrollo en el cargo y la organización” (pp.375). En el mismo sentido, el autor también señala que contribuye a detectar necesidades específicas de capacitación, tomar decisiones administrativas (promoción, desvinculación, etc.), planear las acciones para remover los obstáculos o proveer de recursos de apoyo para mejorar el desempeño, y recalca la responsabilidad de línea y función de staff en la evaluación de desempeño.

Métodos de Evaluación en Grupo

Se basan en la comparación del desempeño del empleado y uno o más de los miembros de la unidad de trabajo que es efectuada generalmente por el supervisor. Su utilidad se relaciona con la toma de decisiones administrativas como remuneraciones o incentivos basados en el desempeño, promociones y otras recompensas extrínsecas. Estos métodos no se ven influenciados por el error de indulgencia y la tendencia a la medición central, poseen una alta confiabilidad, pero no entregan retroalimentación específica al empleado porque habitualmente los resultados no se publican para evitar competencias internas dentro de las unidades de trabajo y además son muy genéricos (Cummings et al., 1985; Wherther, 1995), por lo tanto, es aconsejable combinarlos con otro método de evaluación. Además se debe considerar que: (a) los grupos pequeños de empleados pueden no presentar una distribución normal (Billipokf, 2002; Cummings et al., 1985), (b) el grado de diferencias entre el desempeño de los empleados no se distribuye homogéneamente (Billipokf, 2002), y (c) el orden de presentación de los empleados al momento de realizar la evaluación puede distorsionar los resultados¹⁴.

¹⁴ Por ejemplo, si se observa el desempeño de un empleado con desempeño superior, el siguiente empleado podría ser calificado más negativamente que lo debido; y si un empleado con desempeño bajo es seguido de un trabajador

Los tres métodos de evaluación de grupo más importantes son:

Método de categorización individual. El supervisor elabora una jerarquía del desempeño global de los empleados de mejor a peor (Werther, 1995, Robbins, 1999). Es un procedimiento fácil de implementar y utilizar si el evaluador no tiene que evaluar más de 15 a 20 empleados (Quijano, 1994). La distorsión más usual está dada por las inclinaciones personales y los acontecimientos recientes, pero al existir pluralidad de evaluadores disminuye la subjetividad.

Método de distribución forzada. El supervisor debe ubicar una fracción de los empleados en cierta categoría de desempeño. Este método disminuye el efecto halo, la tendencia a la medición central y el excesivo rigor o indulgencia, pero también puede forzar a que ciertos empleados reciban puntuaciones que no equivalen a su desempeño real. Cabe acotar que los resultados pueden distorsionados por los acontecimientos recientes.

Método de comparación por parejas. El evaluador debe comparar a cada empleado con cada miembro del equipo sobre la base del desempeño global. Dado lo cual, sólo es aplicable a equipos de trabajo pequeños. Posteriormente, se suman las veces en que el desempeño de cada empleado es seleccionado como el mejor de cada par para obtener un indicador y generar una jerarquía. La distorsión proviene de las inclinaciones personales y los acontecimientos recientes. No obstante, este método disminuye la tendencia de medición central y indulgencia o excesivo rigor.

Evaluaciones Psicológicas y Centros de Evaluación

Un equipo de psicólogos evalúa el potencial o desempeño futuro del individuo en cuanto a sus características intelectuales, motivacionales, emocionales, entre otras; mediante entrevistas en profundidad, exámenes psicológicos, entrevistas con supervisores, etc. Estas evaluaciones generalmente se usan para determinar el desarrollo de carrera dentro de la organización. Como su costo es alto, se emplean usualmente para evaluar a gerentes jóvenes y brillantes directivos de alto nivel. El segundo método, por

con desempeño medio, este último podría ser calificado con puntuaciones más altas, no reflejándose así el desempeño real. (King, 1984)

su parte, además de usar las técnicas mencionadas, realiza un estudio de antecedentes personales y ejercicios de simulación. Los resultados son útiles para el desarrollo de carrera y las decisiones de ubicación, teniendo -según los especialistas- una precisión del 75% (Werther, 1995).

Administración por Objetivos

Este método que constituye una alternativa a los sistemas tradicionales de evaluación, tiene por finalidad el desarrollo y motivación de los empleados (Cummings, 1985, Quijano, 1994). En cuanto la aplicación del mismo, en primer lugar, se definen los objetivos a lograr. La definición se realiza conjuntamente entre el supervisor y el subordinado. Es importante notar que la participación del evaluado en el establecimiento de objetivos potencia la responsabilidad y compromiso con los mismos (Quijano, 1994). El método establece que los objetivos deben ser mensurables, alcanzables y definidos con precisión. En segundo lugar, se evalúan periódicamente los progresos en el logro de estos de modo de efectuar los ajustes necesarios (Werther et. al., 1995). Esto implica, a su vez, retroalimentación frecuente. En tercer lugar, se evalúa formalmente el cumplimiento de los objetivos, lo que requiere de una autoevaluación que se plasma en un informe así como de la realización de una entrevista de evaluación. Al respecto, según Quijano (1994), en la entrevista basada en objetivos, se analizan los resultados obtenidos así como se fijan nuevos objetivos y los medios para alcanzarlos. La aplicación de este tipo de entrevista se ha extendido a otros métodos de evaluación. El mismo autor señala que se pueden usar los resultados de escalas basadas en la conducta para fijar objetivos comportamentales.

Las fortalezas y debilidades de éste método dicen relación con:

1. El establecimiento de los objetivos depende de cada individuo, o sea, los objetivos entre trabajadores de una misma unidad pueden ser diferentes. Esto plantea dificultades para vincular el logro de objetivos con la distribución de recompensas, ya que esto alentaría al individuo a seleccionar sólo los objetivos que garanticen o faciliten la obtención de recompensas. Además, se dificulta la comparación entre los empleados (Cummings, 1985).

2. Permite relacionar los objetivos organizacionales e individuales.
3. Los empleados obtienen el beneficio de carácter motivacional de contar con objetivos específicos (Werther et. al., 1995).

Propiedades de los Instrumentos de Evaluación

Thorndike, ya en la década del cuarenta, expuso que los instrumentos de evaluación deberían ser evaluados sobre la base de su confiabilidad, validez, viabilidad y libertad de influencia de sesgos. Si bien las propiedades psicométricas han recibido mayor atención, las propiedades pragmáticas o de utilización son igualmente importantes. (Quijano, 1994; Ramos et. al. 1996)

Validez del Instrumento

La validez, en términos generales, establece si un instrumento mide lo que dice medir. Ésta propiedad es relativa, es decir, es una cuestión de grado más que de todo o nada. En efecto, la validez depende del propósito, la situación de prueba y la población en la cual el instrumento es aplicado (Garrett, 1999, Nunnally y Berstein, 1995).

La validez de un instrumento determina el potencial de generalización científica del mismo. En Psicología y ciencias afines, es de vital importancia por la gran distancia que existe entre el objeto de medición y el instrumento. En otras palabras, la relación entre lo que se quiere medir y lo que se mide no es tan evidente ni fácilmente demostrable (Bartolomé en Rosales, 1997). Si este fuera el caso, no habría necesidad de estudiar la validez de las medidas. Dado lo anterior, según Nunally et. al. (1995), si bien usualmente los instrumentos son validados de modo independiente a su diseño es: "... muy recomendable que cualquiera que desarrolle un instrumento de medición piense en la manera en que se podría establecer su validez al principio y realizar al menos un estudio de validación preliminar para usarlo en la etapa de desarrollo" (pp. 94).

En el área de la evaluación de desempeño, la validez se equipara al concepto de relevancia del trabajo, el cual se define como el grado en que un instrumento mide los aspectos relevantes de un puesto de trabajo determinado. Implica, como señala Cascio en Ramos et. al. (1996) que: "... hay un vínculo claro entre los elementos críticos del trabajo, identificados por el análisis de puestos y las dimensiones en la que los sujetos son evaluados." (pp. 341) Al respecto, Cummings et. al. (1985) comenta que si bien el problema es fácil de comprender y apreciar en un nivel teórico, es extremadamente difícil de resolver en el nivel operativo.

La validez, según Nunally et. al. (1995), tiene tres significados principales, validez de contenido, constructo y concurrente. Todos estos se aplican también al desarrollo de instrumentos de evaluación de desempeño. Es importante considerar que la validación de un instrumento es un proceso continuo que busca establecer si los instrumentos se comportan de la manera que deberían.

Validez de contenido

La validez de contenido se entiende como el grado en que el conjunto de reactivos de un instrumento constituye una muestra adecuada y representativa del contenido específico que se pretende evaluar (Hernández et. al., 1998, pp. 236). Para determinar este tipo de validez se emplean dos procedimientos de juicio experto (Garrett, 1999). El primero se relaciona muy estrechamente con el concepto de validez aparente. Al respecto, Garrett (1999) comenta: "Se dice que un test tiene validez aparente cuando parece medir cualquier cosa que el autor propuso; es decir, lo que él cree estar midiendo" (pp. 392). Los juicios de validez aparente permiten al autor de un instrumento seleccionar los ítems apropiados que constituirán el mismo. Éstos deben ser sometidos posteriormente a juicio experto. En cuanto a este segundo procedimiento, en el caso de la evaluación de desempeño, expertos verifican que el contenido de los ítems realmente se refiera al desempeño. Garrett (1999) también plantea que el juicio experto será satisfactorio si el muestreo de reactivos es amplio y juicioso. En la misma línea, Nunally et. al. (1995) acota que: "Las pruebas que se relacionan con más claridad con el trabajo les van mejor en las

pruebas de validación de contenido” (p.412). Es decir, es de vital importancia basar la construcción y el muestreo de los reactivos en los resultados de algún análisis del puesto de trabajo.

Validez de Constructo

La validez de constructo es la validez más significativa desde un punto de vista científico. En términos generales, los constructos son variables abstractas y latentes más que observables y concretas (Nunally et. al. 1995). Estos determinan qué conductas han de relacionarse para su observación así como se deduce que éstas se verán afectadas, de igual manera, por manipulaciones experimentales (Nunally et. al. 1995, Rosales, 1997). De este modo, los constructos pueden explicar la varianza en los puntajes obtenidos. Es decir, explican las diferencias individuales en los puntajes de las pruebas sobre la base de propiedades psicológicas (Kerlinger, 1998), lo cual permite, en el campo de la evaluación de desempeño, planificar acciones de mejora.

Los constructos varían en el dominio de variables observables asociadas por cuanto éste puede ser pequeño o grande y estar definido en términos generales o específicos. Al respecto, Nunally et. al. (1995) comenta que:

“El tamaño y dominio del constructo y su especificidad se relacionan muy estrechamente, entre mayor es el dominio de observables relacionados con el constructo, es más difícil especificar las variables que pertenecen al dominio. El dominio de observables relacionados de muchos constructo es confuso y los científicos a menudo están inseguros de los completos de sus propios constructos. De forma típica, sostienen creencias firmes acerca de los observables prominentes relacionados con un constructo pero sólo pueden especular que tanto se extiende este último.” (pp. 97)

El procedimiento para desarrollar o validar constructos, según Nunally et. al. (1995), implica seguir uno o más de las siguientes estrategias: (a) especificar el dominio de observables relacionados con el constructo, (b) determinar el grado en que los observables tienden a medir lo mismo o cosas diferentes y (c) realizar estudios para ver si las medidas de un mismo constructo se correlacionan entre sí o con otros constructos. Estas estrategias no son aplicadas en un orden específico.

Especificación del Dominio de Observables. La especificación es una tarea difícil para los investigadores. De hecho, según Nunally et. al. (1995), establece que no existe un método preciso. El mismo autor además postula que delimitar el constructo implica explicitar lo que uno entiende con el uso de palabras como “ansiedad”, en otras palabras, definir un concepto usando palabras más concretas. Esto supone el uso de una teoría más o menos formalizada. Cabe destacar que: “En lugar de definir estrictamente el dominio inicial de observables para el constructo, la naturaleza del dominio por lo general es sugerida por numerosos intentos de desarrollar medidas particulares del constructo. De manera posterior, algunos investigadores intentan una delimitación explícita del dominio del contenido” (Nunally et. al., 1995, pp. 97). Dado lo anterior, en el área de la evaluación de desempeño, se requiere especificar las conductas que lo constituirían. Es vital considerar que como las conductas que lo conforman son situacionales, o sea, varían en función de las circunstancias y además como se relacionan con puestos de trabajo determinados; se ha progresado poco en la dirección de obtener modelos conceptuales del desempeño (Warr y Cornner en Ramos et. al., 1996, pp. 333)

Determinación de relaciones entre observables. Se intenta determinar empíricamente aquí las intercorrelaciones entre observables para ver si miden la misma cosa o no (Nunally et. al. 1995). Las medidas que emplean un mismo dominio de observables son afectadas de la misma manera por los tratamientos experimentales o presentan altas correlaciones en los sujetos en los estudios de diferencias individuales. Si este no es el caso, es posible que estemos ante la presencia de dos constructos diferentes.

Determinación de las relaciones entre constructos. Este paso supone la existencia de una teoría que explicité las relaciones entre dos o más constructos mediante hipótesis científicas (Nunally et. al., 1995). Es importante destacar que, en este punto, se requieren constructos bien definidos.

Cabe destacar, por último, que los procedimientos que se emplean para determinar la validez de constructo son variados, siendo el más destacado el análisis factorial. Así, según Rosales (1997), los estudios exploratorios pueden emplear: (a) análisis correlacionales y estudios de características técnicas de instrumentos de medida, (b) análisis factorial exploratorio, (c) análisis de conglomerados y análisis de

correspondencias. En estudios confirmatorios, se usan: (a) estudios experimentales (b) análisis factorial como técnica de validación de constructo, (c) matriz multirasgo-multimétodo, (d) modelos estructurales causales, (e) análisis factorial confirmatorio, (f) análisis cluster confirmatorio y (g) estudios diferenciales.

Validez de Criterio Externo

La relación entre los resultados de un instrumento que busca ser validado y los resultados de otra medición (criterio externo) determina la validez referida a criterio externo (Hernández, 1998). Garret plantea que: “Un criterio puede ser una medida objetiva del rendimiento o una medida cualitativa tal como el juicio del carácter o de la excelencia de un trabajo hecho” (1999, pp. 393), es decir, en cuanto a la evaluación de desempeño, el criterio externo a usar son los resultados u otras evaluaciones de desempeño.

La elección y obtención de un criterio adecuado es vital y difícil. Al respecto, Garret (1997) plantea que “Tal vez la mayor dificultad con la que se tropieza la validación de los test es el problema de hallar criterios auténticos. A menudo, los criterios tienen que ser necesariamente aproximados e indirectos, pues si dispusiéramos de criterios fácilmente accesibles y confiables, se utilizarían estas medidas en lugar de los test” (pp. 399). En la misma línea, Cronbach en Rosales (1997) añade que: “... al profesional que le interese la validez con fines prácticos (pedagogo, psicólogo) le concierne encontrar no un criterio imperfecto, sino un criterio de cualquier manera, en especial en el campo diagnóstico, puesto que no puede esperar un claro y unitario criterio...”

Existen dos tipos de validez según el criterio usado *validez concurrente* y predictiva. La primera usa un criterio fijado en un mismo tiempo que el instrumento testeado (criterio presente) mientras que la segunda emplea un criterio fijado en el futuro. Esta última es más importante para los instrumentos que se usan para tomar decisiones sobre las personas. Específicamente, en el área de la evaluación de desempeño, es clave para el instrumento que se usa para predecir el desempeño que un empleado presentará en otro puesto de trabajo, lo cual se relaciona estrechamente con la evaluación de potencial.

El procedimiento seguido para determinar la validez de criterio es de índole correlacional. Si la correlación entre los resultados del instrumento y el criterio externo son altas, el instrumento es válido. Al respecto, no se necesitan otras normas o procedimientos (Nunally et. al. 1995).

Confiabilidad

La confiabilidad de un instrumento, en términos generales, es el: "... grado en que su aplicación repetida al mismo sujeto u objeto produce iguales resultados" (Hernández et. al., 1998). La preocupación en la confiabilidad yace en la necesidad misma de obtener medidas válidas así como de confiar en éstas (Guilford, 1984).

Una medida es confiable cuando es completamente exacta, o sea, está libre de errores. No obstante, toda medición contiene error. En el desarrollo y aplicación de instrumentos, se debe minimizar lo más posible y determinar el grado en que éste afecta la exactitud de nuestras mediciones. En términos generales, el error implica la existencia de una desviación o diferencia entre la puntuación obtenida de la aplicación de un instrumento y una puntuación verdadera, la cual es difícil de precisar.

Nunally et. al. (1995) afirma que el error de medición se da mediante procesos sistemáticos y aleatorios. Los primeros afectan a todas las observaciones y pueden ser clasificados en errores constantes o sesgos. Los sesgos son propios de las evaluaciones psicológicas en la cual atributos irrelevantes (raza, género, edad, etc.) influyen la medición. Los segundos se traducen en errores aleatorios, los que son difíciles de detectar. Cabe consignar que la adivinación, la fatiga en pruebas largas, el registro erróneo de los datos son algunos de los factores que se traducen en errores aleatorios.

Medición de la Confiabilidad

Las medidas de la confiabilidad son de tres tipos: confiabilidad de coherencia interna, confiabilidad de formas alternas y confiabilidad prueba - re prueba. Los primeros dos establecen las variaciones entre los reactivos de una prueba mientras que el tercero refiere a variaciones que se dan

en aplicaciones sucesivas de un instrumento. Es importante considerar que algunos factores que se traducen en errores de medición son específicos para un tipo de medida de confiabilidad. De esta manera, las variaciones dentro de una prueba que afectan los índices de confiabilidad coherencia interna y formas alternas pueden explicarse por: (a) el muestreo de reactivos, (b) los efectos de la adivinación, (c) marcar una respuesta por error cuando se elige otra, (d) la lectura errónea debido a una redacción confusa, (e) los efectos de la fatiga en pruebas largas, (f) condiciones inapropiadas de la aplicación (ruido, iluminación, etc.) y (g) sólo para las formas alternas, las pruebas no son exactamente iguales. Por su parte, la variación entre prueba y re prueba puede deberse a un cambio real del sujeto entre pruebas y; en ciertas pruebas, a diferencias en los juicios de evaluadores. En el ámbito de la evaluación de desempeño, cabe apuntar que estas diferencias pueden indicar que los evaluadores tienen diferente acceso al desempeño del evaluado más que baja confiabilidad (Quijano, 1994).

La elección de un tipo de medida depende, según Guilford (1984), del propósito de la investigación así como del uso que se le quiere dar al coeficiente de confiabilidad. Los siguientes son los métodos que se usan habitualmente para estimar la confiabilidad de las medidas:

Método Prueba- Reprueba. Este método es más conocido como método test-rest. Es el método más sencillo para calcular la correlación entre los puntajes de aplicaciones sucesivas. Se obtendrán bajas correlaciones si el intervalo de tiempo entre una aplicación y otra es prolongado. La estabilidad temporal de las medidas depende también del objeto de evaluación. Nunally et. al. (1995) señala que las medidas con estabilidad temporal alta miden de rasgos (Vg. medidas de capacidades cognoscitivas como inteligencia) mientras que las medidas que tienen una estabilidad baja son medidas de estado (Vg. medidas de estado de ánimo). Al respecto, cabe considerar que el desempeño correspondería a una medida de estado más que de rasgo.

