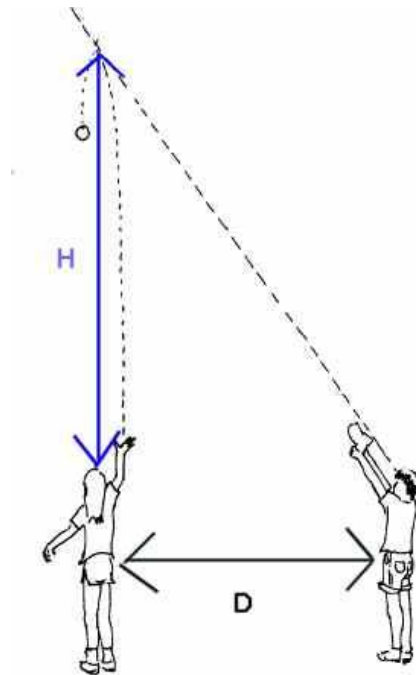


EVALUACIÓN DE LA INSTRUCCIÓN

TEORÍA DE LA EVALUACIÓN



Jorge Everardo Aguilar Morales



ASOCIACIÓN
OAXAQUEÑA DE
PSICOLOGÍA A.C.

Evaluación de la instrucción: Teoría de la evaluación.
Aguilar-Morales, Jorge Everardo
© 2008. Asociación Oaxaqueña de Psicología A.C.
Calzada Madero 1304, Centro, Oaxaca de Juárez, Oaxaca,
México. C.P. 68000
Tel. (951)5142063, (951) 5495923, Fax. (951) 5147646
www.conductitlan.net
E-mail: jorgeever@yahoo.com.mx, comentarios@conductitlan.net
Se promueve la reproducción parcial o total de este documento
citando la fuente y sin fines de lucro.

En caso de citar este documento por favor utiliza la
siguiente referencia:
Aguilar-Morales, J. E. (2006) Evaluación de la instrucción:
Teoría de la evaluación. México: Asociación Oaxaqueña de
Psicología A.C.

Propósito del curso

La eficacia de un sistema educativo solo se puede determinar cuando existen procedimientos sistemáticos y objetivos para evaluar los resultados obtenidos al final del proceso, su ausencia ocasiona que los juicios que se emitan sobre dicho sistema carezcan de importancia por no estar sustentados en datos confiables o que al carecer de validez se limite la posibilidad de tomar decisiones adecuadas sobre los cambios que se necesitan realizar para mejorarlo.

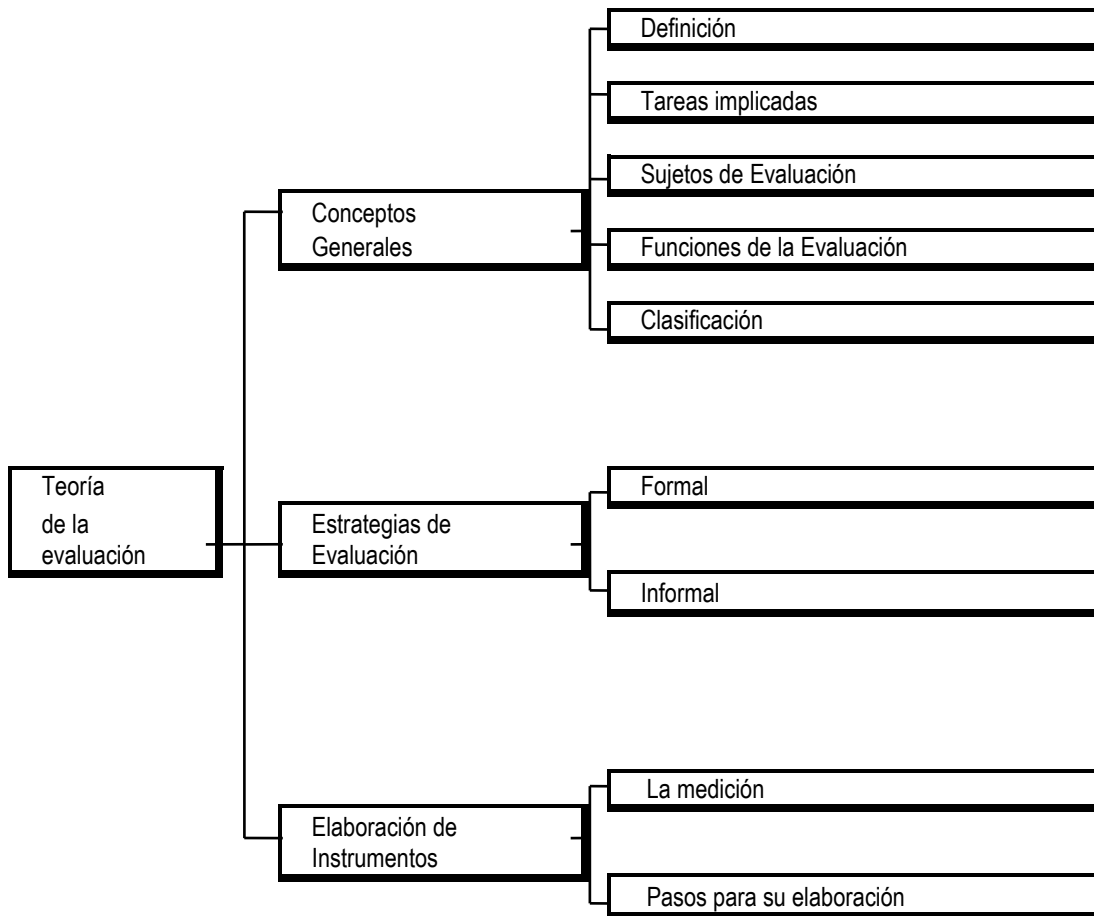
En este curso se proporcionarán los elementos necesarios para que los participantes puedan desarrollar y evaluar una estrategia de evaluación para un objetivo educativo determinado.

Para una mejor comprensión del curso se sugiere que los participantes hayan tomado con anterioridad el curso de Formación de Instructores y que tengan elaborada la carpeta didáctica de un curso, sin embargo, aún cuando no cubran con los requisitos mencionados, este curso le puede resultar de utilidad a todos aquellos que estén relacionados con procesos de evaluación.

ÍNDICE

CONTENIDO	PAG.
UNIDAD I: CONCEPTOS GENERALES DE LA EVALUACIÓN	7
UNIDAD II: ESTRATEGIAS DE EVALUACIÓN	14
UNIDAD III: ELABORACIÓN DE INSTRUMENTOS DE MEDICIÓN	30
BIBLIOGRAFÍA	46

