



Confiabilidad y validez

Tomado de http://www.scielo.org/ve/scielo.php?script=sci_arttext&pid=S0798-97922002000200006&lng=en&nrm=iso&ignore=.html

La calidad de la evaluación depende sobre todo de cómo se aborde la validez y la confiabilidad, “cualidades esenciales que deben estar presentes en el desarrollo del proceso de recoger y analizar la información conducente a garantizar una mayor confianza sobre las conclusiones emitidas, de manera individual y compartida, por el evaluador”.

Estas dos cualidades se abordan desde diferentes perspectivas, siendo la cualitativa la que mejor explica la evaluación en educación

Validez Cuantitativa

La validez cuantitativa, bajo la perspectiva de la teoría positivista, se orienta fundamentalmente hacia las técnicas e instrumentos de medición elaborados por el evaluador, supuestos desarrollados a partir de las teorías planteadas por el filósofo francés Auguste Comte (1798-1857). **En esa perspectiva, la validez se concibe en términos de: (a) precisar lo que se pretende medir, (b) medir lo que se dice medir de cierto rasgo (Comte, 1896).**

Este tipo de evaluación presenta diferentes interpretaciones de validez, referidas a pruebas o tests:

a. Validez de contenido: juicio lógico sobre la correspondencia que existe entre el rasgo del objeto evaluado y lo que se incluye en la prueba, recurriendo a expertos para valorar la adecuación de cada ítem al rasgo a evaluar.

b. Validez de criterio: eficacia de la prueba para comparar un rasgo en cuestión con alguna o algunas variables externas, midiéndola a través de análisis correlacional o de regresión entre las puntuaciones obtenidas en la prueba y en otras basadas en el criterio.

c. Validez de constructo: basada en la integración de cualquier evidencia que fundamenta la interpretación o significado de las puntuaciones de la prueba o test, medida a través del análisis correlacional y de covarianza interítem.

Ahora, al tomar en consideración la pertinencia, el significado y la utilidad de las inferencias específicas derivadas de las pruebas o tests, conjuntamente con las consecuencias sociales y personales, y los valores, se pudiera pensar más bien en unificar las tres categorías de validez en una sola: la de constructo; debido a que las otras dos categorías también contribuyen al significado de las puntuaciones de la prueba o test, reconociéndolas así como aspectos de la validez de constructo (Messick, 1989: 20). De esta manera, la validez de constructo estaría conformada por tres componentes: (a) componente substantivo (confrontación entre la relevancia del juicio de contenido y la representatividad, por un lado, y la consistencia de la respuesta empírica, por el otro); (b) componente estructural (conjugación de la fidelidad del modelo para calificar las características estructurales de las manifestaciones del constructo y el grado de la estructura interítem); y (c) componente externo (grado de la relación empírica entre la prueba o



test con otras pruebas que reflejen las esperadas relaciones implícitas en la teoría del constructo que se evalúa).

La validez de constructo, sin embargo, debe tomar en consideración el uso contextual de la prueba o test y cómo se generan las puntuaciones de la misma. Cronbach (1988) señala que "los argumentos para usar la prueba deben estar claros para extender, de manera persuasiva y posible, la construcción de la realidad y el peso valorativo implícito, y sus aplicaciones" (p. 7). Para recoger información de la validez de constructo, necesaria para justificar la interpretación de la prueba, los datos deben surgir, inicialmente, de la administración y desarrollo de la prueba, para continuar con aquéllos obtenidos más allá del análisis de las puntuaciones, estableciendo así un proceso de búsqueda amplia y minuciosa.

Por su parte, las puntuaciones de la prueba pudieran llevar a emitir juicios erróneos del evaluador sobre el evaluado, al no considerar otros aspectos de tipo humano, social o cultural referidos al evaluado y, además, la información susceptible de ser medida limita su existencia, dejando afuera otros aspectos que pudieran ser de significativa importancia para validar el juicio. Al respecto Messick (1989) señala lo siguiente:

☐ *“los tests no sólo son imprecisos o falibles en virtud de los errores muestrales de medida sino también inevitablemente imperfectos como ideales del constructo que se proponen valorar. Los tests son medidas imperfectas de constructos porque no consideran aspectos que pudieran incluir de acuerdo a la teoría del constructo o incluyen aspectos que pudieran quedar fuera, o ambos. Esto último indica que hay dos tipos de evidencias necesarias para validar el constructo, uno para valorar el grado con el cual las implicaciones del constructo se evidencian con las relaciones empíricas de la puntuación y el otro para argumentar que estas relaciones no son atribuidas a otros constructos diferentes y alternos”* (p. 34).