Método de Formas Paralelas y División por Mitad. El primer método implica la determinación de las correlaciones de los puntajes de formas equivalentes de una prueba. Es de especial utilidad para pruebas psicológicas y de rendimiento educacional. El segundo método, por su parte, establece la división

de una misma prueba en dos mitades, las que posteriormente serán correlacionadas. Generalmente, en una se dejan los ítems pares y, en la otra, los impares. El coeficiente de Spearman-Brown es el más utilizado al respecto. (Garrett, 1990)

Método de Consistencia Interna. Según Guilford (1984), este método es probablemente el que mejor mide la confiabilidad de una prueba. La consistencia interna es calculada principalmente mediante el coeficiente Alfa de Cronbach. En términos generales, el coeficiente alfa refleja el número de reactivos como la correlación promedio de estos. Un alfa muy bajo indica que la prueba es demasiado corta o que los reactivos no tienen nada en común. (Nunally et. al., 1995). Un alfa muy alto, por su parte, según Guilford (1985) sugiere que el instrumento mide un mismo atributo o habilidad, lo cual es propio de medidas homogéneas.

Medidas para Aumentar la Confiabilidad de las Medidas

Las siguientes son medidas que, Nunally et. al. (1995), propone para mejorar la confiabilidad de un instrumento, las cuales pueden ser aplicadas en conjunto o por separado:

1. Aumentar el número de reactivos. Para ello, los reactivos deben ser tan buenos como los del muestreo original. Cabe destacar además que las estimaciones de confiabilidad de las pruebas más largas son más precisas.

2. Escribir con claridad los reactivos.

3. Entregar instrucciones claras para responder la prueba así como respetar las condiciones prescritas para la administración de la misma.

4. Hacer explícitas las reglas de calificación de sujetos tanto como sea posible

5. Entrenar a los evaluadores para que hagan su trabajo

6. Guilford (1984) sugiere también evitar los artículos con un número demasiado pequeño de respuestas.

Normas de Confiabilidad

Los niveles de confiabilidad que se requieren varían según el propósito de la medición así como el grado de validez al que se aspira en función del estado de desarrollo del instrumento. Nunnally et. al. (1995) acota que: “En las primeras etapas de la investigación de la validación predictiva o de constructo, puede ahorrarse tiempo y energía mediante instrumentos que sólo tienen una confiabilidad modesta, por ejemplo, .70... Puede argumentarse que aumentar la confiabilidad mucho más allá de 0.80 en la investigación básica con frecuencia es un desperdicio de tiempo y dinero” (pp. 296). El mismo autor señala que: “Si se van a tomar decisiones con respecto a puntajes de prueba específicos, una confiabilidad de 0,90 es justo lo mínimo, y una confiabilidad de 0.95 se consideraría como una norma deseable” (1995, pp.296)

Interpretación y Uso del Coeficiente de Confiabilidad

El coeficiente de confiabilidad indica la repetibilidad de los resultados, es decir, señala si aplicaciones sucesivas del instrumento al mismo individuo arrojarán los mismos resultados siempre y cuando lo medido no cambie entre las aplicaciones (Nunnally et. al., 1995). El autor consigna también que: “...permiten generalizar de un uso particular del método a una amplia gama de circunstancias relacionadas” (1995, pp. 238)

El coeficiente de confiabilidad además es un índice de la efectividad del instrumento, ya que es necesaria al menos algún grado de confiabilidad para obtener medidas válidas (Garrett, 1999, Nunnally et. al., 1995). Al igual que en el caso de la validez, la confiabilidad no es una propiedad que depende sólo del instrumento, sino también de las condiciones de aplicación y de la población medida.

El coeficiente de confiabilidad no debe ser interpretado mecánicamente. Nunnally et. al.(1995), estipula que se debe tener cuidado con las medidas basadas en correlaciones entre reactivos que son calculadas con paquetes estadísticos como el SPSS. Al respecto, el autor consigna que una baja correlación observada entre reactivos puede reflejar diferencias estadísticas en la distribución relativa de un reactivo en relación con los demás, más que diferencias en su contenido.

Relación entre los Tipos de Validez y Confiabilidad

La relación entre los tres tipos de validez es muy estrecha. De hecho, se podría decir que son sólo facetas de un mismo fenómeno. Al respecto, Nunally et. al. (1995) establece que un instrumento diseñado para validar un constructo puede ser usado como criterio externo en otro estudio. El grado en que una prueba sirve para predicción aumenta la validez del constructo. Por otra parte, en cuanto a la validez de contenido, ésta también apoya la validez de constructo. Así, una mayor especificación del dominio de observables facilita la obtención de una muestra adecuada y representativa de reactivos.

La validez y confiabilidad también mantienen un vínculo muy estrecho dado que ambos, según Garrettt (1999): “refuerzan la eficacia del test” (pp.398). Al respecto, la confiabilidad es un requisito de validez y no viceversa. Finalmente, hay que destacar que hay factores que afectan tanto la validez como la confiabilidad de las medidas. Hernández et. al. (1998) enumera los siguientes: (a) improvisación en el desarrollo de instrumentos, (b) aplicación de instrumentos desarrollados en el extranjero sin mediar ningún procedimiento de estandarización para la población a la cual se pretende aplicar en el presente (c) aplicación de instrumentos que han sido estandarizados en un pasado remoto (e) condiciones de aplicación (ruido, mala iluminación, etc.), (f) errores en formato del instrumento (errores de enumeración de las páginas, letra ilegible, etc.), (h) longitud excesiva de la prueba y (i) inadecuación del instrumento en relación con las personas a las que se aplica. Respecto de este último factor, el autor consigna que: “Utilizar un lenguaje muy elevado para el encuestado, no tomar en cuenta las diferencias de sexo, edad, conocimientos, capacidad de respuesta, memoria, nivel ocupacional y educativo, motivación para responder y otras diferencias en los entrevistados, son errores que pueden afectar la validez y confiabilidad del instrumento de medición” (1998, pp. 240)

Discriminación

La discriminación o sensibilidad en la evaluación de desempeño apunta al grado en que un instrumento permite diferenciar entre sujetos que tienen un desempeño sobresaliente y los sujetos que presentan un desempeño inferior, lo que orienta la toma de decisiones administrativas o de desarrollo. Se

dice así que un ítem tiene poder discriminativo si distingue entre los sujetos que puntúan alto y los que puntúan bajo en el test.

La discriminación se define también como un procedimiento. Al respecto, Tristán establece que ha sido definido siempre en términos muy generales. Según el autor, se considera como: “Un procedimiento útil cuando el número de casos es 100 o más que requiere que dividamos al grupo en dos mitades” (www.ieesa-kalt.com/articulo1_ka.html)”

Quijano (1994), por su parte, va más allá. Establece que la discriminación de un instrumento de evaluación de desempeño se relaciona con tres factores distintos, a saber:

1. El grado de variación que ofrecen las escalas
2. El grado de utilización fáctica que tendrían cada nivel de la escala. Al respecto, acota que los niveles extremos de la escala, a veces, están formulados de tal forma que casi nunca son utilizados.
3. El grupo de referencia que propone el instrumento de evaluación. De modo que si el grupo de referencia es inadecuado, no se podría discriminar entre el bajo y alto desempeño de los empleados evaluados.

La medición de la discriminación se da mediante procedimientos como las correlaciones biserials y tetróricas. Se ha visto también que métodos más rudimentarios como la diferencia de cuartiles extremos son inadecuados, ya que no constituyen una prueba suficiente de discriminación (Guilford, 1984; Tristán, www.ieesa-kalt.com/articulo1_ka.htm)

Aceptabilidad

La aceptabilidad dice relación con el grado en que el sistema de evaluación recibe el apoyo de sus usuarios, tanto evaluadores como evaluados. Al respecto, se dice de “... un sistema que no es aceptado es peor que no realizar medición alguna” (Barnett, 1997, pp. 241), En la misma línea, Ramos et. al. (1996) plantea que: “Si un sistema de valoración no ha sido aceptado por los empleados, con toda probabilidad los datos obtenidos carecerán de valor” (pp.342)

Los sistemas de evaluación exhiben diferentes niveles de aceptabilidad en función de las características específicas de estos. Al respecto, por ejemplo, las evaluaciones de pares tienen baja aceptación (Cederblom et. al., 1980). Se ha encontrado además que la frecuencia de evaluación y el grado de estructuración del sistema aumenta la aceptabilidad del mismo (Orpen en Barnett, 1997).

Para asegurar que los sistemas de evaluación sean aceptados por los usuarios es vital la participación de los mismos en el diseño. Al respecto, se consigna que: “Un sistema que no es aceptado es casi peor que no realizar medición alguna. Esto apunta a la necesidad de involucrar tanto a evaluadores como evaluados en el desarrollo de instrumentos de evaluación de desempeño. La investigación ha establecido que esto incrementa la aceptabilidad del sistema” (Barnett, 1997, pp. 20)

La aceptabilidad también es una propiedad de los instrumentos de medición del desempeño (Barnett, 1997). Desde una perspectiva aplicada, ésta es más importante que las propiedades psicométricas (Quijano, 1994).

La aceptabilidad de un instrumento está determinada por: “... otras características particulares de la situación concreta en que se aplica el instrumento (Vg. historia, cultura de la organización, el método de implantación, etc.)” (pp.292) así como el sistema en el cual se enmarca. De este modo, por ejemplo, la aceptabilidad de un mismo instrumento puede variar en función de los objetivos que tenga el sistema de evaluación. Al respecto, el estudio de la aceptabilidad de un instrumento de evaluación por pares concluyó que éste era más aceptado si tenía por objetivo dar feedback en comparación con tomar decisiones administrativas (Mc Evoy et. al., 1987).

Algunas dimensiones de aceptabilidad de los instrumentos de evaluación de desempeño que han considerado los estudios son percepción de: validez, confiabilidad, libertad de sesgos, valor de feedback, practicalidad y utilidad. Cabe destacar que la aceptabilidad del instrumento no se refiere a las propiedades psicométricas y pragmáticas del mismo sino a la percepción que tienen los usuarios de éstas.

Se entiende por utilidad, el grado en que un sistema de evaluación sirve realmente a los propósitos para los que fue construido (Ramos et. al. 1996) La practicalidad refiere al grado en que el

instrumento es fácil de usar y es comprensible (Ramos et. al. 1996) En relación con la comprensión, otros autores han señalado la importancia de la transparencia, es decir, el grado en que parece clara para el evaluador, y especialmente para el evaluado, la relación entre los comportamientos del trabajador y las puntuaciones obtenidas (Quijano, 1994)

La medición de aceptabilidad se ha dado principalmente mediante la aplicación de cuestionarios a evaluadores y evaluados. Para detectar la aceptabilidad global de un instrumento, se le pide al usuario optar por una de las siguientes alternativas: mantener la aplicación del instrumento, revisar el mismo o suspender su aplicación. Del mismo modo, Mc Evoy et. al. (1987) postula que también se puede preguntar sobre el grado de satisfacción con el instrumento. En cuanto a la medición de las dimensiones, las preguntas cerradas siguen un mismo formato. Por ejemplo, en relación con la percepción de utilidad en relación con el objetivo “Dar feedback”, Cederblom et. al. (1980) preguntaron a profesores universitarios: ¿Qué tan útil fue el último ranking de desempeño dado por sus compañeros para dar feedback sobre sus fortalezas y debilidades?.

Otras propiedades no psicométricas que hay que considerar es la especificidad y operatividad. Barneett, 1997 acota que en cuanto a la primera, “Es el grado en que la medición de desempeño aporta una guía específica a los trabajadores acerca de lo que se espera de ellos y de cómo pueden cumplir con estas expectativas” (pp. 21). Respecto de la segunda el mismo autor señala: “Para que un sistema de evaluación sea viable debe ser fácil de implementar y al mismo tiempo eficiente en el uso de los recursos y del tiempo. El sistema de evaluación debe poseer un diseño tal que su correcta aplicación no exija más tiempo del que los supervisores puedan disponer, ni que exija de éstos el contar con habilidades comunicacionales excepcionales, cuya preexistencia es rara o cuyo desarrollo requiera una capacitación demasiado intensiva y prolongada” (1997, pp. 21).

En síntesis, todo instrumento de evaluación de desempeño debe satisfacer tanto requerimientos psicométricos de validez, confiabilidad y discriminación como requerimientos pragmáticos de aceptabilidad, practicalidad, utilidad, operatividad y especificidad.

Sesgos Propios de la Evaluación del Desempeño

Las distorsiones en la evaluación de desempeño pueden tener variadas fuentes, como ya hemos visto. Las fuentes de error más estudiadas son las que provee el propio instrumento y el evaluador. No obstante, variables organizacionales también pueden ser importantes.

Sesgos del Evaluador

El evaluador es una importante fuente de sesgos en las evaluaciones de desempeño. Esto ha sido el foco de numerosas investigaciones cuyo objetivo es aumentar la exactitud de las mismas y reducir el error de los ratings, ya que estos últimos constituyen problemas éticos, y en Estados Unidos, legales también.

Esto ha generado mejores procedimientos y técnicas, un aumento en la validez y confiabilidad de los instrumentos así como se ha logrado una mejor interpretación de los resultados de la evaluación mediante un mejor entendimiento de las dinámicas del proceso de evaluación (Kavanagh en London, Rowland y Zammuto, 1982)

Prejuicios personales

Esta distorsión consiste en que el evaluador formula juicios a priori basados en estereotipos en relación con la nacionalidad, género, religión, etc. (Werther, 1995). Por ejemplo, diversos estudios muestran que las personas atractivas son percibidas como más inteligentes y portadoras de características positivas (Dion, Bersheild y Walter, 1986 en Billipokf, 2002).

La edad también es una fuente de prejuicios, especialmente por la influencia de explicaciones teóricas que han perdurado en el tiempo¹⁵. La evaluación de desempeño así también puede ser una práctica discriminatoria. El metaanálisis realizado por Avolio y Waldaman (1986) encontró que, en las investigaciones analizadas, los indicadores objetivos de desempeño como los índices de productividad aumentaban con la edad, en cambio, los ratings de compañeros y supervisores mostraban una declinación progresiva con el incremento de la edad de los evaluados, debido a los sesgos que aportaron los evaluadores. Cabe acotar que existe poco estudio de las variables moderadoras de la relación edad-desempeño.

El género y la raza también serían importantes fuentes de prejuicios, pero la investigación al respecto no es concluyente, ya que es difícil determinar si las varianzas relacionadas se deben a sesgos de los evaluadores o a diferencia reales de rendimiento que se relacionan con estas variables u otras enmascaradas (Quijano, 1994).

Efecto de acontecimientos recientes

Los sistemas de evaluaciones que se basan en mayor medida en mediciones subjetivas son más influenciados por los acontecimientos recientes del empleado, ya sean positivos o negativos (Quijano, 1994; Werther, 1995). El método de incidentes críticos disminuiría esta tendencia. Cabe acotar que éste se puede emplear en combinación con otros métodos.

Tendencia a la medición central

La distorsión se produce porque los evaluadores evitan las puntuaciones muy altas o muy bajas, distorsionando éstas para que se acerquen a la media (Werther, 1995), ya sea para evitar suministrar información adicional que pueda requerir el Departamento de Recursos Humanos o dar explicaciones a los

¹⁵ Por ejemplo, el “contextualismo” sostiene que la habilidad y motivación para desempeñar la tarea disminuyen progresivamente por la influencia de factores contextuales como el aburrimiento del trabajo, experiencia significativa y necesidad de estimulación intelectual. Actualmente, los especialistas, por su parte, sostienen que los estudios presentan evidencia mixta. No obstante, las denuncias por discriminación a empleados de mayor edad se han ido incrementando progresivamente.

empleados o por desconocimiento de los evaluadores sobre el desempeño real de sus empleados, mal uso de la metodología de la evaluación o una baja habilidad para diferenciar entre el bajo y alto desempeño que es un factor clave para el éxito del manager (Quijano, 1994; Robbins, 2000). Al respecto, los managers que son "... bajos diferenciadores tienden a ignorar o suprimir las diferencias y perciben el mundo como algo más uniforme de lo que realmente es (Pizam en Robbins, 2000)".

Efecto halo o aureola

Es la tendencia del evaluador a calificar el desempeño global del empleado, basándose en una característica aislada del empleado, ya sea esta positiva o negativa. (Vg.. Responsabilidad, puntualidad, etc.) (Quijano, 1994; Robbins, 2000). Se produce con frecuencia, según Werther et al. (1995), cuando un supervisor que le corresponde evaluar es amigo de algunos de sus subordinados. Los sentimientos de simpatía o antipatía, en esta situación, tendrían más potencial para distorsionar la evaluación. Este efecto también puede darse en relación con el orden en que el evaluador califica a sus subordinados de modo que la valoración del desempeño de un empleado puede influenciar la evaluación de otro empleado que es evaluado a continuación o previamente. Así, por ejemplo, es probable que si un empleado presenta un muy bajo desempeño y el empleado evaluado a continuación presenta un desempeño mediocre, este último finalmente sea evaluado como un empleado con desempeño sobresaliente. Esta distorsión es conocida como halo sucesivo (Quijano, 1994)

Error de indulgencia

Esta distorsión consiste en que el evaluador califica con puntajes muy altos (error de indulgencia positivo) o muy bajos al grupo de los sujetos evaluados (error de tipo negativo). Esto se produce porque el estándar del evaluador no se equipara con el desempeño real de sus subordinados (Quijano, 1994; Robbins, 1999). Los mismos resultados distorsionados también pueden producirse por las expectativas, creencias o necesidades del evaluador. Por ejemplo, el supervisor que desea ser popular entre sus subordinados puede otorgar sistemáticamente puntuaciones muy altas (Werther, 1995).

Forzar la información para adaptarse a los criterios no relacionados con el desempeño

Una práctica frecuente en las organizaciones es realizar una evaluación informal previa que después se intentará equiparar con los criterios de la evaluación formal. Así, basado en la apreciación de otros criterios, como la antigüedad, se intenta que el individuo obtenga una calificación equivalente en el sistema formal (Robbins, 1999).

Otras tendencias que hay que considerar que pueden distorsionar el proceso de evaluación de desempeño son:

1. Patrones de atribución causal de empleados y superiores que tienen un fuerte impacto en la aplicación de la evaluación como en el feedback que se entrega en la entrevista de evaluación. Por ejemplo, los supervisores tienen el patrón de atribución fundamental que consiste en atribuir el bajo desempeño de sus subordinados a sus características personales, en vez de conductas específicas o factores externos (King, 1984; Chiavenato, 2000). Además tanto supervisores como empleados, tendemos a la autojustificación, o sea, nos atribuimos mayor responsabilidad en el éxito que en el fracaso.

2. El insuficiente conocimiento de los evaluados es otra importante fuente de distorsión que es difícil de detectar, ya que, especialmente, los supervisores no reconocen el desconocimiento del personal a su cargo. (Quijano, 1994)

3. La tendencia de los evaluadores a calificar con más altos puntajes a los empleados con los que tendrían características similares. Este es el llamado error de similitud (Robbins, 1999).

4. La personalidad del evaluador también afectaría la valoración del desempeño. Schultz (en Quijano, 1994) expone que: (a) evaluadores con alta autoestima tienden a dar evaluaciones más objetivas mientras que los con baja autoestima tienden a caer en el error de indulgencia negativo y (b) las evaluaciones de evaluadores con altos niveles de ansiedad presentan puntuaciones más extremas.

Sesgos del instrumento

El principal sesgo de los instrumentos de evaluación que ha sido documentado es el criterio único. El desempeño global de un empleado se relaciona con el cumplimiento de varios criterios de desempeño. Si se mide sólo uno criterio, no sólo se genera una evaluación sesgada sino que el desempeño de los empleados puede focalizarse exclusivamente en el criterio seleccionado (Robbins, 2000).