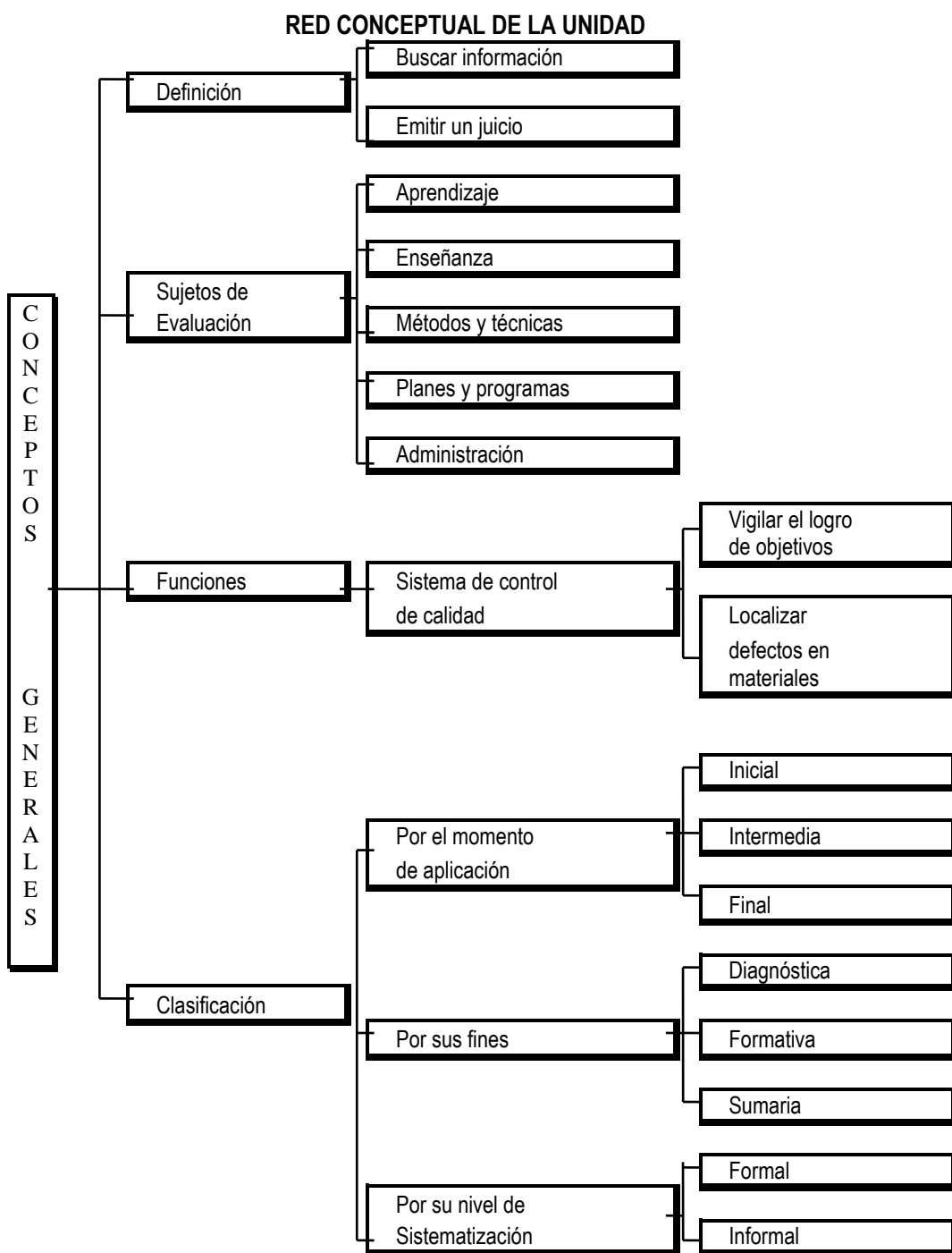
RED CONCEPTUAL DEL CURSO



OBJETIVO GENERAL DEL CURSO

Al término del curso los participantes serán capaces de diseñar y evaluar una estrategia de evaluación para un objetivo educativo determinado.

CONCEPTOS GENERALES DE LA EVALUACIÓN



Objetivo General de la Unidad

Al término de la presente unidad el estudiante será capaz de explicar a través de ejemplos la definición de evaluación, los sujetos de evaluación, sus funciones y su clasificación.

Conceptos Generales de la Evaluación

La definición más sencilla de evaluación es aquella que la define simplemente como la emisión de un juicio acerca de algo.

Definición → En este sentido la evaluación educativa se puede definir como la emisión de un juicio acerca del sistema educativo en su totalidad o de una unidad dentro de este sistema.

Cuando un maestro dice:

“Todos mis alumnos lograron los objetivos de aprendizaje” o

“El taller es una de las formas más efectivas de enseñar a redactar objetivos de aprendizaje”

Lo que está haciendo es emitir un juicio, está evaluando, en este caso, el aprendizaje del estudiante y las técnicas de enseñanza.

La forma en que el maestro llegó a este juicio varía según el tipo de evaluación que haya utilizado, en algunos casos el trabajo será mayor en otros bastarán solo las apreciaciones del maestro pero al final, lo que se busca al evaluar es emitir un juicio a partir del cual se tomen decisiones útiles para mejorar el sistema educativo.

Ahora bien la evaluación implicará siempre necesariamente tres tareas.

Tareas
implicadas en la
evaluación →

1. La búsqueda de información a partir de la cual se emitirá el juicio.
2. La comparación de esta información con un punto de referencia.
3. La emisión del juicio

Aplicar un examen o realizar preguntas a los alumnos durante la clase son formas de obtener información.

Por otra parte los elementos del sistema que pueden ser evaluados son:

Sujetos de
Evaluación →

1. El aprendizaje del estudiante.
2. La enseñanza del maestro.
3. Los métodos y técnicas empleados
4. Los Planes y programas de estudio
5. El subsistema administrativo.

Función
principal
→

La evaluación del sistema educativo o de sus elementos tiene una función principal que es la de “establecer un sistema de control de calidad”, en este sistema de control de calidad se encuentran dos aspectos importantes:

Funciones
importantes
→

1. Vigilar el logro de los objetivos instruccionales por parte de los estudiantes, a fin de determinar que instrucción correctiva o suplementaria se necesita.
2. Localizar defectos en los materiales de la lección y determinar las razones de esos defectos.

Proporcionar al alumno un nuevo material de estudio antes de que continúe con la siguiente unidad o acreditarlo, eliminar un ejercicio en una lección e implementar otro, son decisiones que ejemplifican las funciones de la evaluación.

Otras funciones adicionales son:

Funciones
Adicionales →

1. Facilitar el aprendizaje del estudiante
2. Juzgar el progreso del alumno
3. Tomar decisiones acerca de que métodos y materiales utilizar en la instrucción
4. Determinar la guía y orientación que requiere un estudiante.

La realización de evaluaciones constantes incrementa el aprendizaje del estudiante debido a que las personas aprenden aquello que se les pide que hagan, y las pruebas son una buena oportunidad para que el alumno ejecute lo que se le está enseñando. Para que las evaluaciones faciliten el aprendizaje deben cubrir con dos requisitos:

1. el estudiante debe poder contestar acertadamente un número razonable de problemas y
2. lo que se evalúe debe estar directamente relacionado con los objetivos

Clasificación

La evaluación es a menudo clasificada por el momento en que se realiza, por sus fines y por el grado de sistematización con el que se realiza.

Por el momento en que se realiza es clasificada como:

Inicial : La que se realiza al inicio de una lección o un curso.

Por el momento



Intermedia: La que se realiza durante la lección o el curso.

Final : la que se realiza al terminar la lección o el curso.

Por los fines con que se realiza es clasificada de la siguiente forma:

Diagnóstica: Es aquella que investiga los conocimientos que posee el estudiante al inicio de una experiencia de aprendizaje.

Formativa: Es aquella que se realiza durante la experiencia de aprendizaje con el objeto de realizar los ajustes necesarios en la lección.

Por sus fines



Sumaria: Es aquella en la que se evalúan los conocimientos adquiridos al terminar la experiencia de aprendizaje y busca determinar si se lograron los objetivos.

De acuerdo con el grado de sistematización con que se realiza se puede hablar de dos tipos de evaluación:

Por su nivel de sistematización
➔

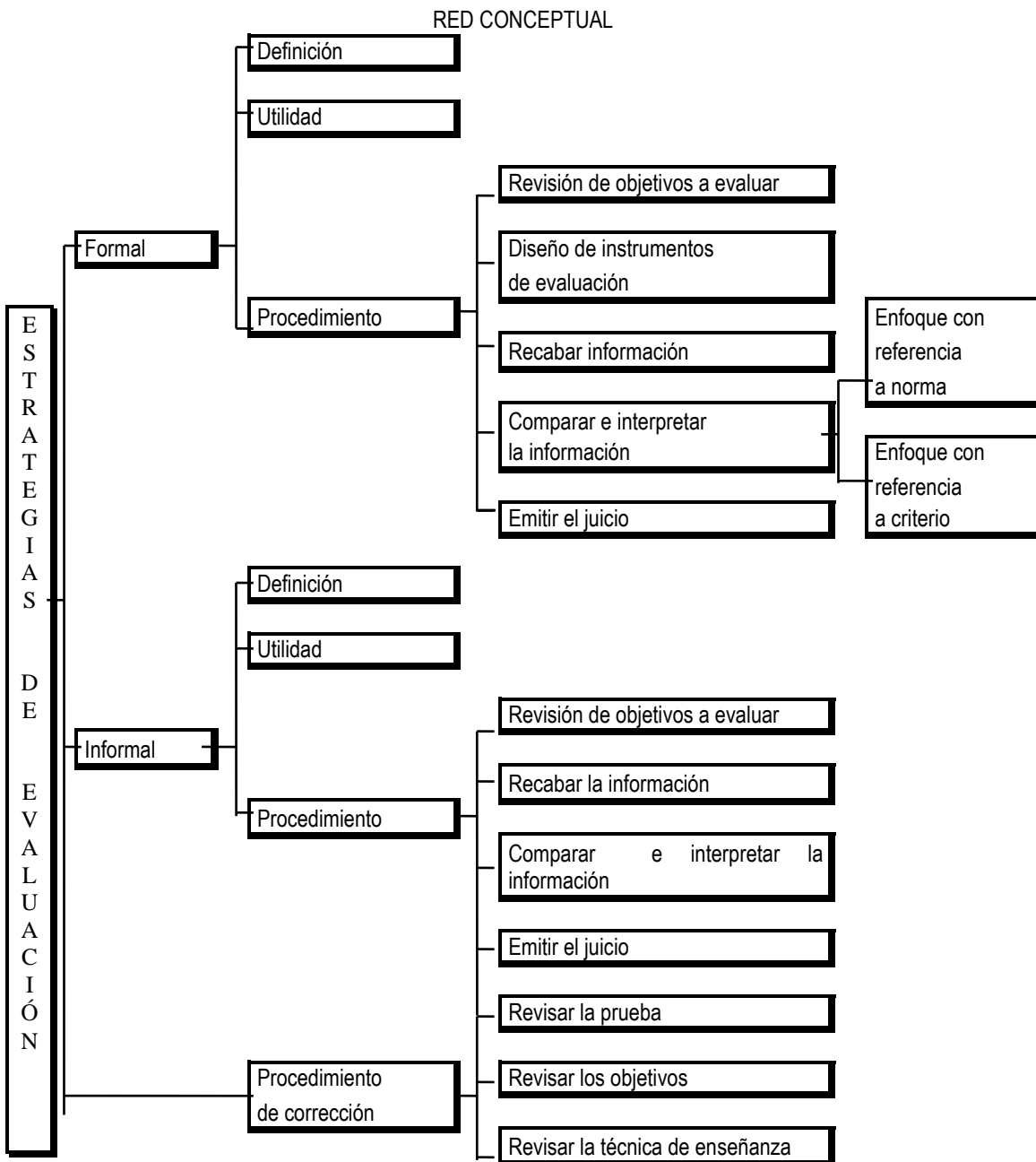
Formal: Que implica la utilización de instrumentos de medición confiables y el desarrollo de un procedimiento determinado.

Informal: En la que se recaba información sin que necesariamente se utilice instrumentos de medición confiables y sin que se siga un procedimiento determinado.

Evaluación

1. Explique la definición de evaluación educativa
2. Diga tres elementos que usted podría evaluar en su labor como instructor.
3. Ejemplifique como desarrolla o podría desarrollar en su labor educativa las tareas implicadas en toda evaluación.
4. Describa una situación en la que se observen las funciones de evaluación educativa.
5. Mencione la diferencia fundamental entre la evaluación formal y a la evaluación informal.

ESTRATEGIAS DE EVALUACIÓN



Objetivo General de la Unidad

Al término de la presente unidad el estudiante será capaz de seleccionar un procedimiento de evaluación apropiado para probar un objetivo educativo determinado.

ESTRATEGIAS DE EVALUACIÓN

Al final de la unidad anterior se planteó que existen dos formas de realizar la evaluación: la formal y la informal. En esta unidad hablaremos de las características de cada una de ellas cuando el elemento del sistema educativo que se pretende evaluar es el aprendizaje del estudiante.

EVALUACIÓN FORMAL

Definición → La evaluación formal busca emitir un juicio a cerca de determinados elementos del sistema educativo siguiendo un procedimiento ya definido y utilizando instrumentos de medición confiables.

Utilidad → Es recomendable utilizar la evaluación formal cuando existen probabilidades de que los alumnos ya posean los conocimientos y habilidades que se les van a enseñar y cuando se quiere realizar una Evaluación lo mas objetiva y confiable posible.

Para realizar una evaluación formal es necesario desarrollar las siguientes actividades:

Procedimiento de una evaluación formal →

1. Revisar los objetivos específicos que se van a evaluar.
2. Diseñar los instrumentos de medición apropiados para el objetivo planteado.
3. Recabar datos con los instrumentos diseñados
4. Comparar e Interpretar los datos
5. Emitir el juicio respectivo

Enseguida se analizan a detalle cada una de estas actividades.

1. Revisión de los objetivos a evaluar

Toda evaluación del aprendizaje del estudiante debe partir de los objetivos de aprendizaje que se establecieron con anterioridad.

Un maestro puede diseñar una prueba a partir del temario de un curso pero como no tiene claro lo que desea que el estudiante haga al final del curso, su prueba no evaluará la situación del estudiante con respecto a los objetivos que se persiguen. En consecuencia lo primero que tendrá que hacer será revisar los objetivos que pretende evaluar

Revisión de objetivos a evaluar →

Si los objetivos son específicos y utilizan verbos observables en su redacción será relativamente fácil establecer la forma en que se evaluará dicho objetivo.

Por ejemplo, si el objetivo que se pretende evaluar señala que:

Ejemplo →

“Al término de la lección el estudiante redactará un escrito en el que utilice las reglas revisadas en clase”

Resultaría claro que para evaluar este objetivo se tendría que solicitar al alumno la elaboración un escrito el cual debería ser revisado por el maestro para determinar si se utilizaron las reglas revisadas en el curso”

Veamos otro ejemplo:

“Al término de la lección el estudiante será capaz de señalar las tomas, los planos y los efectos sonoros de un comercial”

Ejemplo →

En este caso la evaluación consistiría en solicitar al estudiante que de forma verbal o escrita mencione las tomas, los planos y los efectos sonoros de un comercial.

Sin embargo, si el objetivo esta redactado en forma ambigua antes de determinar lo que se evaluará debe volver a redactarse el objetivo pero ahora utilizando términos operacionales o conductuales.

Ejemplo →

Por ejemplo en el caso del objetivo “Al término del curso el alumno habrá adquirido el gusto por la lectura” la forma de evaluar no es clara debido a que en el objetivo en realidad no se especifica la conducta meta. Redactarlo en términos operacionales o conductuales implicaría transformarlo de tal manera que se señale de forma específica la conducta del estudiante.

Redactado en términos conductuales u operacionales el objetivo podría ser de la siguiente forma “ Al término del curso el alumno realizará la lectura de un libro que el haya elegido con una extensión no menor de 100 páginas”

Con un objetivo redactado ya en términos conductuales la técnica de evaluación sería fácil de deducir, en este caso el maestro lo que tendría que hacer es diseñar un procedimiento para comprobar que el alumno realizó la lectura mensual de un libro con una extensión no menor de 100 páginas.

Objetivos a
evaluar en una
lección →

Las evaluaciones tienen como finalidad averiguar si los estudiantes han alcanzado los objetivos de la lección, las evaluaciones finales deben entonces obtener información sobre si los estudiantes lograron el objetivo general. Sin embargo, debido a que las lecciones no están formadas por un solo objetivo sino por muchos objetivos debe procurarse evaluar todos los objetivos involucrados en una lección.

Cuando se evalúan todos los objetivos de la lección se puede contar con información que le permita al maestro saber con exactitud en donde estuvo la falla, es decir, que hizo que el estudiante no alcanzara el objetivo general, para en consecuencia tomar las medidas correctivas pertinentes.

La razón de buscar obtener un diagnóstico tan específico es que en la mayoría de los casos el alumno no logra los objetivos finales debido a que no dominó una subhabilidad o subconcepto necesario para dominar la habilidad o el concepto final.

Validez de la evaluación →

Retomar los objetivos de aprendizaje como primer paso en la realización de una evaluación es lo que le dará validez a los resultados obtenidos. Si una prueba no se apoya en los objetivos de aprendizaje carecerán de validez los resultados que de ella se obtengan. Solo serán útiles los resultados de una evaluación cuando en realidad indiquen en que medida de obtuvieron los objetivos que el maestro planteo al inicio del curso.

Es conveniente recordar que la emisión de un juicio acerca de que tan efectiva resultó una experiencia de aprendizaje dependerá necesariamente de la comparación entre los resultados que el estudiante obtuvo en las pruebas finales y los objetivos de aprendizaje.

2. Diseño de instrumentos de medición

Tres tareas están implicadas en toda evaluación: buscar información sobre lo que se evaluará, comparar esta información con un punto de referencia y emitir un juicio a partir de dicha información. En la evaluación formal el buscar información significa utilizar instrumentos de medición para recabar la información.

En el capítulo dos y tres hablaremos de la medición y el diseño de instrumentos de medición respectivamente, por ahora solo diremos que significa medir y cuales son los instrumentos que con mayor frecuencia se utilizan.

Definición de medición →

Medir significa **asignar un número a objetos o acontecimientos de acuerdo a ciertas reglas**, la función principal de medir es **lograr que ante un mismo evento distintos observadores coincidan en sus observaciones.**

Instrumentos de evaluación →

Usualmente en la enseñanza se utilizan con mucha frecuencia dos instrumentos de medición:

1. **Las pruebas o exámenes de aprovechamiento:** en los que se le presenta al alumno un problema que tiene que resolver, de los resultados obtenidos se deducirá si logro los objetivos.

Registros conductuales o listas de verificación: Que no son mas que cuadros en los que por un lado se anota el criterio a evaluar y por el otro si se cumplió el criterio evaluar, estos instrumentos son recomendables para evaluar el aprendizaje complejo, pueden utilizarse durante la ejecución de la conducta por ejemplo al realizar una exposición o para evaluar el producto de la conducta del estudiante por ejemplo al evaluar la forma en que el estudiante construyo un mueble.

En el anexo se muestran ejemplos

Lo importante de los instrumentos de medición es desarrollarlos con la técnica adecuada para lograr que sean instrumentos válidos y confiables. Mas adelante explicaremos que significan estas características.

3. Recabar información

Momentos de la evaluación →

Cuando el educador ha revisado los objetivos que evaluará y ha diseñado instrumentos de medición válidos y confiables el siguiente paso es utilizar los instrumentos elaborados para obtener información de lo que sucede antes, durante y después de la instrucción.

Evaluación Diagnóstica →

Antes de la instrucción se debe realizar una evaluación diagnóstica que este relacionada con los objetivos de la enseñanza planeada.

Todas las lecciones suponen en el alumno un comportamiento inicial, es decir, suponen que el alumno domina ya ciertas habilidades y conceptos necesarios para que aproveche la lección, también suponen que el alumno no domina aquellas habilidades y conceptos que se le enseñarán.

Utilidad →

La evaluación diagnóstica permite al maestro eficientar su trabajo debido a que en los resultados que obtenga puede descubrir si a los alumnos les faltan ciertos prerrequisitos con los que supuso contarían o dominan ya las habilidades que se pretendían enseñarles, en consecuencia puede decidir modificar sus objetivos de aprendizaje para hacerlos más fáciles o más difíciles.

La evaluación diagnóstica permite, además, que el maestro establezca de manera rigurosa los niveles de conducta de los alumnos antes de comenzar con la enseñanza y pueda por lo tanto atribuir los próximos cambios de conducta de los alumnos directamente a los efectos de su enseñanza.

Evaluación formativa →

Cuando la lección se esta desarrollando es conveniente realizar evaluaciones constantes, el aprovechamiento de los estudiantes debe ser vigilado constantemente a fin de descubrir concepciones falsa y vacíos en la comprensión.

Los maestros tienen que decidir constantemente si deben continuar con el material nuevo o si es necesario que expliquen con mayor amplitud el concepto o habilidad que se esta aprendiendo, de ahí la importancia de evaluar constantemente pues la información que obtengan al realizar cada evaluación les permitirá tomar la decisión correcta.

La decisión final dependerá de si los estudiantes han dominado el concepto y para averiguarlo el maestro debe plantear preguntas, problemas y tareas. La Evaluación debe seguir a toda explicación, lectura, programa, película, ejercicio y actividad en el laboratorio.

Evaluación sumaria →

Finalmente, después de la enseñanza debe realizarse una Evaluación sumaria, la utilidad de esta evaluación consiste en proporcionar información que nos permitirá saber si se lograron los objetivos instruccionales.

Si los objetivos de instrucción se lograron el maestro podrá proporcionar con confianza al alumno el siguiente material de instrucción en caso contrario deberá de realizar la instrucción complementaria o correctiva.

La comparación entre los datos obtenidos en la evaluación diagnóstica y la evaluación final permitirán, como ya se indicó, que el maestro atribuya el éxito obtenido a la instrucción que él realizó. Sin embargo, es requisito para hacer esta comparación que las pruebas utilizadas comprueben todas el mismo tipo de conducta en el estudiante o que en la prueba diagnóstica se incluyan algunas de las mismas preguntas que el docente proyecta emplear.

4. Comparar e interpretar los datos

Para que un maestro pueda emitir un juicio acerca de los resultados que su alumno obtuvo necesita tener un punto de comparación. En el terreno educativo existen principalmente dos enfoques para realizar esta comparación.

Evaluación por norma →

Enfoque por **referencia a norma**: En este el aprovechamiento de cualquier estudiante se describe en relación con otros estudiantes. Decir que Juan obtuvo el tercer lugar en la clase de historia es un ejemplo de la aplicación de este enfoque.

El problema que existe con la utilización de este enfoque es que no señala con exactitud el número de objetivos que el alumno alcanzó, únicamente señala la posición del alumno con respecto al grupo.

Evaluación por criterio →

Enfoque por **referencia a criterio**: en este enfoque el aprovechamiento se describe en relación con una norma absoluta. Es decir, el punto de comparación son los objetivos instruccionales, por lo que se puede saber con exactitud cuales fueron los objetivos instruccionales que el alumno no alcanzó.

Este segundo enfoque es el más recomendable a utilizar pues permite realizar los ajustes al proceso para que se puedan alcanzar con mayor facilidad los objetivos

Después de comparar los resultados del alumno en las evaluaciones respectivas con los objetivos de aprendizaje, el maestro puede determinar no solo el aprendizaje del estudiante, puede incluso determinar que fue lo que estuvo mal.

De como determinar que es lo que fallo y corregirlo se hablará al final de este capítulo.

5. La emisión de un juicio

Emitir un juicio implica que el maestro señale entre otras cosas, si se lograron los objetivos o si la instrucción y los materiales utilizados fueron los adecuados. La emisión de un juicio es una tarea sencilla cuando la información con que se cuenta es confiable. Seguir las indicaciones señaladas anteriormente implicara que las decisiones que se tomen realmente ayuden a mejorar la instrucción y en consecuencia a favorecer el aprendizaje del estudiante.

EVALUACIÓN INFORMAL

Utilidad →

Existen situaciones en las que el procedimiento de evaluación formal no es necesario ni aconsejable. Cuando son pocas las probabilidades de que el estudiante posea los conocimientos y habilidades que se le enseñaran es recomendable utilizar la Evaluación Informal. La Evaluación informal también es recomendable durante el desarrollo de una lección para saber si se continua con la revisión de los temas que comprende o si es necesario realizar antes una nueva explicación y cuando se pretende obtener información cualitativa del proceso educativo.