Sesgos de la organización

La investigación sobre los sesgos del evaluador ha sido extensa, sin embargo, se ha ignorado a la organización como una fuente importante de distorsión que puede afectar la interpretabilidad de la información entregada por la evaluación de desempeño.

La variable organizacional afectaría específicamente la habilidad de los managers de interpretar con exactitud y la comparación de los ratings de desempeño entre dos o más unidades de la organización (London et al., 1982). Así las diferencias contextuales se relacionarían con que diferentes evaluadores podrían poner distintos énfasis en un criterio específico de desempeño de un mismo instrumento. Esto apunta a cómo se define cada criterio y cuánto pesa en las diferentes unidades, departamentos o áreas de la organización. Entonces, en un mismo instrumento, se emplean diferentes modelos de desempeño, lo cual tiene implicancias éticas, legales, administrativas.

La tarea y las variables organizacionales que también pueden afectar la validez y confiabilidad de las evaluaciones, según Kane y Lowler (1979 en London et al., 1982), son: (a) el grado especificación e incertidumbre relacionada con la posición del cargo, (b) el nivel de autonomía del evaluado en el desempeño de tareas, y (c) la frecuencia y cantidad de retroalimentación del desempeño que reciben los evaluados, (d) características sociales como el clima organizacional¹⁶. La investigación sobre estas variables es insuficiente.

¹⁶ Un estudio previo efectuado por Lawler sugirió que climas organizacional deficientes se caracterizaban por bajos niveles de apertura y confianza, lo que distorsionaba los resultados de la rating de desempeño.

Estrategias para Reducir las Distorsiones

En relación con la selección y diseño adecuado de los instrumentos de evaluación, existen tres medidas para reducir distorsiones que deben considerarse:

1. Usar criterios múltiples para determinar el desempeño global (Robbins, 1999).
2. Evaluar comportamientos por rasgos, ya que estos no necesariamente se relacionan con el desempeño en un puesto determinado. Es importante mencionar que la definición de rasgos varía de evaluador en evaluador (Robbins, 1999)
3. Complementar el método de evaluación seleccionado con el registro de incidentes críticos, lo que reduce el efecto de indulgencia, halo y de acontecimientos recientes (Robbins, 1999).
4. Complementar un método de evaluación del desempeño pasado con uno que evalúe el desempeño potencial (Werther, 1995).

En relación con la selección y desempeño de los evaluadores:

1. Entrenar a los evaluadores potenciales. Se ha visto que talleres de uno a tres días e incluso una sesión de explicación de cinco minutos sobre las fuentes de distorsiones más comunes que aporta el evaluador, contribuyen a reducir las distorsiones (Robbins, 1999). Es importante considerar que se requieren sesiones de actualización, ya que los efectos de las capacitaciones tienden a disminuir en el tiempo. Werther et al. (1995) señala que un programa de capacitación debería incluir una explicación sobre la importancia de la evaluación de desempeño y las fuentes de distorsión más significativas, el cómo aplicar el método seleccionado, y ejercicios de simulación.

2. Usar múltiples evaluadores que evalúen sólo los factores en los cuales son competentes. A mayor número de evaluadores, los resultados de la evaluación representarían una curva normal, por ende, la precisión aumenta. Además, es necesario que los evaluadores sólo califiquen las dimensiones en las cuales son competentes (Robbins, 1999). Por ejemplo, aspectos del trabajo en equipo podrían ser evaluados por pares.

Diseño e Implementación de Sistemas de Evaluación

En la investigación-intervención en las organizaciones es posible identificar una serie de fases en el diseño de servicios de gestión de recursos humanos como los sistemas de evaluación de desempeño, selección de personal, de desarrollo de carrera y de distribución de recompensas. Estas fases están incorporadas en el ciclo de diseño, concepto acuñado por la ingeniería y que es aplicado en el diseño de diversos productos (Roe, 1989 en Quijano, 1994). La aplicación del ciclo de diseño permite comprender el cómo la intervención profesional se logra ajustar a las necesidades de la organización y se constituye en más que una mera aplicación de técnicas y herramientas.

Las fases del ciclo de diseño señaladas por Eekels (1983, en Quijano, 1994) son:

Definición del Servicio a Diseñar y Determinación de las Funciones

La primera fase implica el detectar las necesidades de la empresa frente al servicio. De hecho, esta fase se denomina habitualmente “Diagnóstico de necesidades”. En éste, una de las fuentes primordiales de información son los futuros usuarios (Quijano, 1994).

En cuanto a los aspectos o necesidades a considerar en el diseño de un sistema de evaluación, Quijano (1994) identifica:

1. Objetivos de evaluación o fin para el que se va evaluar
2. Objeto de la evaluación. Esto alude a lo qué se va a evaluar, lo que implica determinar si es necesario evaluar resultados o desempeño, ya sea mediante rasgos o conductas así como usar un criterio simple, múltiple o compuesto. Cabe destacar que en cuanto al objeto de evaluación, el análisis de trabajo aporta información valiosa. Al respecto, Wherther et. al (1995) señala que: “Basándose en las responsabilidades y labores en la descripción del puesto, el analista puede decidir qué elementos son esenciales y deben ser evaluados en todos los casos.” (pp. 323).
3. Actores de la evaluación (evaluadores).

4. Instrumentos de evaluación a utilizar, los cuales finalmente son elegidos sobre la base de criterios de utilización y psicométricos.

Análisis

Se derivan del paso anterior, los requisitos y procedimientos asociados al servicio. Se debe tomar en cuenta, las restricciones de recursos, tiempos, legales, las cuales pueden ser señaladas por los mismos usuarios y el cliente o definidas por el investigador. Es importante considerar para la implementación de los servicios de recursos humanos, la viabilidad y grado de resistencia que se pueden enfrentar.

En cuanto a los requisitos identificados, los instrumentos de evaluación, según Mc Lagan en Barnett (1997) deben cumplir con las propiedades psicométricas y de utilización ya descritas, pero también los datos que de estos se obtienen deben ser fácilmente relacionados con las preguntas y necesidades de decisión de los stakeholders (accionistas, empleados, directivos, la comunidad y el gobierno) así como realizar un pilotaje del sistema. Barnett (1997) también apunta que se deben satisfacer las siguientes condiciones:

1. Transparencia de proceso para todos los involucrados, es decir, que los procedimientos implicados en la evaluación sean conocidos y claros para evaluadores y evaluados

2. Documentación reducida y autorreferente, lo cual sirve para optimizar los recursos de tiempo y para que todos los involucrados tengan conocimiento de la naturaleza y alcance del sistema. El documento, conocido como manual del sistema, debe explicitar las normas y procedimientos de evaluación sin ser muy extenso.

3. Perfeccionamiento continuo del sistema en forma participativa.

4. La persona que es evaluada debe participar del proceso de evaluación

Síntesis y Simulación

Se elabora finalmente una propuesta de servicio que implica el uso de conocimiento y técnicas científicas. A continuación, se prueba el sistema diseñado mediante modelos de simulación o contrastación empírica. Es decir, se determinan las propiedades del servicio para ver si éste cumple con los requisitos establecidos.

Evaluación

Se procede a continuación a evaluar el valor que el servicio tiene para los usuarios. La evaluación debe permitir el grado de satisfacción del sistema para el cliente. Es importante notar que la evaluación de los usuarios es diferente a la perspectiva de un profesional. Así, por ejemplo, un profesional que diseñe un sistema puede valorarlo en función de la calidad, confiabilidad, validez mientras que el cliente se preocupa de la utilidad, rentabilidad, rapidez, etc.

Toma de Decisiones

Se decide, por último, si se acepta o no el sistema. Si no es aceptado, se debe volver a la fase 2 o 3 hasta que el sistema sea percibido como satisfactorio.

Evaluación del Desempeño y Política

En relación con el desarrollo de sistemas de evaluación, también debe considerarse como un juego político que da cuenta de las relaciones de poder que se dan entre los actores organizacionales. El psicólogo jugará un rol clave en la validación de éstas. De esta perspectiva, la evaluación de desempeño es un dispositivo de poder en el cual la fijación de criterios de evaluación posibilita el control de los comportamientos de los empleados que, en la mayoría de las organizaciones, son evaluados por miembros que ocupan posiciones de mayor nivel en la estructura jerárquica. Desde esta perspectiva política, la evaluación es una herramienta para distribuir recompensas y castigos. (Gómez - Mejía, Balkin, Crady,

20001). En la misma línea, Paéz (2002) apunta que una de las causas de la falibilidad de la evaluación del desempeño es la existencia, dentro de cada organización, de una estructura política en la cual el proceso de manejo de los acuerdos y desacuerdos entre evaluadores y evaluados y la administración de juicios son manifestaciones de poder. Es así como la evaluación de desempeño, en cuanto al qué se evalúa, cómo, quién, cuándo; nos dice mucho sobre la organización en sí y el proyecto social de ser humano al cual se circunscribe (Pujol, 2000). El psicólogo que diseña, implementa y evalúa sistemas de evaluación de desempeño valida así las relaciones de poder que se dan en cada organización. Cabe señalar que éste deja de ser objetivo para intentar ser imparcial mediante reconocer y equilibrar los intereses propios y de terceros, siendo estos evaluadores y evaluados.

Desde esta perspectiva política, un segundo punto de análisis refiere a los objetivos y roles que el evaluador y evaluado ponen en juego durante el proceso, tanto en la observación, calificación como en el proceso de toma de decisiones. Al respecto, Gómez- Mejía, Balkin y Cardy (2001) señalan que mientras desde una perspectiva racional el objetivo de la evaluación es la precisión; en una perspectiva política la utilidad radica en la agenda y el contexto de la evaluación. De este modo, un evaluador puede calificar con altas puntuaciones a su personal para minimizar conflictos laborales. En cuanto a los roles del evaluador y evaluado, estos distan de ser agentes pasivos en el proceso de evaluación. Ambos entran en un proceso de negociación en la cual, por ejemplo, el evaluado intentan influir sobre el juicio del evaluador mediante un conjunto diverso de comportamientos como halagos, excusas y arrepentimientos. (Gómez- Mejías, et. al., 2001).

Evaluación y Gestión del Desempeño

Como ya se expuso, el acto de evaluar el desempeño¹⁷ es sólo una pieza del proceso de evaluación que tiene por objetivo último mejorar el desempeño del recurso humano. Precisamente, el proceso de valoración agrupa una serie de prácticas regulares que van desde la entrega de feedback hasta la

¹⁷ Evaluar en el sentido de emitir juicios sobre el desempeño que presenta un empleado.

determinación e implementación de acciones de mejora y en las cuales tanto el empleado como el evaluador tienen un rol que jugar.

Es importante destacar que la evaluación del desempeño debe constituirse como una herramienta integrada que concuerda con las políticas y procedimientos de gestión del desempeño entre los cuales destacan el análisis de puesto y la retroalimentación (Barnnett, 1997). En relación con este último, se considera que es parte integral del proceso de evaluación más que una herramienta diferente a la evaluación.

Retroalimentación del Desempeño

Es uno de los puntos más fundamentales del proceso de evaluación porque posibilita el equilibrio de las necesidades organizacionales e individuales (Clement y Malinauskas, 1987). Los empleados necesitan recibir retroalimentación de su comportamiento, de los resultados y objetivos logrados. Si estos no obtienen información significativa y específica de su desempeño es difícil que se logren los objetivos organizacionales (Chiavenato, 2000).

Robbins afirma que: “Una revisión eficaz - una en la cual el empleado perciba que la evaluación es justa, el gerente es sincero y el clima es constructivo- puede provocar que el empleado deje la entrevista con humor más ligero, informado acerca de las áreas de desempeño en las cuales necesita mejorar y salga con la determinación de corregir las deficiencias” (1999, pp. 447). No obstante, los efectos de las entrevistas de revisión sobre la motivación, las actitudes hacia el trabajo y el desempeño posterior no han sido un tema muy explorado (Mohrman, Resnick-West & Lawler, 1989 en Milliman, Mohrman, y Nathan, 1991). Existe consenso sobre la resistencia de los supervisores y defensividad de los empleados ante las mismas (Billikopf, 2002; Werther, 1995). Ésta se explicaría en función de las expectativas del supervisor de tener que dar o recibir retroalimentación negativa y de tener que manejar la diferencia frecuente entre las evaluaciones - formales e informales- que el empleado y el mismo han efectuado. Por todo esto, los controles organizacionales son importantes para asegurar su adecuada realización.

La entrevista de evaluación de desempeño es un proceso de interacción cara a cara que se inicia con la preparación de la misma y termina con la fijación de objetivos y metas para el próximo período de evaluación. No obstante, se ha observado, en sujetos voluntarios (docentes universitarios), que la sola retroalimentación en la cual un asesor describe los comportamientos observados en el ambiente de trabajo del evaluado sin emitir juicios de valor, dar sugerencias y planificar acciones puede generar mejoras significativas en el desempeño (Polanco, 1994).

La planificación de la entrevista incrementa la objetividad del proceso y reduce el estrés de los evaluadores, ya que favorece la entrega de información sobre la evaluación del trabajo específico del empleado y no del empleado mismo. Según King: “Los empleados responden - conscientes o inconscientemente- al grado de esfuerzo que el supervisor invierte en la evaluación.” (1984, pp. 66). Preparar la entrevista implica desde: (a) determinar el espacio y tiempo de la entrevista; (b) compartir, con el empleado evaluado, las expectativas, filosofía de evaluación y registros de evaluación, (c) establecer, en conjunto, una agenda sobre los objetivos y temas a tratar que sea compartida y realista entre el empleado y el jefe; a (d) planificar qué se va a decir y cómo lo va a decir (Chiavenato, 2000; King, 1984).

La entrevista de evaluación misma tiene objetivos múltiples¹⁸. Cabe destacar que como proceso de interacción cara a cara no sólo tiene un componente racional, sino también emocional. Podemos así encontrar diferencias perceptuales, mecanismos de defensa en acción, conflicto, malos entendidos e incluso hostilidad (Clement et. al., 1987). Por ende, la entrevista como una herramienta comunicacional demanda una serie de competencias comunicacionales claves en el supervisor que si son desarrolladas sirven para manejar los problemas potenciales antes referidos. Durante esta reunión, una de las partes más importantes es la fijación de objetivos y metas para el próximo período con fechas de cumplimiento, así

¹⁸ Según Chiavenato (2000) los objetivos de una entrevista de evaluación son: mejorar el desempeño mediante una comunicación clara e inequívoca de los estándares de desempeño (calidad, cantidad, métodos de trabajo y de la importancia de estos), entregar al subordinado retroalimentación clara de su desempeño (fortalezas y debilidades en comparación con el logro de estándares de desempeño), discutir y desarrollar los planes de acción para mejorar el desempeño, estimular las relaciones interpersonales entre el supervisor y subordinado, así como eliminar o reducir discrepancias, ansiedades, tensiones e incertidumbres de los individuos que no gozan de asesoría planificada y bien orientada.

como los mecanismos para hacer un seguimiento del logro de estos. Finalmente, cabe destacar que el clima organizacional influye en la efectividad de la entrevista de evaluación de desempeño. Clement et. al. (1987) destaca que: “Cuando el clima es de alta confianza, apoyo y apertura, evaluadores y subordinados ven que la entrevista mejora. Así reportan un énfasis en el desarrollo del subordinado y en la participación constructiva” (pp. 79)

Análisis y Planeación del Desempeño

El análisis de las causas del rendimiento es una tarea compleja dados los múltiples determinantes del mismo. Sin embargo, no puede obviarse como parte integrante del proceso de evaluación porque el juicio que hace el evaluador sobre las causas del desempeño influye en la evaluación misma y porque las acciones de mejora deben seleccionarse en función de los determinantes de desempeño implicados. De hecho, el análisis y planeación se da a lo largo de todo el proceso de valoración.

El análisis del desempeño se ha orientado principalmente a mejorar el desempeño insuficiente, basándose en los determinantes, motivación, habilidad y oportunidad. El uso de árboles de decisión o la elaboración de lista de determinantes puede facilitar la realización de este análisis. Así, por ejemplo, Gómez et. al. (2001) recomienda al supervisor elaborar una lista con los factores externos más relevantes que son propios de la organización que influyen en el bajo desempeño. La lista posteriormente debe ser revisada por el subordinado. Autores como Maper y Pipe (1984), ofrecen a los supervisores un árbol de decisión como guía. En general, en torno al desempeño insatisfactorio, algunas de las medidas sugeridas son: entrenamiento, feedback frecuente, redefinición de objetivos, instrucciones y advertencias verbales, y/o reasignación o desvinculación cuando las anteriores son insuficientes (Alles, 2000, Gómez - Mejía, Balkin, Crady, 20001). Al respecto, Alles (2000) consigna que los planes de mejora del rendimiento se deben limitar a los casos especiales, o sea, cuando un empleado que presenta bajo desempeño históricamente ha obtenido evaluaciones positivas o cuando el desempeño del mismo se ve afectado por factores externos de causa mayor (Vg. crisis familiar).

Figura N° 5: Determinación de Medidas para Problemas de Desempeño

Causa	Preguntas que hay que Hacer	Posibles Medidas
Capacidad	¿El trabajador ha sido capaz alguna vez de tener buen rendimiento?	Formación Transferencia
	¿Pueden otros trabajadores hacer el trabajo adecuadamente, pero éste no?	Volver a diseñar el trabajo
Esfuerzo	¿Está disminuyendo el rendimiento del trabajador?	Despedir al trabajador
	¿Es el rendimiento menor en todas las áreas?	Reconocer el buen funcionamiento
Situación	¿El rendimiento es errático?	Clarificar las necesidades de los proveedores
	¿Los problemas de rendimiento afectan a todos los trabajadores, incluso aquellos que disponen de los suministros y equipos adecuados?	Cambiar de proveedores Suprimir señales o demandas en conflicto Proporcionar las herramientas adecuadas

Fuente: Gómez -Mejía et. al. (2001, pp. 226)

Sin importar el método de análisis, debe haber acuerdo entre el directivo o supervisor y el evaluado sobre los determinantes involucrados y medidas a tomar. En este punto, la autoevaluación del desempeño cobra especial importancia, ya que favorece no sólo la entrega de feedback sino la planificación de acciones. En relación con la autoevaluación, por ejemplo, Billikopf (2002) sugiere que el empleado conteste una serie de preguntas, entre las cuales se incluye: ¿Qué puedo hacer yo de forma diferente, como su supervisor, para que usted me ayude a cumplir mejor con mis responsabilidades?. En este caso, el empleado sugiere acciones para mejorar el desempeño del jefe. En cualquier caso, se debe tratar de maximizar los factores que afectan positivamente el desempeño y eliminar o minimizar los factores que lo afectan negativamente. Para ello, según Alles (2000), es clave definir un plazo de mejora, por ejemplo, seis meses.

MÉTODO

Tipo de Estudio

La presente investigación corresponde a un estudio exploratorio- descriptivo. El carácter exploratorio está dado por la escasez de investigaciones en el área de la evaluación de desempeño en nuestro país⁴. Por su parte, el carácter de descriptivo se relaciona con el propósito de describir un fenómeno con la mayor exactitud posible (Hernández et. al. 1998), a saber, la evaluación del desempeño de técnicos y asistentes técnicos que laboran en una empresa tabacalera.

Diseño de Investigación

El diseño de la investigación es no experimental - transversal. El diseño no experimental es una indagación empírica y sistemática que se caracteriza por no manipular activa o deliberadamente una situación o sus condiciones y no asignar sujetos o tratamientos al azar (Kerlinger, 1998). Es decir, según el autor, intenta observar los fenómenos en su contexto más natural. Los diseños transeccionales o transversales, por su parte, recogen información y datos en un momento en particular (Hernández et al., 1998).