Definición → La evaluación informal es aquella en la que se emite un juicio sin que necesariamente se haya recabado la información a través de instrumentos de medición confiables y sin seguir un procedimiento determinado. Las actividades que se realizan en una evaluación informal son:

Procedimiento →

1. Revisión de los objetivos
2. Recabar información
3. Comparar e interpretar la información
4. Emitir el juicio

1. Revisión de los objetivos

Para que la evaluación informal sea efectiva debe, al igual que en la evaluación formal, retomar los objetivos de aprendizaje, y a partir de ellos determinar lo que evaluará.

2. Recabar información

La diferencia esencial entre la evaluación formal e informal radica en que en la evaluación informal no es necesario utilizar instrumentos de medición confiables. Para recabar la información en una evaluación informal el maestro puede utilizar procedimientos sencillos.

Estrategias
informales de
Evaluar →

Plantear algunas preguntas sencillas relacionadas con el tema, promover la participación de los alumnos en discusiones o en conversaciones en clase en donde se comente lo que se estudiará son ejemplos de evaluaciones informales.

Al hacer evaluaciones de este tipo procure que participen la mayoría de los alumnos, el que uno de ellos pueda contestar correctamente no significa que todos puedan hacerlo, por lo tanto es recomendable que la muestra que se elija sea lo más numerosa posible. También es conveniente que prepare con anticipación material que pudiera utilizar por ejemplo: las preguntas que usted realizará y las respuestas que tendría que dar el alumno para demostrar que ha dominado el tema.

3. Comparar e interpretar la información

La comparación de la información que se obtenga al evaluar al alumno deberá, igual que en la evaluación formal, ser comparada con los objetivos de aprendizaje. Cabe mencionar que, sin embargo, en este caso ya no se basará en los datos sino la opinión del maestro, que puede muy fácilmente ser afectada por situaciones distintas a las que evalúa.

4. Emitir el juicio

Después de hacer la comparación se deberá juzgar si se cumplieron los objetivos para luego tomar las decisiones correspondientes. La evaluación informal es útil y complementaria de la evaluación formal, es menos exacta y confiable que esta última, pero ayuda a que el maestro tenga una visión rápida y aproximada de algo que para ser evaluado objetivamente requeriría de una inversión de tiempo y trabajo considerable.

CORRECCIÓN DE LA ENSEÑANZA A PARTIR DE LA EVALUACIÓN

Una de las funciones más importante de la evaluación es que permite mejorar de manera constante la enseñanza. Los logros de los estudiantes son un índice directo de la capacidad del educador para programar y poner en práctica la enseñanza. De tal forma que si los resultados son positivos el maestro puede atribuirse a si mismo el éxito obtenido, sin embargo, si los resultados de los alumnos son malos, en lugar de culparlos por su falta de interés, el maestro debe revisar la planeación y la ejecución que realizo de su enseñanza, pues en la mayoría de los casos de el dependen esos resultados.

Interpretación de los datos →

Si un grupo numeroso de estudiantes obtienen bajo aprovechamiento en una prueba para evaluar un concepto o habilidad determinada, esto significa que la instrucción que precedió a la prueba fue inadecuada. El maestro no solo debe volver a enseñar a los estudiantes que tuvieron un desempeño deficiente sino revisar la instrucción de manera que él enseñe de manera más efectiva al utilizar después con otros estudiantes.

Procedimiento de corrección

Revisión de la prueba →

Para revisar la instrucción lo primero que debe revisarse es la prueba con la que se determinaron los niveles de aprovechamiento de los estudiantes. Las actividades incluyen en principio descartar que la prueba contenga errores de construcción y corregirlos en caso necesario, enseguida determinar la razón del deficiente aprovechamiento por ejemplo:

Razones del deficiente aprovechamiento →

Los alumnos pueden haber confundido los términos o conceptos, les falta mencionar las características importantes, no fueron capaces de poner ejemplos, manejan de manera inadecuada los sinónimos y antónimos o símbolos del concepto o no pueden relacionar el concepto con otros.

De igual manera puede suceder que en la ejecución de habilidades los alumnos no cuenten con los requisitos necesarios para desarrollar un procedimiento, que desconocen las situaciones en las que se aplica y sus restricciones, que omitieron alguno de los pasos de los procedimientos o que ejecutaron pasos equivocados.

Revisión de los objetivos →

Tras identificar los errores se deben buscar las causas, la revisión de los objetivos constituye una tarea importante pues es necesario determinar si los objetivos no fueron demasiado ambiciosos.

Si el objetivo en realidad fue ambicioso el procedimiento de corrección consiste en proporcionar al estudiante los conocimientos y las habilidades previas, que son necesarias para lograr el objetivo deseado.

Cuando los objetivos son adecuados para el nivel del estudiante y cuando este posee los requisitos necesarios para desarrollarlos, se puede pensar que tal vez el defecto estuvo en las técnicas de enseñanza o en el material utilizado.

Revisión de la
técnica de
enseñanza →

Entonces hay que localizar el lugar en donde se suponía que el estudiante debería haber aprendido, si el estudiante aprendería de la lectura de varias páginas habrá que releer el contenido de la lectura, o si el conocimiento lo adquirirá tras una discusión grupal habrá que revisar lo que realmente sucedió. De lo que se trata es de encontrar cual fue la dificultad para así obtener la solución.

Errores en la
técnica de
enseñanza →

Lo que sucede con mucha frecuencia es que el concepto no aprendido ni siquiera ha sido enseñado o se le ha prestado solo una atención pasajera, también sucede que las sintaxis y el vocabulario de las explicaciones en una exposición fueron demasiado difíciles, que al estudiante se le hablo de un principio pero no se le dio la oportunidad de aplicarlo, que se esperaba que el estudiante demostrará sus conocimientos en condiciones muy diferentes a aquellas en la cual aprendió, que el estudiante no trabajo lo suficiente, no puso atención o no hizo completa su tarea.

Todos los problemas que se detecten tienen solución, pero antes de resolverse deben identificarse, ese es uno de los propósitos de la evaluación.

RESUMEN

1. Para realizar la evaluación de un objetivo educativo se puede optar por hacerlo de manera formal o informal.
2. La evaluación formal implica cinco actividades: revisar los objetivos a evaluar, diseñar los instrumentos de evaluación, recabar información, comparar e interpretar la información con los objetivos de aprendizaje y emitir un juicio.

3. La evaluación informal implica cuatro actividades: revisión de objetivos a evaluar, Recabar información, comparar e interpretar la información obtenida con los objetivos de aprendizaje y emitir un juicio

4. Los resultados obtenidos por los alumnos se pueden comparar con los resultados obtenidos por sus compañeros o con los objetivos de aprendizaje. Es decir, se puede tomar como referencia una norma o un criterio. En el ámbito educativo resulta mas conveniente utilizar un criterio.