Universo y Muestra

En la presente investigación se emplean dos muestras y universos de sujetos. En la fase de diagnóstico, la primera muestra delimita el personal a entrevistar, y en la fase de simulación, la segunda muestra, los evaluadores y sujetos a evaluar por medio del instrumento diseñado. Esta última muestra corresponde a una unidad pareada de análisis.

⁴ En relación con la investigación en nuestro país, sólo se pudo acceder a cuatro tesis de pregrado. Tres de éstas centradas en el diseño y validación de sistemas de evaluación, una cuarta orientada al análisis de un sistema de evaluación previamente establecido. Ninguna de la tesis referidas considera como sujetos evaluados a personal de línea, técnicos u operarios de empresas productivas.

Universo y Muestra de la Fase de Diagnóstico de Necesidades

La población está dada por todos los empleados del Área de Operaciones y del Área de Recursos Humanos de la Fábrica de Casablanca de Chiletabacos, lo cual equivale a un total de 224 sujetos.

La muestra de tipo no probabilística por cuotas²⁰ está compuesta por 63 sujetos. Las cuotas que fueron fijadas son: (a) 25% de técnicos mecánicos- electrónicos de línea, (b) 25% de asistentes técnicos de línea, (c) 100% de jefes de producción y (d) 100% del Personal de la Gerencia de Recursos Humanos y Operaciones.

Tabla N° 1: Composición de la Muestra de Personal Entrevistado

Categorías	f ₍₁₎	f ₍₂₎	fr ₍₃₎
I. Departamento de Operaciones			
A. Técnicos mecánicos-electrónicos	22	85	25%
B. Asistentes técnicos	33	131	25%
C. Jefes de Producción	5	5	100%
D. Gerente de Operaciones	1	1	100%
Total	61	222	27%
II. Departamento de Recursos Humanos			
A. Gerente de Recursos Humanos	1	1	100%
B. Analista de Recursos Humanos	1	1	100%
Total	2	2	100%

(1) Frecuencia de empleados entrevistados

(2) Frecuencia de empleados por cargo

(3) Frecuencia relativa de empleados entrevistados por cargo

Universo y Muestra de la Fase de Simulación

Sujetos Evaluadores

La población de sujetos evaluadores está dada por todos los jefes de producción que se desempeñen en el Área de Operaciones de la Fábrica de Casablanca de Chiletabacos, lo que equivale a un total de 5 sujetos.

²⁰ En esta muestra, las cuotas se fijan en función de la proporción de ciertas variables sociodemográficas de la población (Hernández, 1998). Cabe acotar que se asignan diferentes porcentajes para las cuotas en función del número total de sujetos por cargo así como se mantiene la proporción de técnicos electrónicos, técnicos mecánicos y asistentes técnicos de cada célula

Dado el reducido número de jefes de producción, la totalidad de estos originalmente evaluarían el desempeño de sus subordinados. No obstante, una serie de condiciones organizacionales como la reducción de la dotación de jefes de producción y la reasignación de 3 de los mismos a otras células, disminuyó el número de sujetos habilitados para evaluar el desempeño en función del conocimiento del cargo y de los sujetos a evaluar. El desconocimiento del puesto de trabajo así como de sus ocupantes aumenta la posibilidad de otorgar calificaciones sesgadas que son difícilmente detectables (Robbins, 2000, Quijano 1994). Dado lo anterior, con la determinación de los criterios de selección y exclusión de la muestra de sujetos evaluadores, sólo en una célula se pudo aplicar el instrumento de evaluación. En relación con el jefe de producción que evaluó a sus subordinados, cabe destacar que ha ejercido por más de dos años como jefe de producción en la Célula Printing, desempeñándose con anterioridad como técnico mecánico en la misma célula. Esto último garantizaría el conocimiento del personal a evaluar, de las tareas a realizar así como del proceso productivo mismo. Finalmente, cabe señalar que la muestra dirigida tiende a un mayor sesgo y menor representatividad respecto de las muestras probabilísticas, sin embargo, es especialmente útil para diseños de investigación en que se debe asegurar que los sujetos posean ciertos atributos o características; en este caso, conocimiento del trabajo y de los evaluados²¹ (Hernández et. al., 1988).

Criterios de Selección de la Muestra de Sujetos Evaluadores

Como criterios de exclusión de la muestra se considera: (a) ejercer el cargo de jefe de producción del Área de Operaciones por 6 meses o más (b) desempeñarse en la misma unidad de trabajo por 6 o más meses y (c) presentar actividad laboral efectiva como jefe de producción durante la duración de este estudio.

²¹ El conocimiento del trabajo y evaluados se relaciona con los criterios a y b de selección o exclusión de la muestra de sujetos evaluadores.

Criterios de Exclusión de la Muestra de Sujetos Evaluadores

Como criterios de exclusión de la muestra se considera: (a) ejercer jefe de producción del Área de Operaciones por menos de 6 meses, (b) desempeñarse en la misma unidad de trabajo por menos de 6 meses y (c) no presentar actividad laboral efectiva como jefe de producción durante la duración de este estudio.

Sujetos a Evaluar

La población se constituye por todos los técnicos y asistentes técnicos del Área de Operaciones de la Fábrica de Casablanca de la Compañía Chilena de Tabaco, lo que equivale a un total de 216 sujetos ²².

La muestra dirigida se define por todos los técnicos mecánicos, técnicos electrónicos y asistentes técnicos de línea de la Célula Printing, lo que equivale a un total de 30 sujetos²³.

Criterios de Selección de la Muestra de Sujetos Evaluados

Como criterios de selección de la muestra se considera: (a) ejercer como técnico electrónico, técnico mecánico o asistente técnico del Área de Operaciones por 6 meses o más (b) desempeñarse en el actual puesto de trabajo por menos 6 meses, (c) desempeñarse en la unidad de trabajo (Célula Printing) por 6 meses o más, y (d) presentar actividad laboral efectiva como técnico o asistente técnico durante la duración de este estudio.

Criterios de Exclusión de la Muestra de Sujetos Evaluados

Como criterios de exclusión de la muestra se considera: (a) ejercer como técnico electrónico, técnico mecánico o asistente técnico fuera del Área de Operaciones, (b) desempeñarse en el actual puesto de trabajo por menos 6 meses, (c) ejercer en la unidad de trabajo (Célula Printing) por menos de 6, y (d) no presentar actividad laboral efectiva como técnico o asistente técnico durante la duración de este estudio

²² Este número es estimativo y corresponde al período de trabajo de campo.

²³ Todos los técnicos electrónicos - mecánicos y asistentes técnicos de CCT son varones.

Tabla N° 2: Composición de la Muestra de Personal Evaluado

Características	f
I. Cargo	
A. Técnico Electrónico	2
B. Técnico Mecánico	4
C. Asistente Técnico	25
II. Edad	
A. Menor de 20 años	0
B. 21 a 30 años	3
C. 31 a 40 años	9
D. 41 a 50 años	15
E. 50 a 60 años	4
F. Mayor de 60 años	0
II. Estado Civil	
A. Casado	23
B. Soltero	8
III. Educación	
A. Educ. Bás. Inc.	0
B. Educ. Bás. Com.	0
C. Educ. Med. Inc.	0
D. Educ. Med. Com.	0
E. Educ. Técnica Inc.	7
F. Educ. Técnica Com.	17
G. Educ. Sup. Inc.	1
H. Educ. Sup. Com.	6
IV. Antigüedad en la Compañía	
A. Menos de 1 año	2
B. 1 a 10 años	9
C. 11 a 20 años	12
D. 21 a 30 años	4

Definición de Variables

Desempeño

Definición Conceptual

Conjunto de conductas, presentadas por un empleado durante un periodo de tiempo establecido, que contribuyen al logro de objetivos de la organización así como al buen funcionamiento de la misma.

(Ramos et. al., 1996)

Definición Operacional

Puntaje bruto calculado mediante la sumatoria de las puntuaciones obtenidas en cada una de las dimensiones de la variable desempeño.

Dimensiones

1. Comunicación Oral.
2. Comunicación Escrita.
3. Escucha Activa.
4. Relaciones Interpersonales.
5. Orientación a Resultados.
6. Resolución de Problemas
7. Innovación.
8. Cooperación.
9. Desarrollo de Pares.
10. Aprendizaje.
11. Prevención de Riesgos.
12. Manejo de Situaciones Críticas.
13. Mantención Correctiva.
14. Mantención Preventiva.
15. Operación de Máquina.
16. Aseguramiento de la Calidad

Cabe señalar que el puntaje por dimensión se calcula mediante la sumatoria de las respuestas señaladas por el evaluador a los ítems correspondientes a una dimensión según la variable escalar: nunca, casi nunca, regularmente, casi siempre y siempre.

La asignación de valores a los grados de respuesta es la siguiente:

1. Nunca: 1 punto.
2. Casi nunca: 2 puntos.
3. Regularmente: 3 puntos.
4. Casi siempre: 4 puntos.
5. Siempre: 5 Puntos.

Indicadores por Dimensión²⁴

1. Comunicación oral: (a) Dice lo que quiere decir en forma clara y directa, (b) Usa adecuada y éticamente la información que dispone y (c) Se asegura que su interlocutor comprendió lo expresado

2. Comunicación escrita: (a) Escribe en forma clara, concisa y directa, y (b) Se asegura que el documento o los datos sean recibidos oportunamente y comprendidos cómo corresponde.

3. Escucha Activa: (a) Presta atención a quien habla y (b) Se asegura que comprende lo que se le dice.

4. Relaciones Interpersonales: (a) Entiende los motivos por los cuales una persona hace, siente o dice tal o cual cosa, (b) Rechaza amablemente las peticiones de los demás si son poco razonables o si se oponen a los intereses de la empresa o los propios, (c) Pide colaboración o recursos adecuadamente, (d) Hace cumplidos en el momento mismo, (e) Pide disculpas cuando ha cometido un error o ha pasado a llevar a alguien, y (f) Hace críticas constructivas, concisas, en el momento mismo

5. Orientación a resultados: Se preocupa por el logro de los objetivos de su unidad y la empresa

6. Resolución de Problemas: (a) Reconoce rápidamente los problemas que debe solucionar por sí mismo o en conjunto con la unidad, (b) Determina correctamente la(s) causa(s) del problema por sí mismo o en conjunto con el equipo y (c) Hace rápidamente lo que hay que hacer para solucionar problemas o mejorar los resultados sin una orden directa o sugerencia.

²⁴ Estos indicadores son los resultantes de la validación de contenido efectuada por los jueces expertos.

7. Innovación: (a) Identifica oportunidades de mejoras en el proceso productivo y (b) Desarrolla y propone nuevos métodos y procedimientos.

8. Cooperación: Cooperación activa, eficazmente y por iniciativa propia con tareas que permitan el logro de objetivos de su unidad de trabajo, así como de otras áreas o departamentos de la empresa

9. Desarrollo de pares: (a) Se interesa por el desarrollo de habilidades y la adquisición de conocimientos en sus compañeros, (b) Enseña de una manera clara y dinámica y (c) Evalúa el proceso de instrucción para mejorar

10. Aprendizaje: (a) Se preocupa por su desarrollo profesional, (b) Genera o participa activamente en actividades de aprendizaje, (c) Aplica continuamente lo aprendido a su trabajo

11. Prevención de riesgos: (a) Toma precauciones adicionales cuando realiza trabajos potencialmente peligrosos, (b) Aplica rigurosa y oportunamente todos los procedimientos de seguridad para la operación y mantenimiento de máquina, así como del manejo y traslado de materiales y sustancias

12. Manejo de Situaciones Críticas: (a) Detecta rápidamente las situaciones de emergencia y determina los procedimientos a seguir y (b) Responde oportuna y adecuadamente en situaciones de emergencia.

13. Mantenimiento Correctivo: (a) Diagnostica eficazmente la falla(s), y (b) Repara rápidamente la falla(s).

14. Mantenimiento Preventivo: Aplica oportuna y rigurosamente las pautas de mantenimiento programadas.

15. Operación de Máquina: (a) Comprueba la disponibilidad y especificaciones técnicas de materias primas e insumos necesarios para la producción, (b) Manipula adecuadamente la máquina y sus sistemas anexos y (c) Ejecuta adecuadamente el programa de producción según el tipo de producto, cantidad, tiempo de fabricación, siguiendo las especificaciones del manual de operación.

16. Aseguramiento de la Calidad: (a) Monitorea y modifica los parámetros de proceso para prevenir fallas de calidad, (b) Realiza eficazmente inspecciones de calidad en el producto de acuerdo a los procedimientos establecidos y (c) Toma muestras representativas del producto de acuerdo a la periodicidad y los procedimientos de muestreo fijados para su posterior análisis de calidad

Validez

Definición Conceptual

Grado en que un instrumento mide lo que pretende medir (Hernández, 1998)

Definición Operacional

Validez de constructo, validez de contenido y validez concurrente del instrumento.

Dimensiones

1. Validez de constructo.
2. Validez de criterio externo
3. Validez de contenido

Indicadores

1. Validez de constructo: (a) Coeficiente de Pearson de correlación ítem-escala, (b) Coeficiente de Pearson correlación escala-test, y (c) Coeficiente de Pearson de correlación ítem-test.
2. Validez concurrente: (a) Coeficiente de Pearson de correlación puntaje global del instrumento- puntaje obtenido en el método de categorización individual, (b) Coeficiente de Pearson de correlación de puntaje global del instrumento- puntaje obtenido en el método de distribución forzada.
3. Validez de Contenido: mantención, eliminación y revisión de reactivos por juicios de expertos.

Confiabilidad

Definición Conceptual

Grado en que la aplicación repetida de un instrumento al mismo sujeto produce los mismos resultados ante igualdad de condiciones de aplicación (Nunally, 1995).

Definición Operacional

Coefficiente Alfa de Cronbach.

Aceptabilidad

Definición Conceptual

Grado en que un instrumento es aceptado por evaluados y evaluadores. (Ramos et al., 1996)

Definición Operacional

Grado en que el evaluador percibe que el instrumento es útil, práctico, confiable, válido, sensible y libre de sesgos.

Dimensiones

1. Percepción de utilidad.
2. Percepción de practicalidad.
3. Percepción de confiabilidad.
4. Percepción de validez
5. Percepción de discriminabilidad
6. Percepción de libertad de influencia de sesgos.

Indicadores

1. Percepción de utilidad: (a) Percepción de utilidad para tomar decisiones de promoción, despido y asignación a puestos de trabajo, (b) Percepción de utilidad para dar feedback del desempeño, (c) Percepción de utilidad para distribuir incentivos salariales y (d) Percepción de utilidad para detectar necesidades de capacitación.

2. Percepción de practicalidad: (a) Percepción de facilidad de comprensión, (b) Percepción de facilidad de uso.

3. Percepción de confiabilidad: (a) Percepción de consistencia de las puntuaciones entre distintos evaluadores y (b) percepción de consistencia de las puntuaciones en el tiempo

4. Percepción de validez: Percepción de valoración real del desempeño

5. Percepción de discriminabilidad: Percepción de discriminación entre los empleados con bajo nivel de desempeño y los empleados de alto nivel de desempeño

6. Percepción de libertad de influencia de sesgos: (a) percepción de indulgencia, (b) percepción de rigorismo, (c) percepción de tendencia a la medición central, (d) percepción de sesgo por amistad, (e) percepción de efecto halo, (f) percepción de prejuicios personales y (g) percepción de efecto de acontecimientos recientes.

Técnicas y Procedimientos de Recolección de Datos

Diagnóstico de Necesidades

El diagnóstico de necesidades, como primera fase del ciclo de diseño de prácticas de recursos humanos, se aboca a delimitar las necesidades de evaluación. En esta fase, este estudio utiliza la entrevista focalizada para la recolección de datos.

Objetivos y Pauta de Entrevista

Las entrevista de diagnóstico de necesidades de esta investigación tiene por objetivos determinar las necesidades de los futuros usuarios frente al sistema de evaluación de modo de diseñar un sistema de evaluación que cumpla con los requisitos psicométricos y de utilización así como disminuir las resistencias y potenciar el compromiso de jefes de producción, técnicos y asistentes técnicos frente al diseño de la propuesta de evaluación.

La entrevista focalizada está constituida por 15 preguntas, 3 abiertas y 12 cerradas (ver Apéndice C). Los temas que se abordan finalmente como parte de la entrevista son: necesidades de evaluación, aceptabilidad²⁵, aspectos varios²⁶, interés en la evaluación²⁷ y satisfacción con la entrevista de diagnóstico. Cabe destacar que los dos primeros temas son especialmente relevantes para el desarrollo de sistemas e instrumentos de evaluación.

Tabla N° 3: Preguntas y Contenidos de la Entrevista de Diagnóstico de Necesidades

Contenidos	f ⁽¹⁾	Preguntas
I. Necesidades de Evaluación		
A. Objetivos	1	1
B. Objeto	2	2, 3
C. Método	1	4
D. Evaluadores	1	5
II. Aceptabilidad	1	12
III. Aspectos varios		
A. Frecuencia de evaluación	1	6
B. Capacitación – Información sobre el sistema de evaluación	2	7, 11
C. Feedback	3	8, 9, 10
D. Otras necesidades	1	13
IV. Interés en la Evaluación de Desempeño	1	14
V. Satisfacción con la Entrevista	1	15

(1) Total de preguntas

²⁵ Si bien la aceptabilidad no constituye una necesidad de evaluación propiamente tal, no deja de ser importante considerarla como foco de interés del diagnóstico. Cabe destacar que si bien la aceptabilidad es propiedad de una medida de desempeño ya desarrollada, se estima que es necesario preguntar directamente al futuro usuario sobre qué es lo que necesita una evaluación de desempeño para ser aceptada por éste. De esta manera, parece razonable que si en el diseño se siguen las especificaciones señaladas por evaluadores y evaluados, la aceptabilidad de la propuesta de evaluación estaría asegurada. Lo anterior además posibilitaría una real participación del usuario en el diseño de sistemas de evaluación.

²⁶ Estos aspectos si bien no resultan claves para el diseño de instrumentos de evaluación, son útiles para el desarrollo del sistema de evaluación en el cual va inserto.

²⁷ Se considera relevante determinar el interés por el desarrollo de un sistema de evaluación en técnicos y asistentes técnicos de modo de prever la resistencia ante el proceso de implementación.

Procedimiento Seguido en el Diagnóstico de Necesidades

Para realizar estas entrevistas, se siguen los procedimientos que se describen a continuación. Sin embargo, hay consignar que, en primera instancia, se mantiene una reunión con el Departamento de Recursos Humanos y el Gerente de Operaciones para determinar el interés y el compromiso de la gerencia en evaluar el desempeño de técnicos y asistentes técnicos. En esta reunión se realiza una presentación introductoria al tema de evaluación de desempeño en cuanto a su importancia, puntos claves a definir (objetivos, evaluadores, objeto, etc.) y al procedimiento general que se sigue para la construcción e implementación de un sistema. A continuación, para obtener el apoyo de los jefes de producción, se efectúa una presentación similar que cuenta nuevamente con la asistencia del Departamento de Recursos Humanos y el Gerente de Operaciones. Se debe consignar que, en esta última reunión, se aprueba la construcción e implementación de un sistema de evaluación y la participación de técnicos y asistentes técnicos en este proceso -propuesta que había sido rechazada anteriormente por el Departamento de Recursos Humanos-, así como jefes de producción y la Gerencia expresan algunas de sus necesidades y expectativas frente al diseño e implementación del mismo.

Preparación de la Entrevista

Se definen, en primer lugar, los temas a tratar dada la revisión bibliográfica correspondiente y la discusión en reuniones anteriormente señaladas. En segundo lugar, con el propósito de construir la pauta de entrevista, se procede a revisar documentos de la empresa que resultaban pertinentes como el “Job Analysis: Technician and Assistant Technical” (Chiletabacos, 2001) y la “Declaración de Principios de Empleo” (ver Apéndice A). En segundo lugar, también se efectúa un análisis preliminar de la viabilidad de potenciales métodos, evaluadores, frecuencias de evaluación. Así, por ejemplo, en relación con el método de evaluación, se descartan como alternativas de pregunta el uso del método de centros de evaluación y el método de investigación de campo. Posteriormente, la pauta es revisada por dos jueces expertos externos y por tres expertos internos. La revisión de los jueces externos establece finalmente la pertinencia conceptual y claridad de las preguntas y alternativas.

En relación con la determinación de la muestra de tipo no probabilística por cuotas, se emplea el procedimiento de tómbola para la selección de técnicos y asistentes técnicos. A continuación, cada jefe de producción, empleando una pauta especialmente diseñada, comunica a la totalidad de los técnicos y asistentes técnicos a su cargo la realización de esta investigación y de las entrevistas mismas, comunicando además los sujetos a entrevistar.

Aplicación de la Entrevista

La realización de las entrevistas tiene lugar en los meses de diciembre a enero del 2003, en instalaciones de la empresa y dentro de la jornada laboral. Cabe notar que en la presentación, nuevamente se señalan a los entrevistados los objetivos de la entrevista y las etapas del diseño de un sistema de evaluación, recalcando la importancia de la participación de todos los actores organizacionales; así como se garantiza la confidencialidad de la entrevista. En el cierre, por su parte, se agradece la participación del entrevistado y se responden todas las consultas vinculadas con la evaluación de desempeño.

Instrumento de Evaluación de Desempeño

Descripción del Instrumento de Evaluación de Desempeño

El instrumento de evaluación propuesto consta de 114 ítems (ver Apéndice D), agrupados en 16 dimensiones de desempeño (Apéndice E). Cada dimensión tiene de 1 a 6 indicadores y 3 a 14 reactivos. A su vez, cada indicador tiene de 1 a 11 reactivos asociados. Cabe destacar que del total de reactivos, 57 se relacionan con dimensiones generales de desempeño de dominio y 57 con dimensiones técnicas²⁸.

La confiabilidad y validez así como la viabilidad del instrumento mismo se describen en el apartado de resultados.

²⁸ Esta clasificación se equipara a la establecida por el Modelo de Campbell. Es decir, se puede afirmar que la mitad de los reactivos se relacionan con tareas específicas del puesto de trabajo mientras que la otra mitad lo hace con tareas generales, incluyendo las tareas de comunicación verbal y escrita.

Tabla N° 4: Indicadores y Reactivos por Dimensión

Indicadores	f (1)	f (2)	Reactivos
Comunicación Oral	3	5	8, 24, 25, 62, 63
Comunicación Escrita	2	3	47, 57, 87
Escucha Activa	2	5	5, 53, 61, 72, 93
Relaciones Interpersonales	6	14	1, 7, 10, 13, 21, 27, 32, 37, 38, 49, 52, 58, 99, 108
Orientación a Resultados	1	4	30, 41, 85, 104
Resolución de Problemas	3	7	40, 43, 51, 88, 89, 110, 111
Innovación	2	4	28, 39, 46, 91
Cooperación	1	3	64, 92, 112
Aprendizaje	3	6	3, 29, 48, 81, 105, 113
Desarrollo de Pares	3	6	2, 44, 60, 69, 76, 90
Prevención de Riesgos	2	9	4, 14, 16, 26, 56, 59, 67, 70, 94
Manejo de Situaciones Críticas	2	5	79, 92, 101, 102, 114
Mantención Correctiva	1	12	15, 18, 55, 78, 84, 86, 95, 100, 103, 106, 107, 109
Mantención Preventiva	2	7	6, 12, 17, 23, 36, 65, 80
Operación de Máquina	3	10	9, 19, 22, 35, 50, 66, 68, 73, 74, 77
Aseguramiento de Calidad	3	14	11, 20, 31, 33, 34, 42, 45, 54, 71, 75, 82, 83, 96, 98

(1) Frecuencia de indicadores

(2) Frecuencia de reactivos

Procedimiento de Construcción

Procedimiento Previo a la Construcción del Instrumento: Diseño de un Prototipo de Sistema de Evaluación

Para construir el instrumento, se procede primero a diseñar el prototipo de sistema de evaluación, atendiéndose los resultados del diagnóstico, antecedentes empíricos y teóricos así como ciertas directrices y particularidades de la organización²⁹. Para ello, se definen primero los objetivos del sistema de evaluación; en segundo lugar, el criterio de evaluación y, con posterioridad, se evalúan distintas alternativas de métodos y evaluadores. Se determinan además los procedimientos a seguir como parte del proceso de evaluación y los requerimientos de utilización a considerar. (Ver Apéndice F y G)

²⁹ Así, por ejemplo, en la determinación del método, se presta consideración a la cultura evaluativa y el proceso de capacitación de la organización. En efecto, la decisión de evaluar mediante una Escala de Observación del Comportamiento se toma sobre la base de: (a) la congruencia de ésta con el método de evaluación del desempeño de los jefes de producción (se emplea una Escala Anclada en el Comportamiento) y (b) los contenidos de los programas de capacitación, los cuales son genéricos (no se imparten cursos diferenciales según el nivel de desempeño).

Delimitación de Dimensiones de la Variable Desempeño

La selección de las dimensiones de medición se basa principalmente en los resultados del análisis del trabajo que realizó la empresa³⁰ y en los resultados de la entrevista de diagnóstico de necesidades. Cabe considerar que todas las dimensiones establecidas por el análisis del puesto de trabajo, realizado en el 2001, exceptuando “Administración de Operaciones”, se mantuvieron en el instrumento como dimensiones de desempeño específicas³¹. Al respecto, es necesario precisar que sólo se efectúa una reagrupación de algunos de los componentes de éstas. Específicamente, los componentes que apuntaban a tareas no específicas de los puestos de técnico y asistente técnico y que eran transversales a todas las células, se constituyen en dimensiones generales de desempeño³². De este modo, se determinan finalmente las 16 dimensiones de desempeño.

Éstas se definen conceptualmente, a continuación, sobre la base de una revisión de definiciones dadas por el mencionado análisis del trabajo, los instrumentos de evaluación del desempeño de la empresa y de otras organizaciones así como bibliografía pertinente. Cabe destacar que las definiciones debían ser lo suficientemente claras para ser comprendidas por evaluadores y evaluados.

³⁰ La metodología del análisis del puesto de trabajo contempló la realización de un total de 48 entrevistas estructuradas. Se entrevistaron 40 técnicos/asistentes técnicos y la totalidad de los jefes de producción (8 sujetos). Se emplearon diferentes pautas de entrevista para entrevistar a técnicos/asistentes técnicos y jefes de producción (Ver Apéndice C). Es de suma relevancia consignar que las pautas de entrevista, especialmente la que se aplicó a los jefes de producción, estaba orientada a determinar incidentes críticos.

Se identificaron así las siguientes dimensiones claves de desempeño para el cargo de técnico y asistente técnico: seguridad, mantención preventiva, mantención correctiva, operación de máquina, aseguramiento de la calidad, y administración de operaciones. Dentro de éstas, se determinaron como componentes relevantes: Escucha activa, Resolución de problemas, Relaciones Interpersonales, Cooperación, Toma de decisiones, Análisis de problemas, Desarrollo de Pares, Entrega de feedback, Comunicación, Administración del tiempo, Sentimiento de productividad, Planificación, Organización/Dirección de reuniones y Liderazgo.

³¹ Estas son: Operación de Máquina, mantención preventiva, mantención correctiva, aseguramiento de la calidad, prevención de riesgos y manejo de situaciones críticas.

³² Las dimensiones generales son: comunicación oral, comunicación escrita, escucha activa, relaciones interpersonales, orientación a resultados, resolución de problemas, innovación, cooperación, desarrollo de pares y aprendizaje. Es necesario destacar que las dimensiones de dominio general son importantes porque encierran conductas de rol, pero además una serie de conductas prosociales que la empresa valora.

Construcción de Reactivos

Definidas las dimensiones de desempeño, se procede a diseñar los reactivos en función de:

1. Los resultados del análisis del trabajo y la descripción de puesto asociada.
2. El “Estudio Revisión de la Efectividad Organizacional ³³” (International Survey Research Offices Worldwide, 2001). Se detectó, en este cuestionario, un conjunto de reactivos que se relacionaban directamente con conductas prosociales que la empresa valora.

3. Reactivos de distintos instrumentos que valoran dimensiones similares de desempeño. Son especialmente útiles, el Instrumento de Evaluación de Motorola Italia, el Sistema de evaluación de Competencias diseñado por la Universidad Diego Portales (2000), Trayecto Técnico Profesional en Industrias de Procesos (Instituto Nacional de Educación Tecnológica, 2002), Trayecto Técnico Profesional en Producción Agropecuaria (Instituto Nacional de Educación Tecnológica, 2002), Currículo del ciclo formativo de grado superior, correspondiente al título de Técnico Superior en Producción de Madera y Mueble (Boletín Oficial de Navarra, 1997).

4. Los resultados de las entrevistas de Evaluación del Programa Pago por Habilidades³⁴ y de las entrevistas de diagnóstico de necesidades de evaluación. Como en la definición de dimensiones y construcción de los reactivos finalmente no se cuenta con la participación de técnicos y asistentes técnicos de modo de realizar una entrevista de incidentes críticos³⁵, se procede a revisar las transcripciones de las

³³ Este estudio se aplicó a todas las filiales de British American Tobacco. Se orientó a evaluar la gestión de las palancas o drivers de las personas (liderazgo, aprendizaje, cultura y talento) así como prácticas y temas relacionados con la gestión de recursos humanos, a saber: Información y Comunicación, Responsabilidad Corporativa, Estructura, Pago y Beneficios, Trabajo en Equipo, Respeto por nuestros Empleados, Seguimiento a la Encuesta y Fusión. Fue de especial utilidad la definición de los drivers de las personas pues orientan la gestión de recursos humanos, supliendo una estrategia formal de gestión.

³⁴ Si bien no se encontraron muchos incidentes críticos, estas entrevistas orientaran el diseño de reactivos por cuanto se mencionan conductas que se dan en el ejercicio cotidiano del cargo.

³⁵ En el desarrollo de esta investigación, ciertas condiciones organizacionales -dadas en el transcurso de la investigación- determinan la selección de las técnicas de recolección de datos y los procedimientos a seguir. En efecto, la reestructuración del Área de Operaciones incluye también el despido de un número importante de técnicos y asistentes técnicos. Dado el deterioro del clima laboral, el Área de Recursos Humanos decide que los técnicos y asistentes técnicos, por el momento, no continúen participando en la construcción e implementación de la propuesta de evaluación.

entrevistas señaladas de modo de aislar posibles descripciones de incidentes críticos que se hubiesen entregado en el desarrollo de las entrevistas mismas.

Determinación de Ponderaciones

En relación con la ponderación de las dimensiones de desempeño, se opta por ponderar a todas por igual, ya que este es un instrumento de índole exploratorio y no se lleva a cabo un análisis factorial.

Validación de Contenido y por Criterio Externo

Validación de contenido. Se utiliza el criterio de 6 jueces expertos, 5 externos a la organización y uno interno. Se determina así la claridad y la pertinencia conceptual de un ítem a una dimensión determinada y de las definiciones de las mismas. Se eliminan, incorporan y modifican los reactivos en los cuales al menos 2 jueces externos presentan consideraciones similares. Una vez efectuadas las modificaciones pertinentes, el experto interno determina finalmente: (a) la relevancia y claridad de cada reactivo, (b) la pertinencia de un reactivo a una determinada dimensión y (c) la posibilidad de observación de cada conducta establecida en los reactivos. Cabe señalar que en ambas revisiones, los expertos podían indicar reactivos que debían ser incluidos de modo de representar totalmente el dominio de contenidos que se pretende medir.

Participan en el proceso, los siguientes profesionales y académicos

1. Psicólogo. Profesor de Cátedra de la Escuela de Psicología de la Universidad de Valparaíso y Jefe del Departamento de Recursos Humanos de la Cámara de Diputados
2. Psicóloga. Profesora de Cátedra de la Escuela de Psicología de la Universidad de Valparaíso. Experta en Psicología Organizacional
3. Psicólogo. Profesor de Psicología Organizacional de la Universidad del Mar. Experto en Recursos Humanos y desarrollo de sistemas de evaluación del Departamento de Personal y Servicios del Senado.

4. Ingeniero Electrónico. Profesor de Cátedra de la Universidad Federico Santa María

5. Ingeniero Civil. Experto en Prevención de Riesgos. Consultor de S&AR Consultoría en Prevención de Riesgos

6. Ingeniero Mecánico. Jefe de Producción de la Célula de Printing de Chiletabacos.

Validez Concurrente³⁶. Una vez terminada la fase de observación del desempeño, se usan dos métodos para evaluar el desempeño de modo de establecer la validez concurrente: el método de categorización individual conocido como ranking de desempeño y el método de distribución forzada³⁷. En cuanto al primer método, el evaluador elabora un ranking de mejor a peor desempeño. A continuación, para evaluar con una distribución forzada³⁸, el jefe de producción procede a ubicar a cada empleado en función de su desempeño global en una las distintas categorías que se muestran a continuación³⁹.

Tabla N° 5: Determinación de Categorías del Método de Distribución Forzada

Categoría de desempeño	f ₍₁₎	fr
Desempeño muy satisfactorio	3	10%
Desempeño satisfactorio	6	20%
Desempeño regular	13	40%
Desempeño insatisfactorio	6	20%
Desempeño muy insatisfactorio	3	10%

(1) f en función de empleados evaluados

³⁶ Como el método de evaluación de desempeño escogido se orienta a medir el desempeño pasado, corresponde determinar la validez concurrente.

³⁷ Se aplica el método de formas alternas para ver confiabilidad. El coeficiente obtenido fue aceptable.

³⁸ El segundo método se aplica en virtud de la resistencia inicial del evaluador a emplear el método de categorización individual.

³⁹ Porcentajes de empleados fijados por la distribución típica expuesta por Cummings et. al (1985)

Procedimiento de Aplicación

El jefe de producción, en la primera semana de abril del 2004, completa las hojas de respuesta del instrumento de evaluación⁴⁰, registrando cualquier observación en torno al instrumento mismo (Ver Apéndice H). Para asegurar la confidencialidad de la evaluación, la aplicación tuvo lugar fuera de las dependencias de la empresa. Cabe destacar que antes de evaluar el desempeño y como parte de la fase de simulación, se efectúa una sesión de capacitación en la cual se entregó el manual del sistema de evaluación al evaluador. Con posterioridad a la capacitación, el jefe de producción observa el desempeño entre diciembre a marzo del 2004.

Entrevista de Evaluación de Aceptabilidad

Para evaluar el instrumento diseñado desde la perspectiva del usuario, se opta por aplicar una entrevista focalizada al evaluador. Ésta tiene por objeto explorar la aceptabilidad del instrumento desarrollado. Cabe destacar que ésta es la técnica de recolección de datos más apropiada porque permite abordar temas predefinidos y ahondar en otros que surjan en el desarrollo de la misma.

La entrevista aplicada consta de 22 preguntas (ver Apéndice C). Los puntos que se tratan en relación con la aceptabilidad son percepción de: utilidad, practicalidad, discriminabilidad, confiabilidad, validez y de libertad de influencia de sesgos. Se efectúan también tres preguntas generales.

Tabla N° 6: Contenido y Preguntas de la Entrevista de Aceptabilidad

Contenido	Preguntas
I. Percepción de Utilidad	1, 2, 3 y 4
II. Percepción de Practicalidad	5 y 6
III. Percepción de Sensibilidad	12
IV. Percepción de Validez	10 y 11
V. Percepción de Confiabilidad	7, 8 y 9
VI. Percepción de Libertad de Influencia de Sesgos	13, 14, 15, 16, 17, 18 y 19
VII. General	20, 21 y 22

⁴⁰ El jefe previamente debía revisar la hoja de registro de incidentes críticos, el manual y las instrucciones del cuaderno del instrumento.

En relación con la construcción de la pauta de entrevista, las preguntas se construyen sobre la base de la revisión bibliográfica pertinente. Posteriormente, esta pauta es revisada por 2 expertos externos para observar la claridad y pertinencia de las preguntas. Estos podían sugerir eliminar, revisar o añadir preguntas. Una vez realizados los cambios pertinentes, se entrevista al jefe de producción en dependencias de la empresa.⁴¹.

Técnicas de Análisis de Datos

Análisis de las Necesidades de Evaluación

En el análisis de datos de las entrevistas de diagnóstico, se efectúa un análisis estadístico de tipo descriptivo y un análisis de contenido de tipo cuantitativo. En relación con el primer análisis, se obtiene una distribución de frecuencias absolutas, relativas y acumuladas. En las preguntas que contenían una escala tipo Likert, se suma al análisis de frecuencias, el cálculo de la moda, desviación estándar, mediana, y media. En el análisis de las preguntas abiertas, se realiza un análisis cuantitativo del contenido de tipo cuantitativo, cuya unidad de análisis fue el tema⁴².

Análisis de las Propiedades del Instrumento

Validez

Para determinar la validez de constructo, dado el tamaño de la muestra, se aplica un análisis de las correlaciones ítem-test, ítem-escala (dimensión de desempeño), escala-test mediante el uso del programa SPSS.

Para determinar la validez concurrente, se efectúa un análisis de las correlaciones de los puntajes globales del instrumento propuesto con los resultados del método de categorización individual y método de distribución forzada, usándose el programa SPSS.

⁴¹ La entrevista tiene lugar después que el jefe de producción evaluó el desempeño de su personal por medio de la escala de observación de comportamientos desarrollada

⁴² El tema se define como una oración, enunciado respecto de algo. (Hernández et. al. 1999)

Confiabilidad

Para determinar la confiabilidad, se calcula el coeficiente Alfa de Cronbach del instrumento y por escala (dimensión) mediante el programa SPSS/PC para Windows.

Aceptabilidad

La entrevista se analiza mediante un análisis de contenido de tipo cuantitativo, usando las mismas categorías empleadas para analizar las entrevistas de diagnóstico de necesidades.