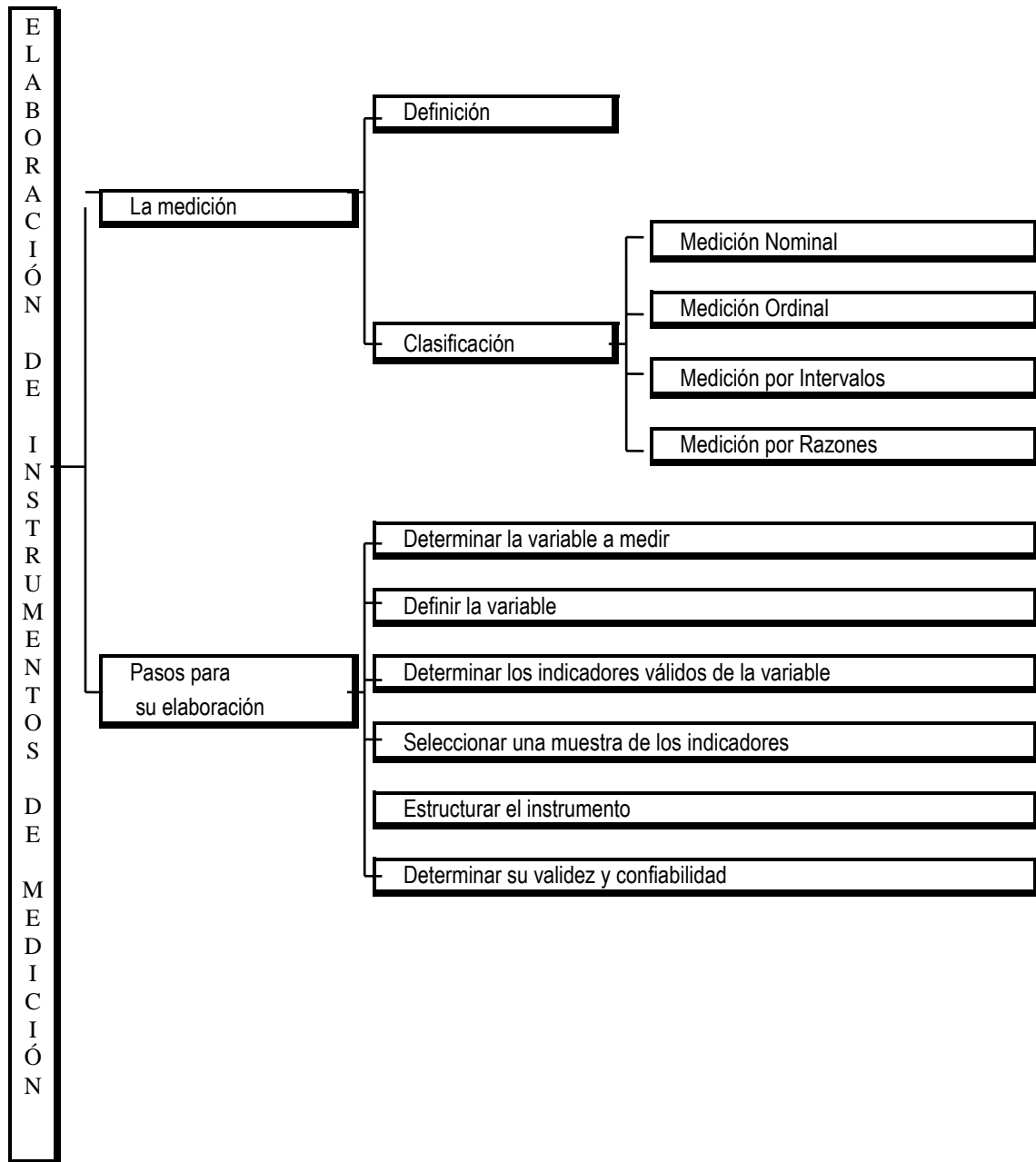
5. Cuando el alumno no logra los objetivos el error puede estar en la prueba, en los objetivos o en la técnica de enseñanza. Los tres elementos deben revisarse para detectar el error.

Evaluación

1. Ejemplifique las diferencias entre evaluación formal y evaluación informal, haciendo referencia a cada uno de los procedimientos que se deben de seguir.
2. Redacte un objetivo de aprendizaje de tal manera que se a útil para formular un instrumento de evaluación.
3. Explique cuales son los instrumentos principales en la evaluación educativa.
4. Que estrategias se den utilizar en la evaluación formal e informal para recabar información.
5. Explique los dos enfoques existentes para comparar e interpretar los datos al realizar evaluaciones.
6. Señale las razones por las ala evaluar a un estudiante sus resultados pueden ser negativos y que es lo que se tendría que corregir.

ELABORACIÓN DE INSTRUMENTOS DE MEDICIÓN

RED CONCEPTUAL



OBJETIVO GENERAL DE LA UNIDAD

Al término de la unidad el alumno será capaz de elaborar un instrumento de medición válido y confiable que le sirva para evaluar un objetivo educativo determinado.

ELABORACIÓN DE INSTRUMENTOS DE MEDICIÓN

En una evaluación objetiva o formal el diseño y la utilización de instrumentos de medición válidos y confiables es un aspecto crucial.

La utilización de este tipo de instrumentos es lo que permite que el maestro deje de realizar una tarea eminentemente filosófica y pase al terreno de la ciencia.

Los datos obtenidos por esta vía permiten tomar las decisiones correctas y realizar los ajustes necesarios para lograr los objetivos instruccionales.

LA MEDICIÓN

Definición y función de la medición →

Medir significa: Asignar un número a objetos o acontecimientos de acuerdo a ciertas reglas y su función principal es lograr que ante un mismo evento distintos observadores coincidan en sus observaciones.

Medir en el ámbito educativo permite que se cuente con una base objetiva, exacta y comunicable para describir, diferenciar y clasificar las características y conductas de los estudiantes.

La medición también permite cierta economía al manejar información y encontrar relaciones que de otra forma nunca se hubieran encontrado.

Aspectos medibles en la educación →

Podemos medir dos aspectos:

1. Las características de los estudiantes
2. Las conductas de los estudiantes

La estatura de una persona es un ejemplo de característica, en tanto que la respuesta a un test es una conducta, el CI de un estudiante es una característica mientras que sus respuestas a un test de inteligencia constituyen un ejemplo de una conducta. Debido a que tanto las características como las conductas pueden tomar distintos valores, se les denomina **variables**.

Definición de Variable →

Existen dos tipos de mediciones que se pueden realizar:

Tipos de mediciones →

1. Las mediciones cuantitativas
2. Las mediciones cualitativas

El puntaje CI de 135 es un ejemplo de medición cuantitativa, en cambio la categoría de "superior al término medio" es un ejemplo de medición cualitativa.

En general en el estudio de la conducta los científicos prefieren la medición cuantitativa a la medición cualitativa.

Anteriormente se mencionó que medir es asignar números a determinados objetos de acuerdo a ciertas reglas. La regla que se utilice también determina un tipo de medición.

1. Medición Nominal:

Definición →

Este tipo de medición es el nivel mas bajo de la medición consiste en poner nombres o rótulos a un conjunto de comportamientos o características. La regla que se utiliza es simplemente que un objeto es distinto de otro.

Cuando damos nombres para diferenciar una categoría de la otra pero sin establecer ningún orden entre ellas estamos usando el tipo de medición nominal.

Por ejemplo para designar las dos categorías que comprende la variable "sexo" debemos utilizar las denominaciones "Hombres" y "Mujeres". Las palabras "Hombres" y "Mujeres" son rótulos que se utilizan para designar dos categorías de la variable sexo y no implican que una de ellas sea mejor que la otra, tampoco tiene un orden implícito, es decir, solo estamos haciendo una medición nominal.

Entonces, la medición nominal es el proceso de diferenciar conductas y características dentro de las variables y de asignar a aquellas nombres o rótulos categoriales no ordenables.

Es importante mencionar que para designar las categorías de una variable se deben especificar las características de cada una de ellas y **las categorías deben ser mutuamente excluyentes**, es decir, que al clasificarlas características o conductas de las personas en las distintas categorías estos solo deben poder incluirse en una o en otra.

2. Medición Ordinal

Definición →

La **medición ordinal** se refiere al tipo de medición que hacemos cuando establecemos un orden en las categorías que comprende una variable. La regla que se utiliza es que los objetos tienen un determinado orden pero se encuentra a distancias desiguales.

Por ejemplo, las categorías “muy de acuerdo”, “de acuerdo” y “en desacuerdo” expresan una relación de orden en cuanto al grado de acuerdo de determinada cuestión.

Para que esta clasificación se vuelva ordinal y no solamente nominal se necesita, además, la utilización de números que nos indiquen el orden de las distintas categorías. Por ejemplo podríamos asignar un valor de 1 a la categoría “muy de acuerdo”, 2 a la categoría “de acuerdo” y 3 a la categoría “en desacuerdo”. Es decir, números que nos indicarán la dirección de la escala y el valor de cada categoría.

En las mediciones ordinales **debemos asegurarnos que las categorías en verdad indiquen un orden y que lo que se ordena pertenezca a aspectos de una misma variable.**

3. Medición Por Intervalos

La medición ordinal señala la dirección de determinadas categorías, pero las distancias entre esas categorías son desiguales. Una medición es más precisa si los puntos de la escala son equidistantes.

Definición →

La **medición por intervalos** se refiere al tipo de medición que hacemos cuando la distancia existente entre las categorías o puntos de una escala son iguales. La regla que en este caso se utiliza es que los objetos que se miden pertenecen a una misma categoría, se encuentran en determinado orden y a distancias iguales.

En una escala ordinal la distancia existente entre los puntos es desigual. Por ejemplo, el intervalo entre inferior y bueno, posiblemente no es igual al intervalo entre bueno y excelente. Por lo tanto aún cuando se designen números a los puntos de una escala ordinal estos no pueden sumarse debido a que los puntos no son equidistantes.

En una escala de intervalos los puntos están a la misma distancia por ejemplo la distancia entre los 20° y los 30° centígrados es igual a la distancia entre los 50° y los 60° centígrados.

Características →

Por lo tanto en una escala de intervalo los puntajes pueden ser aditivos, sustractivos, multiplicativos o divisibles.

Los test educativos que se preparan con una metodología rigurosa son ejemplos de escalas de intervalos y los puntajes que obtengan los alumnos podrán ser analizados de igual forma que los puntajes de una escala de intervalo.

4. Medición Por Razones

Definición →

La medición por razones constituye el cuarto y más alto nivel de medición. Esta escala contiene lo que se llama un punto cero absoluto, que es el punto en donde no existe puntaje. La regla en este caso señala que los objetos pertenecen a la misma categoría, que se encuentran ordenados, a una distancia igual uno de otro y que existe un cero absoluto.

Las escalas de razones existen en las ciencias físicas porque las variables tales como la altura el peso y tiempo tienen punto cero absoluto.

Puesto que la variable de peso tiene un punto cero absoluto es correcto pensar que una persona que pesa 80 kilos pesa el doble que una de 40 kilos. No sería correcto suponer que una persona cuyo CI es de 100, es el doble de inteligente que una persona cuyo CI es de 50. Suponer lo anterior equivaldría a suponer erróneamente que la variable inteligencia tiene un punto cero absoluto. Por lo tanto dos personas que tienen un CI de 50 no tienen en conjunto la misma capacidad intelectual que un individuo cuyo CI es de 100. Sin embargo, dos personas que pesan 50 kilos cada una, pesan en conjunto lo mismo que otra que pesa 100 kilos.

En los test educativos el punto cero es arbitrario, por lo tanto, tales test no dan puntajes de una escala de razones. **En el estudio de la conducta no se puede trabajar con variables con punto cero absoluto, en este terreno no es posible efectuar mediciones empleando las escalas de razones.**

PASOS PARA ELABORAR UN INSTRUMENTO DE MEDICIÓN

Para elaborar un instrumento de medición válido y confiable es recomendable seguir los siguientes pasos:

1. Determinar cual es la variable que se desea medir.

En este caso lo que se medirá dependera de lo que señalan los objetivos educativos determinados y se realizará con el procedimiento anteriormente señalado.

2. Definir la variable.

Se puede realizar una **definición conceptual** por ejemplo definir a la creatividad como la capacidad de dar respuestas únicas u originales y sobre esta definición hacer la consecuente medición, sin embargo esta definición resulta abstracta y dificultaría a dos personas ponerse de acuerdo sobre lo que se pretende medir. Por tal razón se prefieren las definiciones operacionales.

La **Definición Operacional** es una definición que determina las operaciones que comprende la medición de una variable. Ejemplo de Definición Operacional sería determinar diez maneras de usar ladrillos de vidrio como medida de la creatividad.

El segundo paso entonces para medir una variable es realizar una definición operacional de dicha variable.

3. Determinar indicadores válidos de la variable

Aquí de lo que se trata es de tomar algunas características o conductas específicas como indicadores de lo que deseamos medir.

Si lo que deseamos es medir la inteligencia primero tendríamos que dar una definición conceptual : “capacidad para aprender de la experiencia” enseguida deberíamos señalar que variables representan un “aprendizaje por medio de la experiencia”, la memoria y la capacidad para generalizar pudieran representar solo dos de tales variables. Pero aún estas dos variables tienen un nivel de abstracción alto así que tendríamos que deducir indicadores fácilmente observables y medibles.

La capacidad para repetir números de adelante para atrás y viceversa y la capacidad para dar información pudieran ser indicadores de la memoria, en cambio la capacidad para predecir las consecuencias de una acción inmediata pudiera ser un indicador de la capacidad para generalizar.

Cabe mencionar que la definición operacional simplifica las cosas porque ya indica desde la misma definición el indicador que se deberá tomar.

4. Seleccionar una muestra de esos indicadores.

Como los indicadores de una variable son numerosos y no se podrían incluir todos en un test, es necesario realizar un muestreo para seleccionar los indicadores más representativos, los cuales serán lo que se incluirán en el test.

5. Estructuración Del Instrumento

Al estructurar el instrumento lo mas importante es que se elabore un instructivo que permita que todas las veces que se aplique, se haga en las mismas condiciones, es decir que se busque estandarizar el procedimimiento.

6. Determinar la validez y confiabilidad del instrumento.

a) Validez de los tests.

La validez de un test se refiere al grado en que la prueba mide lo que pretende medir.

Definición
→

La **validez de contenido** de se refiere a la pertinencia y representatividad del contenido del test con respecto a la definición de la variable que se va a medir.

Por ejemplo un indicador válido o pertinente para medir inteligencia de acuerdo a lo revisado en el tema anterior pudiera ser la reproducción memorizada de dibujos, de la misma forma que la altura de una persona sería un indicador menos válido de inteligencia que su habilidad para resolver problemas de aritmética.

En lo referente a la representatividad revisemos este otro ejemplo, en un test de rendimiento diseñado para medir la aptitud para la lectura, el lenguaje y la aritmética se necesitaria una muestra que representara cada aptitud. Si tal test tuviera 100 ítems, 80 para aritmética y 20 para lectura y lenguaje, la muestra de estas aptitudes sería viciada. Para que fuera representativa la muestra debiera contener una proporción igual de cada aptitud.

La limitación indiscriminada de los indicadores de la variable reduce la validez de la medición, además las definiciones de las variables extremadamente amplias aumentan la dificultad de seleccionar los indicadores de la variable, de ahí la importancia de utilizar las definiciones operacionales.

Cuanto mas amplios son los objetivos educacionales, mayor será el número de indicadores incluidos en el test, de lo que se desprende que la pertinencia y representatividad del contenido deberán estimarse sobre la base de los objetivos educacionales.

El muestreo inicial de las conductas como indicadores se basa en un análisis lógico de la pertinencia de los indicadores con respecto a los objetivos educacionales. Tal análisis requiere:

1. El conocimiento de los objetivos educacionales y
2. Indicadores potenciales de estos objetivos.

Entonces, expertos en el tema juzgan generalmente los ítems a incluirse en los test de rendimiento. Su juicio establece la validez de contenido del test.

La validación por el consenso de expertos, es decir la validez de contenido, requiere un análisis interno del test. El consenso entre los expertos en el contenido es un proceso judicial y no es un procedimiento experimental. La validez de un test por medios experimentales se hace comparando los resultados que una persona obtiene en el test con un criterio que es externo al test mismo.

La determinación de la validez de un test usando un criterio externo constituye un proceso empírico y no intelectual. El término empírico se usa para aludir a procedimientos directos para probar la validez de un test.

Validez Con
corrente
→

La validez de un test que se basa en la comparación de los puntajes del test con criterios externo concurrentes, es decir que se obtienen al mismo tiempo se llama **validez concurrente**.

La validez concurrente de un test es un proceso empírico que consiste en comparar los puntajes que se obtiene en un test con los puntajes que se obtienen al mismo tiempo mediante el juicio de expertos o mediante un test ya validado anteriormente.

Si los puntajes entre las dos medidas correlacionan se habla de que el test tiene una validez concurrente, si no el test carece de este criterio de validez y deberá ser modificado.