RESULTADOS

Necesidades de Evaluación

Objetivos

Tabla N° 7: Objetivos Generales

Categorías	Técnicos /Asistentes Téc.		Jefes de Producción		Gerente Oper./ RRHH	
	f	fr	f	fr	f	fr
I. Tomar decisiones administrativas	35	56%	5	28%	3	27%
II. Tomar decisiones de mejora del desempeño	4	6%	3	17%	2	18%
III. Dar feedback	8	13%	5	28%	3	27%
VI. Otros/No aplicable	16	25%	5	28%	3	27%
Total	63	100%	18	100%	11	100%

Tabla N° 8: Objetivos Específicos

Categorías	Técnicos /Asistentes Téc.		Jefes de Producción		Gerente Ope./ RRHH	
	f	fr	f	fr	f	fr
I. Tomar decisiones administrativas						
A. Promover	8	10%	3	14%	1	6%
B. Despedir	7	8%	3	14%	1	6%
C. Reasignar	4	5%	0	0%	3	17%
D. Distribuir recompensas	31	37%	2	9%	2	11%
II. Tomar decisiones de mejora del desempeño						
A. Capacitar	4	5%	3	14%	2	11%
B. Coaching	0	0%	0	0%	1	6%
III. Dar feedback	8	10%	5	23%	3	17%
IV. Planificar el desarrollo de carrera	9	11%	2	9%	1	6%
V Investigar	0	0%	0	0%	1	6%
VI. Otros	12	14%	4	18%	3	17%
Total	83	100%	22	100%	18	100%

Se observa en las tablas presentadas que si bien la mayoría del primer grupo y casi un tercio del segundo y tercer grupo⁴³ opina que es necesario diseñar un sistema de evaluación para apoyar la toma de decisiones administrativas, no existe consenso entre los grupos de entrevistados sobre qué decisiones administrativas se deberían tomar. Al respecto, el objetivo específico más mencionado por técnicos y asistentes técnicos es “Distribuir incentivos”. Cabe destacar que el objetivo general señalado es más mencionado que tomar decisiones de mejora. Finalmente, se debe acotar que el objetivo “Dar feedback” es la respuesta más referida por el segundo y tercer grupo de entrevistados.

Criterios y Dimensiones

Tabla N° 9: Resultados a Medir

Categorías	Técnicos/ asistentes técnicos		Jefes de Prod.		Gerente Oper./ RRHH	
	f	fr	f	fr	f	fr
I. Cantidad	30	55%	5	100%	3	100%
II. Calidad	41	75%	5	100%	3	100%
III. Oportunidad en la entrega	33	60%	3	60%	3	100%
IV. Otros						
A. Costos	23	42%	2	40%	1	33%
B. No aplicable	7	13%	0	0%	0	0%

⁴³ En lo sucesivo, nos referiremos al grupo de técnicos y asistentes técnicos como el primer grupo, al grupo de jefes de producción, segundo grupo y al grupo conformado por el gerente de Operaciones, gerente de Recursos Humanos y analista de Recursos Humanos, el tercer grupo.

Tabla N° 10: Dimensiones de Desempeño a Evaluar

Categorías	Técnicos/ asistentes técnicos		Jefes de Producción		Gerente Oper./ RRHH	
	f	fr	f	fr	f	fr
I. Técnicas						
A. Operación de máquina	36	65%	5	100%	3	100%
B. Mantenimiento	41	75%	5	100%	3	100%
C. Aseguramiento de la calidad	44	80%	4	80%	3	100%
D. Administración de Operaciones	10	18%	2	40%	0	0%
E. Seguridad	31	56%	3	60%	2	67%
F. Otras	2	4%	0	0%	0	0%
II. Generales						
A. Adaptabilidad	22	40%	4	80%	1	33%
B. Análisis de problemas	40	73%	5	100%	2	67%
C. Aprendizaje	28	51%	3	60%	2	67%
D. Decisión	37	67%	4	80%	2	67%
E. Disciplina	13	24%	2	40%	3	100%
F. Integridad	30	55%	4	80%	2	67%
G. Orientación al logro	22	40%	3	60%	3	100%
H. Planificación y organización	14	25%	1	20%	2	67%
I. Cooperación	45	82%	5	100%	3	100%
J. Tolerancia al estrés	19	35%	1	20%	1	33%
K. Relaciones interpersonales	31	56%	4	80%	3	100%
L. Visión del negocio	17	31%	2	40%	3	100%
M. Innovación	36	65%	3	60%	3	100%
N. Desarrollo de pares	30	55%	3	60%	2	67%
N. Uso y administración de bienes	14	25%	0	0%	0	0%
O. Otras						
O.1. Comunicación	31	56%	3	60%	2	67%
O.2. Asistencia	3	5%	0	0%	0	0%
O.3. Puntualidad	3	5%	0	0%	0	0%
O.4. Administración del tiempo	0	0%	0	0%	1	33%
O.5. Liderazgo	0	0%	0	0%	2	66%

Tabla N° 11: Ponderación de Criterios⁴⁴

Categorías	Técnicos/ asistentes técnicos						Jefes de Producción						Gerente Operaciones/				
	\bar{X}	S	Mín.	Máx.	Mod.	Med.	\bar{X}	S	Mín.	Máx.	Mod.	Med.	\bar{X}	S	Mín.	Máx.	Med.
I. Resultados	37	20	5	100	40	35	35	9,3	25	50		33	31	10	20	40	33
II. Dimensiones de desempeño técnicas	37	16	0	80	30	33	35	8	25	45		33	26	19	5	40	33
III. Dimensiones de desempeño generales	26	13	0	60	30	30	30	12,3	20	50	20	30	43	29	20	75	33

En cuanto al objeto de evaluación, es posible establecer que:

1. Los resultados y el desempeño resultan ser criterios válidos para los miembros de la organización.

2. En relación con los resultados, es importante destacar que, en los tres grupos, la mayoría de los entrevistados considera que los resultados operacionales -cantidad, calidad y oportunidad en la entrega- deberían ser medidos. Específicamente, se observa que calidad es el resultado más señalado por los tres grupos. Cabe acotar que la totalidad de los jefes de producción, por su parte, señala también el resultado “cantidad” mientras que todos los entrevistados del tercer grupo mencionan estos dos más “oportunidad en la entrega”.

3. En cuanto a dimensiones del ámbito técnico, la más señalada por los tres grupos es “Aseguramiento de calidad”. Cabe resaltar que la totalidad de los entrevistados del segundo y tercer grupo señalan también la dimensión “Operación de máquina” y “Mantenimiento”. Por otro lado, en cuanto a las dimensiones generales, la más seleccionada por los tres grupos de entrevistados es “Cooperación”. “Análisis de problemas” es apuntada también por todos los jefes de producción. Finalmente, cabe acotar que las dimensiones “Innovación”, “Disciplina”, “Relaciones interpersonales”, “Visión del negocio” más “Cooperación” son las más señaladas por los entrevistados del tercer grupo.

⁴⁴ Las ponderaciones se fijan en función de porcentajes. De este modo, la suma de los porcentajes asignados a los tres factores es igual a 100%.

4. El primer y segundo grupo asignan mayor ponderación a las dimensiones técnicas del desempeño y a los resultados mientras que el tercer grupo lo asigna a las dimensiones generales.

Métodos

Tabla N° 12: Métodos de Evaluación

Categorías	Técnicos /Asistentes Téc.		Jefes de Producción		Gerente Ope./ RRHH	
	f	fr	f	fr	f	fr
I. Escalas gráficas de calificación						
A. De acuerdo	27	49%	4	80%	2	67%
B. En desacuerdo	19	34%				
C. No sabe	9	17%	1	20%	1	33%
II. Método de selección forzada						
A. De acuerdo	27	49%	1	20%	1	33%
B. En desacuerdo	20	38%	1	20%	2	67%
C. No sabe	7	13%	3	60%		
III. Lista de verificación						
A. De acuerdo	15	28%	4	80%	2	67%
B. En desacuerdo	32	58%			1	33%
C. No sabe	7	13%	1	20%		
IV. Métodos de evaluación de grupos						
A. De acuerdo	19	36%	1	20%	1	33%
B. En desacuerdo	32	58%	3	60%	2	66%
C. No sabe	3	6%	1	20%		
IV. Administración por objetivos						
A. De acuerdo	39	70%	5	100%	5	100%
B. En desacuerdo	7	13%				
C. No sabe	9	17%				
V. Método de incidentes críticos						
A. De acuerdo	22	40%	3	60%	3	100%
B. En desacuerdo	26	13%				
C. No sabe	7	47%	2	40%		
VI. Escalas basadas en la conducta						
A. De acuerdo	35	63%	5	100%	3	100%
B. En desacuerdo	6	11%				
C. No sabe	14	26%				

Se puede observar en la tabla anterior que la mayoría de los entrevistados de los tres grupos se muestra a favor de evaluar mediante la administración por objetivos y las escalas basadas en la conducta, siendo el primero más señalado en los tres grupos. Por su parte, la mayoría de cada grupo no aprueba la aplicación de métodos de evaluación por grupos.

Evaluadores

Tabla N° 13: Evaluadores Potenciales

Categorías	Técnicos /Asistentes Téc.		Jefes de Producción		Gerente Ope./ RRHH	
	f	fr	f	fr	f	fr
I. Jefe de célula	14	25%				
II. Evaluado	2	4%				
III. Pares	2	4%				
IV. Jefe de célula y evaluado	12	22%	2	40%	1	33%
V. Jefe de célula y pares	6	11%				
VI. Jefe de célula, evaluado y pares	14	25%	2	40%	1	33%
VII. Otros						
A. Personal externo a la empresa	2	4%				
B. Personal de otras Áreas	2	4%			1	33%
C. Recursos Humanos	1	2%				
D. Jefe, pares, empleados y clientes internos			1	20%		

En cuanto a los datos presentados, las respuestas más seleccionadas por los entrevistados del primer grupo son el “Jefe de célula” y el “Jefe, el evaluado y pares”. En relación con el segundo y tercer grupo, ésta última junto con la combinación “Jefe y evaluado” son las repuestas más señaladas.

Aceptabilidad

Tabla N° 14: Especificaciones Generales⁴⁵

Categorías/subcategorías	Técnicos/ Asistentes		Jefes de Producción		Gerent. Oper. /RRHH	
	f	f r	f	f r	f	f r
I. Especificaciones de propiedades de las medidas de Evaluación						
A. Validez	7	5%				
B. Practicalidad	2	1%	1	6%	1	14%
C. Libertad de sesgos	10	7%	1	6%		
II. Especificaciones del sistema de evaluación						
A. Objetivos	11	8%	3	19%		
B. Criterio	4	3%				
C. Evaluadores	20	14%	1	6%		
D. Frecuencia	2	1%	1	6%		
III. Especificaciones del proceso de evaluación						
A. Observación	18	13%				
B. Evaluación	4	3%				
C. Feedback	27	19%	1	6%		
D. Otros	15	11%	3	19%	3	43%
IV. Especificaciones del proceso de diseño						
A. Simulación	15	11%	4	25%	2	29%
F. Otro	4	3%	1	6%	1	14%
Total	139	100%	16	100%	7	100%

Tabla N° 15: Especificaciones de Propiedades de las Medidas de Evaluación

Categorías/subcategorías	Técnicos/ Asistentes téc.		Jefes de Producción		Gerent. Oper. /RRHH	
	f	f r	f	f r	f	f r
I. Validez general						
A. Relevancia respecto del trabajo	7	37%				
II. Practicalidad						
A. Facilidad de comprensión	2	11%				
B. Facilidad de aplicación			1	50%	1	100%
III. Libertad de influencia de sesgos						
C. Libertad del sesgo por Amistad/simpatía	10	53%	1	50%		
Total	19	100%	2	100%		

⁴⁵ En esta tabla y en las posteriores, se muestran sólo las categorías que fueron señaladas por algún entrevistado.

Tabla N° 16: Especificaciones del Sistema de Evaluación

Categorías	Técnicos/ Asistentes téc.		Jefes de Producción		Gerent. Oper. /RRHH	
	f	f r	f	f r	f	f r
I. Objetivos del Sistema						
A. Despedir			2	29%		
B. Promover	4	11%	2	29%		
C. Distribuir recompensas	6	16%				
D. Otro	2	5%	1	14%		
II. Criterio a evaluar						
A. Resultados	1	3%				
B. Comportamientos	2	5%				
C. Rasgos de personalidad	1	3%				
III. Evaluadores						
A. Jefe	6	16%				
B. Autoevaluación	1	5%				
C. Jefe y compañero	3	8%				
D. Jefe, pares y empleado	10	26%	1	14%		
IV. Frecuencia						
A. Anual	1	3%				
B. Semestral	1	3%	1	14%		
Total	38	100%	7	100%		

Tabla N° 17: Especificaciones del Proceso de Evaluación

Categorías	Técnicos/ Asistentes téc.		Jefes de Producción		Gerent. Oper. /RRHH	
	f	f r	f	f r	f	f r
I. Observación						
A. Frecuente	4	6%				
B. Realizada en turnos de día y noche	2	3%				
II. Evaluación						
A. Conjunta empleado-jefe	2	3%				
B. Con detención-tiempo	2	3%				
III. Feedback						
A. Valencia positiva y negativa	8	13%				
B. Frecuente	5	8%				
C. Participativo	15	24%				
D. Fundamentado	6	10%				
IV. Otro						
A. Mecanismo de Revisión/apelación	6	10%				
B. Cumplimiento de normas y procedimientos del Sistema de Evaluación	2	3%			1	33%
C. Implementación de acciones	5	8%	3	60%	2	67%
D. Confidencialidad de resultados			2	40%		
E. Evaluación no anónima	3	5%				
f. No aplicables	3	5%				
Total	63		5	100%	3	100%

Tabla N° 18: Especificaciones del Proceso de Diseño

Categorías	Técnicos/ Asistentes téc.		Jefes de Producción		Gerent. Oper. /RRHH	
	f	f r	f	f r	f	f r
I. Simulación						
A. Capacitación a los evaluadores	3	14%	1	17%	2	33%
B. Capacitación a los evaluados	13	62%	2	33%	2	33%
E. Análisis de los resultados de la aplicación piloto			2	33%	1	17%
II. Otro						
A. Participación de los evaluados en el diseño	2	10%	1	17%	1	17%
B. Diseño dirigido por asesor externo a la empresa	3	14%				
Total	21	100%	6	100%	6	100%

En torno a la aceptabilidad de la evaluación de desempeño, los resultados muestran que:

1. Las respuestas más mencionadas, en el grupo de técnicos y asistentes técnicos, refieren al proceso de evaluación. En cambio, en el segundo y el tercer grupo, las respuestas vinculadas con el proceso de diseño son señaladas con mayor frecuencia.

2. En cuanto a las propiedades del sistema, la validez, practicalidad y libertad de influencia de sesgos son mencionadas. Al respecto, ninguna de éstas es indicada por los tres grupos. Cabe destacar en cuanto libertad de influencia de sesgo, la amistad/simpatía es el único sesgo mencionado.

3. En relación con las especificaciones del sistema de evaluación, es importante notar que la respuesta más señalada por técnicos y asistentes establece que el jefe, los pares y los propios evaluados evalúen el desempeño. Cabe consignar que el “jefe” es la segunda respuesta más indicada.

4. En torno a las especificaciones del proceso de valoración, la respuesta más referida por el segundo y tercer grupo es la necesidad de implementar las acciones de mejora. Por su parte, los entrevistados del primer grupo mencionan la entrega participativa de feedback con mayor frecuencia. Es decir, técnicos y asistentes esperaran tener un rol más activo en la entrega de

feedback, o sea, tener la posibilidad de expresar sus puntos de vista en torno a la evaluación y las acciones a tomar.

5. Finalmente, en cuanto al proceso de diseño, cabe enfatizar que la necesidad de capacitar a los evaluados sobre el sistema de evaluación es la respuesta más referida en los tres grupos de entrevistados.

De las Necesidades de Evaluación y de Aceptabilidad

En síntesis, en relación con las necesidades detectadas⁴⁶, se establece que:

1. Dada la falta de consenso al interior de los grupos de entrevistados y entre estos, la evaluación de desempeño debe responder a objetivos múltiples. Especialmente, debe servir para tomar decisiones administrativas y dar feedback a los técnicos y asistentes técnicos sobre su desempeño.

2. La evaluación de desempeño debe acompañarse de una evaluación de resultados. En cuanto a la evaluación del desempeño, deben evaluarse dimensiones técnicas y generales del trabajo. Al respecto, si se usa como criterio inicial una frecuencia relativa de respuesta por sobre el 51% en los tres grupos de entrevistados, las siguientes son dimensiones a evaluar: “Mantenimiento”, “Aseguramiento de la calidad”, “Operación de máquina”, “Innovación”, “Integridad”, “Relaciones Interpersonales”, “Seguridad”, “Cooperación”, “Análisis de problemas”, “Aprendizaje”, “Desarrollo de Pares” y “Comunicación”. Cabe destacar que las dimensiones técnicas y generales deben tener la misma ponderación en el instrumento. Finalmente, en cuanto a los resultados, la calidad, la oportunidad en la entrega y la cantidad deben medirse.

3. Se debe evaluar el desempeño por medio del método administración de objetivos y/o por una escala basada en la conducta.

⁴⁶ Ver Apéndice I para otros resultados del diagnóstico de necesidades.

4. El desempeño debe ser evaluado por el jefe de célula en conjunto con mismo evaluado y los pares o por los dos primeros.

5. Para desarrollar una evaluación de desempeño que sea aceptable por los usuarios potenciales, se deben considerar múltiples aspectos. Al respecto, se debe poner especial atención al desarrollo del sistema de evaluación y el proceso de valoración en el cual se inserta instrumento así como al proceso que se sigue para diseñar éste. En lo tocante a éste último, se debe capacitar a los evaluadores y evaluados así como garantizar la participación de los primeros en el diseño de las medidas de evaluación.

Propiedades Psicométricas del Instrumento

Validez

Validez de Contenido

En cuanto al proceso de validación de contenido del instrumento, en una primera revisión, los jueces externos determinaron que -de un total 246 reactivos presentados en el instrumento inicial (129 reactivos de dominio técnico y 115 de dominio general)- 53 ítems debían ser revisados (22% del total de reactivos), 4 eliminados (2% del total de reactivos), 2 agregados (1% del total de reactivos) (Ver Apéndice J).

Tabla N° 19: Ítems Revisados, Eliminados o Agregados por los Jueces Externos

Dimensión	Ítems revisados		Ítems eliminados		Ítems agregados	
	f	fr	f	fr	f	fr
Comunicación Oral	6	11%				
Comunicación Escrita	5	9%			2	100%
Escucha Activa	2	4%	1	25%		
Relaciones Interpersonales	6	11%				
Orientación a Resultados	2	4%				
Resolución de Problemas	4	8%				
Innovación	6	11%				
Cooperación	1	2%				
Aprendizaje	4	8%				
Desarrollo de Pares	5	9%				
Prevención de Riesgos	2	4%	1	25%		
Manejo de Situaciones Críticas	2	4%				
Gestión de la Seguridad						
Mantenimiento Correctiva	2	4%	1	25%		
Mantenimiento Preventiva	2	4%				
Gestión de la Mantenición			1	25%		
Operación de Máquina	2	4%				
Aseguramiento de Calidad	2	4%				
Gestión de la Calidad	0					
Total	53	100%	4	100%	2	100%

Como se puede observar en la tabla anterior, la mayoría de las dimensiones incluía reactivos que debían ser revisados, exceptuando las siguientes dimensiones: Gestión de la Calidad, Gestión de la Mantenición y Gestión de la Mantenición.

Una vez modificados, agregados o eliminados los reactivos correspondientes, el juez interno procedió a eliminar los reactivos que no eran relevantes, pertinentes o que presentaban dificultades para su observación; reduciéndose el número de reactivos de 244 a 114.

Tabla N° 20: Ítems Mantenidos y Eliminados por el Juez Interno

Dimensión	Ítems mantenidos		Ítems eliminados	
	f	fr	f	fr
Comunicación Oral	5	4%	6	5%
Comunicación Escrita	3	3%	8	6%
Escucha Activa	5	4%	5	4%
Relaciones Interpersonales	14	12%	9	7%
Orientación a Resultados	4	4%	1	1%
Resolución de Problemas	7	6%	4	3%
Innovación	4	4%	6	5%
Cooperación	3	3%	1	1%
Aprendizaje	6	5%	8	6%
Desarrollo de Pares	6	5%	11	8%
Prevención de Riesgos	9	8%	15	12%
Manejo de Situaciones Críticas	5	4%	2	2%
Gestión de la Seguridad	0	0%	11	8%
Mantenimiento Correctiva	12	11%	2	2%
Mantenimiento Preventiva	7	6%	13	10%
Gestión de la Mantenimiento	0	0%	14	11%
Operación de Máquina	10	9%	3	2%
Aseguramiento de Calidad	14	12%	5	4%
Gestión de la Calidad	0	0%	4	3%
Total	114	100%	130	100%

La tabla N° 20 muestra que todas las dimensiones contenían reactivos que no eran relevantes para el desempeño del cargo o presentaban dificultades para su observación. Al respecto, cabe destacar que tres dimensiones fueron eliminadas, a saber gestión de la mantención, gestión de la calidad y gestión de la seguridad⁴⁷.

⁴⁷ Si bien éstas fueron definidas como dimensiones claves por el análisis del trabajo que efectuó la empresa y tanto la gerencia como los jefes de producción consideran que es relevante desarrollar éstas en el personal, en la revisión del juez interno se eliminan los reactivos de dichas dimensiones por la poca factibilidad de observar los comportamientos contenidos en los reactivos, ya que todavía no se han delegado las tareas asociadas a éstas dimensiones.

Validez de Constructo

Tabla N° 21: Correlación Ítems- Test

Dimensión	N° ítem	Coef. de Pearson	Dimensión	N° ítem	Coef. de Pearson	Dimensión	N° ítem	Coef. de Pearson
Comunicación oral	8	0,646**	Innovación	28	0,700**	Mantención Correctiva	15	0,521**
	24	0,530**		39	0,588**		18	0,433*
	25	0,211		46	0,462*		55	0,35
	62	0,606**		91	0,700**		78	0,35
	63	0,214		64	0,517**		84	0,194
Comunicación escrita	47	0,301	Cooperación	97	0,513**		86	0,441*
	57	0,614**		112	0,601**		95	0,479**
	87	0,496**	Desarrollo de pares	2	0,802**		100	0,007
Escucha activa	5	0,554**		44	0,636**		103	0,007
	53	0,442*		60	0,773**		106	0,163
	61	0,527**		69	0,472**	107	0,384*	
	72	0,572**		76	0,603**	109	0,384*	
93	0,569**	90	0,674**	Operación de máquina	9	0,692**		
Relaciones Intepersonales	1	0,478**	Aprendizaje		3	0,704**	19	0,597**
	7	0,530**			29	0,446*	22	0,584**
	10	0,512**			48	0,598**	35	0,581**
	13	0,458*			81	0,675**	50	0,633**
	21	0,600**			105	0,606**	66	0,675**
	27	0,579**	113		0,708**	68	0,672**	
	32	0,527**	Prevención de riesgos		4	0,265	73	0,525**
	37	0,642**			14	0,435*	74	0,730**
	49	0,631**			16	0,505**	77	0,622**
	52	0,585**		26	0,344	11	0,735**	
58	0,555**	56		0,066	20	0,384*		
38	0,428*	59		0,305	31	0,518**		
99	0,526**	67		0,623**	33	0,399*		
108	0,462**	70		0,517**	34	0,191		
Orientación a resultados	30	0,609**	94	0,531**	42	0,479**		
	41	0,630**	Manejo de situaciones críticas	79	0,727**	Aseguramiento de Calidad	45	0,626**
	85	0,556**		92	0,595**		54	0,596**
104	0,597**	101		0,567**	71		0,485**	
Resolución de problemas	40	0,635**		102	0,3		75	0,571**
	43	0,409*	114	0,609**	82		0,129	
	51	0,430*	Mantención Preventiva	6	0,01		83	0,678**
	88	0,334		12	0,563**		96	0,550**
	89	0,334		17	0,483**		98	0,606**
	110	0,592**		23	0,576**		** Correlación significativa al 0,01 nivel de significancia	
111	0,596**	36		0,654**	* Correlación significativa al 0,05 nivel de significancia			
		65		0,369*				
		80	0,605**					

Se observa que todas las correlaciones de los ítem son positivas y la mayoría es significativa. En efecto, de un total de 114 ítems, 81 (71%) presentan correlaciones significativas con un α 0,01 y 15 ítems (13%) con α 0,05.

Tabla N° 22: Correlación Ítem- Dimensión

Dimensión	N° ítem	Coef. de Pearson	Dimensión	N° ítem	Coef. de Pearson	Dimensión	N° ítem	Coef. de Pearson
Comunicación oral	8	0,701**	Innovación	28	0,951**	Mantención Correctiva	15	0,476**
	24	0,800**		39	0,953**		18	0,781**
	25	0,464**		46	0,792**		55	0,768**
	62	0,711**		91	0,951**		78	0,768**
	63	0,555**		64	0,926**		84	0,534**
Comunicación escrita	47	0,534**	Cooperación	97	0,932**		86	0,334
	57	0,913**		112	0,953**		95	0,659**
	87	0,789**	Desarrollo de pares	2	0,856**		100	0,618**
Escucha activa	5	0,797**		44	0,813**		103	0,618**
	53	0,648**		60	0,858**		106	0,614**
	61	0,877**		69	0,790**	107	0,773**	
	72	0,827**		76	0,809**	109	0,773**	
	93	0,872**		90	0,764**	9	0,917**	
Relaciones Intepersonales	1	0,722**	Aprendizaje	3	0,922**	19	0,923**	
	7	0,698**		29	0,770**	22	0,811**	
	10	0,588**		48	0,671**	35	0,811**	
	13	0,561**		81	0,746**	50	0,848**	
	21	0,759**		105	0,637**	66	0,840**	
	27	0,827**		113	0,719**	68	0,882**	
	Operación de máquina	32	0,544**	Prevención de riesgos	4	0,628**	73	0,894**
		37	0,675**		14	0,786**	74	0,915**
		49	0,704**		16	0,605**	77	0,916**
		52	0,739**		26	0,703**	11	0,792**
		58	0,587**		56	0,203	20	0,667**
		38	0,695**		59	0,706**	31	0,806**
		99	0,773**		67	0,672**	33	0,686**
108		0,681**	70		0,778**	34	0,545**	
Orientación a resultados	30	0,764**	94	0,802**	42	0,707**		
	41	0,850**	Manejo de situaciones críticas	79	0,861**	Aseguramiento de Calidad	45	0,765**
	85	0,783**		92	0,950**	54	0,561**	
	104	0,783**		101	0,937**	71	0,447*	
Resolución de problemas	40	0,788**		102	0,867**	75	0,34	
	43	0,722**	114	0,836**	82	0,357		
	51	0,741**	Mantención Preventiva	6	0,634**	83	0,605**	
	88	0,902**		12	0,937**	96	0,817**	
	89	0,902**		17	0,914**	98	0,823**	
	110	0,914**		23	0,920**	** Correlación significativa al 0,01 nivel de significancia		
	111	0,881**		36	0,803**	* Correlación significativa al 0,05 nivel de significancia		
		65		0,870**				
		80		0,715**				

Se observa que la mayoría de los ítems presenta correlaciones positivas y significativas. Específicamente, 108 ítems (95) arrojan correlaciones significativas con α 0,01 y un ítem presenta correlación significativa con α 0,05.

Tabla N° 23: Correlación Dimensión – Test

Dimensión	Coefficiente de Pearson
Comunicación Oral	0,681**
Comunicación Escrita	0,677**
Escucha Activa	0,668**
Relaciones Interpersonales	0,774**
Orientación a Resultados	0,755**
Resolución de Problemas	0,588**
Innovación	0,678**
Cooperación	0,581**
Aprendizaje	0,814**
Desarrollo de Pares	0,814**
Prevención de Riesgos	0,559**
Manejo de Situaciones Críticas	0,641**
Mantención Correctiva	0,483**
Mantención Preventiva	0,549**
Operación de Máquina	0,715**
Aseguramiento de Calidad	0,770**

** : coeficiente significativo al 0.01 nivel de significancia

La tabla anterior muestra que todas las dimensiones presentaron correlaciones positivas y significativas (α 0,01).

En resumen, los resultados obtenidos en cuanto a las correlaciones ítem-test, ítem-escala, escala-test indicarían que el instrumento posee un nivel aceptable de validez de constructo. No obstante, estos resultados son de carácter preliminar, ya que la determinación de este tipo de validez implica la realización de un análisis factorial.

Tabla N° 24: Correlación del Instrumento con Métodos de Evaluación de Grupo

Medidas de desempeño	Coefficiente de Pearson
Método de Categorización Individual- Instrumento	0,759**
Método de Distribución Forzada-Instrumento	0,858**

** : Correlación significativa al 0.01 nivel de significancia

La correlación entre el instrumento y los métodos de evaluación es muy alta. Cabe destacar que la correlación entre los métodos de evaluación de grupo fue 0,809. Se puede afirmar, por ende, que el instrumento aplicado en la muestra de técnicos y asistentes técnicos posee un nivel aceptable de validez concurrente

Confiabilidad

El coeficiente Alfa de Cronbach del instrumento es 0,973. Por su parte, los índices de confiabilidad por dimensión son los siguientes:

Tabla N° 25: Confiabilidad por Dimensión

Dimensión	Índice*
Comunicación Oral	0,638
Comunicación Escrita	0,534
Escucha Activa	0,852
Relaciones Interpersonales	0,902
Orientación a Resultados	0,794
Resolución de Problemas	0,918
Innovación	0,933
Cooperación	0,93
Aprendizaje	0,835
Desarrollo de Pares	0,89
Prevención de Riesgos	0,823
Manejo de Situaciones Críticas	0,913
Mantenimiento Preventiva	0,92
Mantenimiento Correctiva	0,847
Operación de Máquina	0,964
Aseguramiento de Calidad	0,884

122 * Alfa de Cronbach

Los resultados indican que el instrumento posee una confiabilidad muy alta según las normas descritas por Nunally (1999). Además, se puede apreciar que las dimensiones presentan una confiabilidad alta, exceptuando de Comunicación Oral, Comunicación Escrita y Orientación a Resultados. El coeficiente α de éstas indica una confiabilidad moderada.

Evaluación del Usuario: Aceptabilidad

Percepción de Utilidad

Tabla N° 26: Percepción General de Utilidad

Categorías	f	fr
A. Muy Alta	4	100%
B. Alta		
C. Media		
D. Baja		
E. Muy Baja		
Total	4	100%

La utilidad percibida para todos los objetivos fijados es muy alta. Específicamente, se observa que:

1. El instrumento se percibe como muy útil para “Tomar decisiones administrativas”, ya sea de colocación o de compensación. En cuanto a tomar decisiones de colación, el jefe de producción señala el método (escala diseñada) así como el proceso de valoración en cuanto a la etapa de observación y de entrega de feedback como factores que contribuyen a la utilidad. El evaluador plantea así:

“Esto no quiere decir tal puntaje, tal decisión. No es una receta... Por ejemplo, si tengo un empleado muy bueno y en una evaluación sale mal porque tiene una hija muy enferma, no puedo despedirlo así como así. También en los asensos, hay que tener cuidado porque depende siempre del cargo que está vacante... La reasignación es otro punto porque hay que tener claro que cargos y qué tareas puedes rotar. *Esto no lo tiene el cuestionario porque las preguntas sólo dicen qué tan seguido hace esto o lo otro.* Pero igual les daría tareas que sean más complejas a los que tienen mejor desempeño... Por eso, hay que observar a la persona y escucharla en la reunión de evaluación”

En cuanto al objetivo “Distribuir incentivos”, se alude sólo al método (escala diseñada). Al respecto, el jefe de producción comenta: “...porque tengo entendido que los resultados del cuestionario se expresan en puntajes. Se me ocurre que se podría establecer que tanto puntaje, tanto incentivo”.

2. En torno al objetivo “Tomar decisiones de mejora”, el evaluador se refiere al método (escala diseñada) y la entrega de feedback. Al respecto, éste relata:

“Con los puntajes, por pregunta o dimensión, yo puedo ver que persona tiene que ir a un curso X. Hoy no hay criterios claros, sólo se mandaba por el sueldo. También me ayuda en esto, la reunión de feedback (cuando el técnico o asistente dice cómo podría mejorar) y la matriz de decisión porque todo se tiene que capacitar. En el nivel celular, me hace ver los cursos que hay crear, sacar o cambiar para que se adapten a las necesidades del momento. También las mismas preguntas son pueden servir para hacer mallas”.

3. En cuanto al objetivo “Dar feedback”, se alude al método (escala diseñada). En relación con lo anterior, el evaluador plantea:

“Antes no podía decir nada porque no tenía argumentos... Ahora, con los resultados, puedo decir específicamente que tiene que seguir haciendo y que no, cuáles son las áreas débiles y fuertes de la persona. *Puedo también dar ejemplos con las mismas preguntas del cuestionario*”

Como se observa, la percepción de utilidad del instrumento se relaciona con el método de evaluación así como con el proceso de valoración. Según lo expresado por el evaluador, no sólo es determinante el método de evaluación, sino la reunión de evaluación o feedback también contribuiría al logro de los objetivos del sistema.

Percepción de Practicalidad

Tabla N° 27: Percepción General de Practicalidad

Categorías	f	fr
A. Muy Alta	1	50%
B. Alta	1	50%
C. Media		
D. Baja		
E. Muy Baja		
Total	2	100%

La percepción de practicalidad del instrumento es alta a muy alta. Al respecto, en cuanto a la facilidad de comprensión, se alude al proceso de diseño; específicamente, al procedimiento de validación de contenido. Por su parte, en lo tocante a la facilidad de aplicación, se hace referencia al método (instrumento diseñado). El evaluador afirma así: “Alta porque es fácil contestar el cuestionario... *No es muy alta porque uno se demora contestando.* En la primera mitad, me demoré 20 minutos por persona y, en la segunda tirada, como 15 minutos...”

Percepción de Confiabilidad

Tabla N° 28: Percepción General de Confiabilidad

Categorías	f	fr
A. Muy Alta		
B. Alta	3	100%
C. Media		
D. Baja		
E. Muy Baja		
Total	3	100%

La percepción de confiabilidad general del instrumento es alta. Las respuestas del entrevistado aluden al método empleado (instrumento diseñado) así como al proceso de diseño y de valoración. Específicamente, la capacitación y la observación del desempeño son mencionadas. En relación con lo expuesto, el jefe de producción afirma:

“Alta porque observe durante tres meses el desempeño. Además, traté dentro de lo posible, no caer en los errores que vimos en la capacitación... En la evaluación, eso sí siempre existe la posibilidad que me haya equivocado respecto a una persona. Pero, en general, creo que lo que coloqué es así.

Alta porque el cuestionario es bien objetivo. No es como poner muy bueno o malo por cada factor. Pero... por ejemplo, un mecánico se puede demorar 10 minutos haciendo la reparación a una máquina, para mí eso es rápido, pero puede que no sea tan así para otro jefe. *El desafío ahora es crear criterios más específicos.* Ejemplo: “tal falla, tanto minutos se tiene que demorar”. Igual eso es difícil porque hay muchos factores externos que influyen, pero serviría de referencia.”

Alta porque no las personas no cambian tanto... igual depende de las acciones que se tomen... Por ejemplo, si lo mando a un curso de comunicación, espero que en la próxima evaluación mejore en esto. Espero que se vean tendencias de mejora”

Percepción de Validez

Tabla N° 29: Percepción General de Validez

Categorías	f	fr
A. Muy Alta	1	50%
B. Alta	1	50%
C. Media		
D. Baja		
E. Muy Baja		
Total	2	100%

El instrumento es percibido como válido a muy válido. Al respecto, sólo el método (escala diseñada) y el proceso de diseño son mencionados. En cuanto a éste último, se alude al uso de los resultados del análisis del trabajo que efectuó la empresa y el procedimiento de validación seguido. Así, el evaluador señala:

“Muy alta porque el cuestionario es bien completo. Se pregunta desde cómo se lleva un técnico o asistente conmigo y sus compañeros a si opera adecuadamente las herramientas...”

“Alta porque se tomaron en cuenta los resultados del Job Analysis. Se eliminaron, además, todas las preguntas que eran no eran importantes en este momento... Pero hay que tomar en cuenta que eso puede cambiar. Algo que se saque, por ejemplo, la gestión de mantención, puede ser importante a futuro.”

Percepción de Discriminabilidad

Tabla N° 30: Percepción General de Discriminabilidad

Categorías	f	fr
A. Muy Alta		
B. Alta		
C. Media	1	100%
D. Baja		
E. Muy Baja		
Total	1	100%

La percepción de la capacidad del instrumento para discriminar entre el desempeño satisfactorio y el desempeño insatisfactorio fue media. Al respecto, el método usado así como el proceso de valoración en cuanto a la observación del desempeño son mencionados. De esta manera, el jefe de producción sostiene que: “Media... me tinca porque marqué “regularmente” en casi todos los casos. Para mí es super difícil poner siempre o nunca porque uno no está ahí todo el tiempo...”

Percepción de Libertad de Influencia de Sesgos

Tabla N° 31: Percepción General de Libertad de Influencia de Sesgos

	f	fr
A. Muy Alta	5	71%
B. Alta	2	29%
C. Media		
D. Baja		
E. Muy Baja		
Total	7	100%

La percepción de libertad de influencia de sesgos varia entre alta a muy alta. La respuesta más frecuente es muy alta, lo que quiere decir que el jefe de producción opina que las puntuaciones estarían libres de los sesgos propios de los evaluadores, con excepción del sesgo por amistad- simpatía y la tendencia a la medición central. En estos, la percepción de libertad de influencia fue “alta”.

Se señala el método y la observación del desempeño en relación con la libertad de influencia de sesgos. El evaluador también alude a su baja habilidad para diferenciar como fuente potencial de distorsión. Así, el jefe de producción en relación con la tendencia a la medición central afirma que:

“Reconozco que a mí me cuesta ver los extremos. Me cuesta ver las cosas muy positivas o negativas... para mí toda la gente tiene desempeño más o menos igual, ni muy bueno, ni muy malo. Igual hay algunas excepciones, pero son contadas con los dedos de una mano... no me atreví a marcar siempre y nunca porque uno, salvo contadas excepciones, no puede decir “tú nunca hiciste esto” porque uno no está todo el tiempo mirando a la persona.”

En cuanto al sesgo por amistad o simpatía, el método usado así como la observación del desempeño son mencionadas como relevantes para su control. Así, el jefe de producción señala que:

“Alta...porque el cuestionario no es muy subjetivo como poner bueno o malo... Evaluar me sirvió para darme cuenta que los técnicos y asistentes que me caían mejor, no tenían necesariamente un buen desempeño... hay gente que me di cuenta que sacarían mejor nota de lo que tenía pensado. Si hubiera tenido que subir de puesto a alguien, no la hubiera elegido... elegiría a los primeros porque pensaría que cachan más, están más comprometido con la pega, etc.”

Cabe destacar que, en cuanto a su contribución a restringir la influencia de sesgo, la observación del desempeño es aludida en relación con el efecto de acontecimientos recientes y de halo así como por prejuicios personales. Al respecto de estos, se percibió una muy alta libertad de sesgo.

Aspectos Generales

En relación con el instrumento, el evaluador señala la validez en cuanto a la factibilidad de observar los comportamientos contenidos en algunos reactivos y la operatividad como debilidades. Al respecto, el jefe de producción acota que: “... algunas preguntas no se aplican a todos, por ejemplo, un técnico malo no capacita... no lo puedo tener como instructor. Son sólo 2 o 3 los que ejercen como instructores. También me demoré contestando el cuestionario. Igual si uno lo piensa bien 15 o 20 minutos por persona no es tanto, pero si son 30 personas las que uno tiene a cargo...”.