Fuentes de Invalidez



Hay varios factores que pueden disminuir la validez concurrente de un test dos de los mas importantes son:

1. El grado de concordancia entre los contenidos del test y los objetivos educacionales del maestro.
2. El tiempo transcurrido entre las dos mediciones.

Por lo que es recomendable:

1. Realizar una validez de contenido de los dos criterios con que se realizará la medición y
2. Procurar que la dos medidas se tomen al mismo tiempo.

Es indudable que el juicio del maestro variara cada vez que haga una evaluación de un criterio a diferencia de las pruebas estandarizadas que permitiran obtener información de manera mas eficiente y sin variaciones atribuidas a quien mide.

A diferencia de la validez concurrente, la validez predictiva toma como criterio externo una medida que se obtendrá en el futuro y compara los resultados que un individuo obtiene en el momento actual con los que obtiene despues de cierto tiempo.

Si los resultados obtenidos en el momento actual correlacionan en un alto grado con los resultados que se obtendrán en el futuro se habla que el test tiene validez predictiva y que puede considerarse como un instrumento útil para predecir el éxito futuro de una persona.

Las operaciones para establecer la validez concurrente y la validez predictiva son esencialmente las mismas, la diferencia reside en el intervalo entre la administración del test al individuo y la medición de su posición con respecto al criterio externo.

Son ejemplos de pruebas con validez predictivas aquellos inventarios de intereses vocacionales utilizados para seleccionar estudiantes o trabajadores.

Fuentes de Invalidez



Por último es necesario mencionar dos fuentes de invalidez muy comunes en las pruebas educativas:

1. El que la estructura de la prueba permita al estudiante contestar correctamente sin tener el conocimiento y habilidad o
2. El que la estructura de la prueba haga que el estudiante se equivoque aun cuando haya dominado la habilidad o el conocimiento.

Una revisión de los reactivos en donde mas se equivocan los estudiantes permitirá disminuir estas fuentes de error y en consecuencia aumentar la validez de la prueba.

b) Confiabilidad del instrumento

Definición



La confiabilidad de un instrumento se refiere al grado en que un individuo obtiene un puntaje en una prueba que se apege con exactitud a la verdadera situación de la persona con respecto a la variable que se mide y si la información que se proporciona sobre esa situación es coherente.

Cuando se aplica una prueba los resultados pueden depender de la posición de la persona con respecto a la variable que se mide, pero pueden depender tambien de otras situaciones ajenas a la variable.

Por ejemplo puede suceder que el puntaje que obtenga un estudiante en una prueba sea afectado por las condiciones del aula, las instrucciones que se le dieron, el estado emotivo en el que se encontraba, el material en el que se imprimió la prueba, etc. Si este fuera el caso la prueba no sería confiable y por lo tanto sus resultado carecería de utilidad.

Una prueba es confiable cuando los resultados del estudiante dependen de su posición con respecto a la variable que se mide y no de otros factores.

En la medida que los resultados de un estudiante ante distintos reactivos que evalúan una misma habilidad sean congruentes entre sí y con criterios externos, en esa medida se estará hablando de confiabilidad del instrumento.

La confiabilidad de una prueba puede obtenerse a través de distintos procedimientos empíricos, por ejemplo:

1. comparando los puntajes obtenidos por un mismo estudiante en una prueba que se aplica en dos momentos diferentes,
2. comparando los resultados que obtiene el estudiante en dos pruebas distintas pero equivalentes
3. comparando los resultados que en un estudiante obtiene en la medida realizada por dos observadores independientes pero que utilizan un mismo instrumentos o
4. comparando los resultados obtenidos por un estudiante en cada una de las dos mitades de una misma prueba.

Estos procedimientos requieren de la utilización de técnicas estadísticas, sobre todo de la utilización de un coeficiente de correlación .

Sin embargo existen dos procedimientos sencillos de utilizar para determinar la confiabilidad de un instrumento de medición.

Si se trata de una prueba objetiva, el procedimiento consiste en determinar el porcentaje promedio de aciertos de cada objetivo para luego revisar y modificar aquellos en los que los promedios se desvían considerablemente del resto de los promedios.

En cambio si se trata de un registro conductual o una lista de verificación el procedimiento consiste en obtener dos medidas de dos observadores independientes que utilicen el mismo instrumento y luego utilizar esta formula:

$$\text{Confiabilidad} = \frac{\text{Número de acuerdos}}{(\text{Número de acuerdos} + \text{Número de desacuerdos})} * 100$$

Si los resultados son superiores a 79% se puede decir que el instrumento es confiable, en caso contrario sería conveniente realizar una modificación en cada uno de los rasgos que se evalúan, procurando definir en términos conductuales cada uno de los reactivos.

Una estrategia que aumenta la confiabilidad de los instrumentos, consiste en realizar su aplicación de la forma más estandarizada posible, es decir buscando que siempre se haga de la misma forma.

RESUMEN

1. Medir es asignar números a objetos o acontecimientos de acuerdo a ciertas reglas.
2. La medición permite lograr que ante un mismo evento distintos observadores coincidan en sus observaciones.
3. Existen cuatro tipos de mediciones de acuerdo a la regla que se utilice : medición nominal, ordinal, por intervalos y por razones.
4. En el estudio de la conducta no se pueden realizar mediciones utilizando una escala de razones.
5. Para elaborar un instrumento de medición existen seis pasos: Determinar cual es la variable a medir, definir la variable, determinar indicadores de la variable, seleccionar una muestra de esos indicadores, estructurar el instrumento y finalmente determinar su validez y confiabilidad.

6. Es importante que un instrumento de medición se aplique siempre en las mismas condiciones.
7. La validez de un instrumento se refiere a que este mida lo que realmente se quiere medir.
8. La validez de contenido consiste en hacer un análisis lógico entre lo que mide el instrumento y lo que se pretende medir.
9. La confiabilidad de un instrumento se refiere a que este obtenga datos que realmente correspondan al fenómeno estudiado y no a las características del instrumento.
10. En examen la forma más sencilla de determinar su confiabilidad es observar los porcentajes de aciertos en cada reactivo. Aquellos reactivos en donde el porcentaje de acierto sea muy bajo deberán ser modificados.

EVALUACIÓN

1. Explique la definición de medir
2. Ejemplifique los dos aspectos que se pueden medir en la educación
3. Ejemplifique cada una de las escalas de medición
4. Explique cuáles son los instrumentos para elaborar un instrumento de medición.
5. Señale la diferencia entre definición conceptual y operacional y ponga un ejemplo.
6. Explique el concepto de validez y confiabilidad de un test
7. Explique la fórmula para obtener la confiabilidad de un test.
8. Elabore un instrumento de evaluación para un objetivo educativo determinado.

NOTA: El contenido de este capítulo fue tomado de: Introducción a la medición en psicología y educación. Manual autoprogramado de Samuel Levien y Freeman F. Elzey.

BIBLIOGRAFÍA

SOBRE CONCEPTOS GENERALES DE LA EVALUACIÓN

- Enfoques y principios teóricos de la evaluación. Fernando Huerta Carreño. Editorial Trillas.
- Instrumentos de medición del rendimiento escolar, Fernando Carreño Huerta. Editorial Trillas.
- Didáctica General. Asociación Nacional de Universidades e Instituciones de Educación Superior.

SOBRE ESTRATEGIAS DE EVALUACIÓN

- Planeamiento de la enseñanza. W.J. Popham y E.L. Baker. Editorial Paidós.
- Psicología Educativa. Richard C. Anderson y Gerald W. Faust. Editorial Trillas.

SOBRE MEDICIÓN

- Psicometría Aplicada. Robert L. Thorndike. Editorial Limusa
- Introducción a la medición en psicología y en educación. Samuel Levine y Freeman F. Elzey. Editorial Paidós.
- Teoría Psicométrica. Jum C. Nunnally. Editorial Trillas.
- La forma del pensamiento matemático. Antología de James R. Newman. Editorial Grijalbo