En relación con las fortalezas, la validez y utilidad del instrumento son mencionadas. Estas fortalezas junto con el proceso de diseño y de valoración son referidas en relación con la satisfacción con el instrumento. Específicamente, se mencionan la participación en la entrevista de diagnóstico y en la validación de contenido, la sesión de capacitación y la observación del desempeño del personal a cargo”. Cabe resaltar que el nivel de satisfacción es muy alto. Al respecto, el evaluador afirma que:

“Mucho... Primero, esto me obligó ver a la gente haciendo su trabajo en el piso. Me di cuenta de varias cosas que no tenía idea porque no eran problemas graves, pero creo que igual influyen el proceso. Ejemplo, nadie busca las pautas de mantención preventiva que hay que aplicar. Se hace todo por intuición o experiencia. Segundo, repito de nuevo, el cuestionario es super completo, va de lo soft a lo hard... Tercero, esto me permite dar feedback... para que estén más tranquilos y también me da un piso para decidir qué hay que hacer para mejorar, quién se queda, quién se va. Cuarto, creo que fue positivo haber tenido la oportunidad de revisar las preguntas y las variables, dar mi opinión al principio, resolver las dudas en la reunión...”

Dado lo anterior, la recomendación entregada en cuanto a revisar, aplicar o suspender la aplicación del instrumento es aplicar. Sin embargo, recomienda, en lo posible, disminuir el tiempo empleado en la evaluación misma.

DISCUSIÓN

Resultados

Necesidades de Evaluación

En la fase de diagnóstico nos permitió explorar las necesidades de evaluación que presentan los miembros de la organización. Al respecto, la mayoría de los resultados muestran una gran diversidad de respuestas, lo que sugiere un bajo grado de acuerdo al interior y entre los grupos de entrevistados. En contraposición, los resultados complementarios resultan más concluyentes. Se puede observar en ellos, un alto grado de acuerdo al interior de los grupos y entre los grupos de entrevistados. Dado lo anterior, se concluye que, en primera instancia, cualquiera alternativa de diseño que sea seleccionada no logra satisfacer a todos los miembros de la organización. En segunda instancia, parece haber más consenso sobre los aspectos accesorios de la evaluación de desempeño que sobre los que constituyen el meollo de la misma. Es entonces tarea del profesional de recursos humanos seleccionar la mejor alternativa de evaluación, o sea, la que más se adapte a las particularidades de la organización de modo de diseñar e implementar un sistema de evaluación e instrumento que sea aceptado por los usuarios y que asegure el cumplimiento de los requisitos de validez y confiabilidad.

En relación con el valor de la entrevista focalizada para el diagnóstico de necesidades, cabe finalmente destacar que además de explicitar las necesidades y expectativas de los usuarios, posibilita la participación activa de los mismos desde la primera fase del ciclo de diseño. Dado el alto grado de satisfacción con la misma, el formato de entrevista resulta valioso en el desarrollo de sistemas de evaluación orientados a satisfacer las necesidades de los usuarios.

Propiedades Psicométricas

Los resultados del análisis de las propiedades psicométricas son auspiciosos, pero no concluyentes.

El análisis de correlaciones ítem-test sugiere que los reactivos de la prueba dan cuenta del constructo desempeño por cuanto todas las correlaciones son positivas y la mayoría son altamente significativas ($\alpha 0,01$), aún cuando pertenezcan a dimensiones de dominio técnico o general. En cuanto a las últimas, es importante notar que contienen reactivos orientados a evaluar conductas extra-rol. Esto indicaría que el desempeño no sólo está conformado por conductas que contribuyen al logro de los objetivos de la organización sino por conductas que potencian el buen funcionamiento de la organización. Cabe notar que un reducido número de reactivos presentan coeficientes regulares y un reactivo requiere ser interpretado con cautela, ya que presenta un coeficiente de Pearson menor a 0,2 (Ítem N° 6). Es posible, según lo señalado por Guilford (1984) que estos reactivos estén contaminados por algún factor o factores no controlados o que se han mantenido constantes. Al respecto, según las observaciones del evaluador, la mayor parte de estos reactivos presentan dificultades para su observación. Por ejemplo, en relación con el ítem N° 6, el evaluador señala que ninguno de sus subordinados “*Busca, al inicio de su turno, la pauta de mantención preventiva*” aún cuando los procedimientos de mantención lo establecen con claridad y los empleados han sido capacitados al respecto. Desde una óptica psicométrica o de investigación, debería considerarse descartar este reactivo, ya que éste reflejaría una inconsistencia con el constructo medido. Desde una óptica aplicada, no obstante, el reactivo es valioso para la evaluación de desempeño. En efecto, permite al jefe de producción identificar los procedimientos que no están siendo seguidos por el personal. Al respecto, es importante destacar que todos los reactivos que presentan coeficientes regulares pertenecen a dimensiones técnicas, a saber: prevención de riesgos, mantención preventiva, mantención correctiva y aseguramiento de calidad, así como no presentan objeciones en cuanto a su pertinencia y claridad por parte de los expertos externos y del jefe de producción. Finalmente, cabe acotar que se aplican las mismas consideraciones en cuanto a la correlación ítem -escala. En efecto,

la mayoría de los ítems que presenta baja correlación también arroja bajos coeficientes de correlación ítem-test.

En relación con la validez de contenido, es importante resaltar que en la revisión del evaluador se elimina la mitad de los reactivos del instrumento inicial por considerarse irrelevantes para el puesto de trabajo o presentar dificultades para su observación. A la luz de los resultados de la entrevista de aceptabilidad y las observaciones registradas en torno al instrumento, es posible conjeturar que la necesidad de disminuir la longitud de la prueba influyera en la decisión de eliminar un número tan elevado de reactivos. Del mismo modo, se deduce que el procedimiento seguido para validar el contenido no garantiza la completa eliminación de los reactivos que presentaban dificultades de observación. Dado lo anterior, parece recomendable que éste procedimiento incluya un breve periodo de observación de los comportamientos contenidos en los reactivos como un paso previo a la aplicación del instrumento.

En cuanto a la validez de criterio externo, los resultados sugieren que el instrumento representa una buena medida del desempeño global. De esta manera, a mayor puntaje obtenido en el instrumento, mejor posición ocupa el empleado en el ranking de desempeño y en la lista de distribución forzada. Se observa además que el instrumento diseñado presenta una correlación más elevada con el método de distribución forzada, el cual resulta más empático para el evaluador. Esto sugiere que la consideración de las necesidades de los usuarios, en lo referido a la selección del criterio externo, puede incidir en los índices de validez de criterio que un investigador obtiene. Cabe acotar que la ventaja que ofrece el instrumento propuesto frente a los métodos de evaluación de grupo se refiere a la utilidad para dar feedback significativo y para detectar necesidades de capacitación.

La aplicación del instrumento en la muestra definida, por otra parte, arroja un muy alto nivel de confiabilidad, lo que indica que los resultados serían generalizables a la población. A su vez, indica que si no se introducen acciones de mejora y/o no se entrega feedback a los empleados sobre su desempeño, los puntajes que obtendría un mismo individuo en aplicaciones sucesivas del instrumento serían similares. Dado lo expuesto y siguiendo las normas de confiabilidad de Nunnally (1999), los evaluadores así podrían

tomar decisiones sobre la base de los resultados obtenidos en la aplicación del instrumento propuesto, ya que estos serían exactos.

Un coeficiente de Cronbach tan alto puede ser explicado por diversos factores que dicen relación con la metodología implementada. En primer lugar, cabe considerar que las evaluaciones fueron efectuadas por un sólo evaluador. En segundo, la participación del evaluador, como juez interno, en el procedimiento de validación de contenido permite descartar o revisar los reactivos que resultaban poco claros, disminuyendo así las posibilidades de interpretación errónea. En tercer lugar, la realización de una sesión de capacitación y la posterior observación del desempeño por parte del evaluador posiblemente reducen posiblemente los efectos de los sesgos propios de los evaluadores. Si bien los resultados siempre enmascaran un cierto efecto de halo u otro sesgo. De hecho, el mismo evaluador reconoce tener una baja habilidad para diferenciar entre el desempeño superior e inferior, lo que se traduciría en una tendencia a la medición central. Por último, si bien la longitud de la prueba asegura el cumplimiento de las normas de confiabilidad, aumenta también el tiempo requerido para completar el instrumento, lo cual es percibido negativamente por el evaluador. De esta manera, como no resulta idóneo aumentar el número de reactivos en las dimensiones con confiabilidad moderada, se podría eliminar los reactivos de las dimensiones de confiabilidad moderada que presentan una baja correlación con los demás reactivos de la dimensión o aumentar el número de reactivos en dichas dimensiones, reduciendo el número de reactivos en dimensiones con una muy alta confiabilidad. Lo anterior requiere de la previa aplicación de un análisis factorial, ya que no se debe aumentar la confiabilidad de las medidas a costa de la validez.

La discriminación de los reactivos, por su parte, no se obtiene dado el tamaño de la muestra. Al respecto, cabe considerar que las correlaciones tetrónicas entre el ítem y el puntaje total necesitan muestras de más de 100 sujetos. Se ha visto también que métodos más rudimentarios como la diferencia de cuartiles extremos son inadecuados, ya que no constituyen una prueba suficiente de discriminación.

Aceptabilidad

La evaluación del usuario indica claramente que el instrumento es aceptable para el evaluador. Al respecto, se debe subrayar la importancia del proceso de valoración y del proceso de diseño del mismo. Así, la entrega de feedback y la observación del desempeño se constituyen en otra pieza más del proceso de valoración. Del mismo modo, un rol activo del evaluador en el proceso de diseño contribuye a la aceptabilidad. Cabe resaltar que la imposibilidad de implementar por completo el proceso de valoración en cuanto a la entrega de feedback y análisis del desempeño y planificación de mejoras, determina que las respuestas reflejen las expectativas del evaluador más que juicios sobre eventos observados.

El constructo “Aceptabilidad” así resulta útil tanto para explorar las necesidades del usuario como su satisfacción respecto de la propuesta de evaluación de desempeño. Cabe considerar que no todas las dimensiones de aceptabilidad son relevantes. Al respecto, en la fase de diagnóstico, sólo la validez, practicalidad y libertad de influencia de sesgos son señaladas por los entrevistados mientras que, en la fase de evaluación, el evaluador señala la utilidad, validez así como la operatividad. En cuanto a esta última, si bien no es considerada una dimensión de aceptabilidad, el límite con la practicalidad es difuso. Cabe destacar que muchas de las especificaciones del proceso de valoración y diseño que son referidas en ambas fases se asocian a más de una dimensión de aceptabilidad. En relación con todo lo anterior, hay que considerar que la aceptabilidad como concepto no tiene un alto nivel de formalización.

La aceptabilidad del instrumento, en cierto modo, está garantizada por la consideración de las especificaciones que los usuarios se efectúan en la fase de diagnóstico. Así, si bien es complejo desarrollar medidas de evaluación que sean aceptables en función de la multiplicidad de aspectos a considerar, explorar estos aspectos como parte del diagnóstico es una herramienta sumamente útil.

Recomendaciones

Para Desarrollar Evaluaciones de Desempeño

En relación con el desarrollo de instrumentos y sistemas de evaluación, se puede tomar como referencia el proceso de diseño y aplicación seguido, superando algunas barreras que se presentan en este estudio. Al respecto, se propone:

1. Realizar un diagnóstico de necesidades que incluya un análisis del puesto de trabajo, el cual considere la aplicación de una entrevista de incidentes críticos a jefes y empleados. De esta manera, no sólo se identifican las dimensiones de desempeño a evaluar sino posibles indicadores y reactivos a emplear, asegurándose así no sólo la relevancia respecto del trabajo sino contribuyendo a la aceptabilidad del sistema. Cabe destacar, por otra parte, que se puede utilizar el formato de entrevista de diagnóstico usado en esta investigación, si las alternativas establecidas corresponden a las particularidades de la organización.

2. Si es posible tras la redacción de los reactivos, fijar un periodo breve de tiempo para que evaluadores y evaluados observen si las conductas contenidas en los reactivos se dan en el ejercicio del cargo. Estos, a su vez, pueden participar como expertos internos en el proceso de validación de contenido.

3. Capacitar a los mismos sobre el sistema de evaluación.

4. Explorar no sólo las propiedades psicométricas, sino la aceptabilidad que tiene el sistema de evaluación para los usuarios tanto evaluadores como evaluados. Opcionalmente, se puede estudiar la utilidad y practicalidad del sistema. Cabe destacar que la aceptabilidad se preocupa sólo de la percepción de estas propiedades más que la utilidad y practicalidad misma.

Para la Aplicación y Validación de la Propuesta de Evaluación

En cuanto a la aplicación del instrumento propuesto para evaluar el desempeño de personal de línea en otras organizaciones, en primer lugar, se requiere efectuar un diagnóstico de necesidades para asegurarse que las necesidades de la empresa y de los miembros concuerdan con las establecidas por este

estudio. En segundo lugar, también es importante verificar si la población de sujetos a evaluar tiene características similares a las de la población de esta investigación. Tercero, efectuar un estudio de validación en la primera aplicación, ya que los resultados del análisis de validez y confiabilidad son preliminares. Al respecto, se recomienda, en cuanto al estudio posterior de la confiabilidad, continuar usando del coeficiente Alfa de Cronbach, ya que el desempeño de los empleados varía en el tiempo en función de la entrega de feedback así como de la implementación de otras acciones de mejora que se deben contemplar como parte de la fase de simulación. Por otra parte, en cuanto al estudio de validez, sin lugar a dudas, éste debe contemplar la aplicación de un análisis factorial de tipo exploratorio no sólo para determinar los factores que explican la varianza, sino también para establecer ponderación de las dimensiones que sea más adecuada.

En cuanto a la etapa de simulación y/o posterior aplicación del sistema, se sugiere además seguir el mismo proceso aplicado en este estudio y adicionalmente: (a) entregar de feedback, (b) en relación con los evaluados, efectuar una sesión de capacitación y reuniones de información para evaluadores y evaluados (c) implementar otros procedimientos como la revisión de la evaluación ante discrepancias entre el evaluador y evaluado y (d) desarrollar un sistema formal de gestión de desempeño para los empleados con desempeño superior e inferior que homologue los criterios para el análisis del desempeño, toma de decisiones, implementación de acciones y evaluación de los progresos.

Cabe destacar, por último, si bien el instrumento se desarrolla para ser aplicado a todas las células de Operaciones de la empresa, es validado sólo en una la Célula Printing. Es importante acotar que como los procesos productivos varían de una célula a otra, específicamente van desde el acondicionamiento de materias primas, la elaboración de cigarrillos hasta la impresión de paquetes y envoltorios; el instrumento podría ser validado en otras industrias manufactureras, específicamente en poblaciones de operadores de maquinaria de producción y de impresión.

En cuanto a la empresa en la cual se ambienta este estudio, se sugiere que el sistema sea aplicado en todas las células de modo de efectuar un análisis de validez, confiabilidad y discriminación más concluyente, siguiendo las recomendaciones descritas. En cuanto a la implementación debe contar además con la realización de una sesión de capacitación para todos los evaluadores y evaluados así como con la participación de estos en la fase de evaluación. Se recomienda además aplicar el instrumento, atendiendo al proceso de valoración al cual se circunscribe. Al respecto, es relevante la presencia de un clima organizacional positivo, la realización de una observación prolongada del desempeño, la entrega de feedback sobre el desempeño a los evaluados, el análisis de los resultados de las evaluaciones así como la planificación e implementación de acciones de mejora. La evaluación de desempeño además se deberá insertar como una herramienta integrada de la Estrategia de Recursos Humanos, por lo tanto, se deberá asegurar su congruencia con las prácticas de recursos humanos; especialmente de selección, promoción y asignación de puestos de trabajo, distribución de recompensas y beneficios, desarrollo del talento así como del manejo de las relaciones y canales de comunicación.

Si se siguen estas recomendaciones, el instrumento permitiría equilibrar las necesidades del individuo y de la empresa. Así, por una parte, facilitaría a la organización optimizar su desempeño mediante la identificación de los comportamientos específicos que no se presentan con la frecuencia esperada así como las dimensiones del desempeño que representan oportunidades de mejora. Por otra parte, contribuiría a satisfacer las necesidades del individuo en relación con la necesidad de feedback y de justicia en la distribución de recompensas y castigos. De este modo, el mensaje que se enviaría estable que las contribuciones de los empleados al logro de objetivos y el cómo realizan éstas son importantes; es decir, el desempeño de cada individuo realmente importa y debe ser gestionado como un activo clave. En cierto grado, transformaría además la interacción entre el jefe de producción y su personal, ya que abre un canal de comunicación permanente sobre las expectativas que uno tiene en relación con el desempeño el otro.

Líneas de Investigación

Las líneas que abre esta investigación son principalmente tres. En primera instancia, futuros estudios deben centrarse en medir y mejorar la validez y confiabilidad del instrumento propuesto para evaluar el desempeño del personal de línea de otras industrias manufactureras, proceso que por lo demás nunca termina. En relación con lo anterior, estos estudios pueden centrarse además en determinar los sesgos del evaluador y de la organización que afectarían las evaluaciones obtenidas mediante el instrumento propuesto. En segunda instancia, la investigación puede abordar el proceso de diseño y de valoración en cuanto a su aporte a las propiedades psicométricas y de utilización de los instrumentos, lo cual orientaría el futuro desarrollo de propuestas de evaluación de desempeño. Finalmente, la línea de investigación más compleja se puede orientar a estudiar los determinantes y predictores del desempeño, ya sean factores internos como externos al individuo. Al respecto, el Modelo de Campbell es tentativo, pero representa un buen punto de partida. La última línea de investigación presenta un mayor valor social, ya que es la identificación de los determinantes y su posterior manejo, lo que nos permite mejorar el desempeño laboral.

Limitaciones y Fortalezas de la Investigación

A modo de Epílogo, en relación con las fortalezas de la investigación, cabe mencionar que la principal fortaleza radica en la metodología empleada, la cual se orienta a satisfacer las necesidades de los usuarios y los requisitos psicométricos. Esto constituye un avance frente a la investigación tradicional que se centra sólo en la validez, discriminación y confiabilidad y que no reconoce la importancia del rol de evaluadores y evaluados en el diseño de propuestas de evaluación. Al respecto, si bien la aceptabilidad es un constructo poco estudiado, la investigación propone una técnica útil para abordarla como criterio de diseño y evaluación.

En cuanto a las debilidades, en primer lugar, hay que señalar la falta de estudios en nuestro país en el área de evaluación de desempeño. En segundo lugar, si bien se siguieron todas las recomendaciones para controlar los sesgos de modo de garantizar la validez y confiabilidad de las medidas, se evalúa solamente la percepción de libertad de influencia de sesgos del evaluador. Queda así sin determinar en qué medida los resultados realmente presentan distorsiones debidas al desempeño del jefe de producción como evaluador y/o a condiciones organizacionales. En tercer lugar, es importante acotar que diversos factores situacionales de la misma empresa restringen finalmente los alcances de esta investigación. Al respecto, son relevantes la limitación de la participación de los evaluados en fase de análisis, simulación y evaluación así como el tamaño final de la muestra de sujetos evaluadores. Dado esto último, el instrumento se valida en sólo una célula así como no se realiza un análisis factorial exploratorio y un análisis de discriminación, los cuales hubieran permitido obtener un análisis más concluyente de las propiedades psicométricas de la propuesta de evaluación.