Por su parte, Campbell y Stanley (1972) distinguen dos tipos de consideraciones para la validez cuantitativa asociada a la evaluación del aprendizaje: la validez interna y la externa.

a. Validez interna: concebida como la mínima imprescindible sin la cual es imposible interpretar alguna evaluación. Aquí es preciso considerar variables de tipo externo, como por ejemplo:

- acontecimientos específicos ocurridos entre mediciones sucesivas
- procesos internos del evaluado como cansancio, desmotivación, desinterés, etcétera
- influjo que la administración de una prueba o test ejerce sobre los resultados de otro posterior
- cambios en los instrumentos de medición o en los evaluadores que pueden producir variaciones en las mediciones que se obtengan

b. Validez externa: la cual plantea la interrogante sobre la posibilidad de generalización. Los factores que amenazan esta validez son: (a) el efecto reactivo o de interacción de las pruebas o tests (aumento o disminución de la sensibilidad o la calidad de la reacción del evaluado al estímulo de un aprendizaje); (b) el efecto de interacción (sesgo de selección de técnicas e instrumentos de evaluación y el tipo específico de aprendizaje); (c) efectos reactivos de la evaluación



(impedimentos para hacer extensiva las técnicas e instrumentos de evaluación); y (d) las interferencias de las evaluaciones múltiples (efectos que se producen cuando al evaluado le aplican evaluaciones múltiples y sucesivas, persistiendo aquéllos de procesos y resultados de evaluaciones anteriores).

Confiabilidad Cuantitativa

La concepción de confiabilidad de la evaluación, bajo el enfoque cuantitativo, hace referencia a la estabilidad o consistencia interna en las técnicas e instrumentos; y reside en establecer la medida en que se puede replicar la evaluación: una exigencia al evaluador para que, utilizando otros métodos y estrategias, llegue a idénticos resultados. La prueba y el test alcanzan un elevado coeficiente de confiabilidad si los errores de medida se minimizan; es decir, cuando toda medida de confiabilidad indica qué proporción de la varianza total de las puntuaciones es varianza de error.

Existen tres tipos de confiabilidad en la evaluación cuantitativa:

- la repetición de la prueba con tiempo suficiente para que el evaluado olvide los ítems, calculando el coeficiente de correlación que permite conocer el grado de confiabilidad de dicha prueba;
- las formas paralelas, técnica similar a la anterior pero con pruebas no idénticas sino con ítems equivalentes; y
- división en mitades, con el cual se calcula un coeficiente de consistencia interna, mediante la correlación entre las puntuaciones de ambas mitades.

Validez Cualitativa

La validez, bajo el enfoque cualitativo, concierne a la exactitud con que las conclusiones representen efectivamente la realidad empírica y si los constructos diseñados representan categorías reales de la experiencia humana (Hansen, citado en Pérez, 1994: 80).

La credibilidad, asociada a la validez, está condicionada a la estrecha relación del evaluador con el evaluado y con el contexto sociocultural donde se realiza la evaluación, tomando en consideración sus características más relevantes.

La validez depende, a su vez, de una actitud de aproximación de igual a igual, entre el evaluador y el evaluado, revisando permanentemente las conjeturas planteadas para que resulte en todos los casos, y desarrollando conjuntamente un proceso de intersubjetividad progresiva sobre el propio proceso de construcción.

Tomando en consideración el enfoque cualitativo, existen las siguientes estrategias para asegurar la validez en materia de evaluación:



a. Triangulación: consiste en recoger información desde diferentes puntos de vista, realizando comparaciones múltiples de una misma evaluación y combinando metodologías en su análisis; pudiendo existir, según Denzin (1979: 140), las siguientes modalidades:

- triangulación de tiempo (información recogida en diferentes momentos para su contrastación)
- triangulación de espacio (técnicas transculturales)
- niveles combinados de triangulación (análisis del nivel individual, del nivel interactivo o grupal, y del nivel de colectividades, organizacional, cultural o social)
- triangulación teórica (empleo de diferentes perspectivas coincidentes en una misma evaluación)
- triangulación del evaluador (verificación intersubjetiva entre varios evaluadores, a través del contraste de la información)
- triangulación metodológica (colección de técnicas o métodos diferentes, utilizados sobre el evaluado).

b. Saturación: Según Hopkins (1985), ella consiste en reunir las pruebas y evidencias suficientes para garantizar la credibilidad de la evaluación; lo cual se logra repitiendo la evaluación o revisando el proceso involucrado.

c. Validez respondiente o negociación: se desarrolla contrastando los resultados obtenidos por el evaluador con los del evaluado, así como contrastando su interpretación.

Goetz y LeCompte (1988), por su parte, plantean dos tipos de validez asociados a la evaluación:

a. Validez interna: consiste en conocer si el evaluador observa realmente lo que cree observar. Las variables que amenazan esta validez serían las mismas descritas por Campbell y Stanley (1979), sólo que su tratamiento es diferente debido a que la evaluación se realiza en escenarios no sujetos a mediciones cuantitativas ni a esquemas rígidos preestablecidos. En el caso de la historia y la maduración, el evaluador debe determinar la información que se mantiene estable en el tiempo así como la mutable. Para conocer los cambios, recurre a la replicación y comparación sistemática de la evaluación. Sobre la influencia del evaluador, éste debe controlar los factores personales en el proceso de recoger la información, la cual debe ser lo menos visible posible, sin prejuicios ni posturas preestablecidas. La selección y regresión las controla el evaluador a través de un inventario de subgrupos, acontecimientos y escenarios.

b. Validez externa: está referida a constatar si los constructos y postulados abstractos creados, perfeccionados o comprobados por el evaluador son aplicables a los del evaluado. Ella depende de la identificación y descripción de las características más importantes del evaluador y de las evaluaciones, de tal manera que puedan ser comparadas con otras posteriores de tipo similar, para lo cual quizás sea necesario conocer exhaustivamente sus atributos. Para contrarrestar esto, es recomendable desarrollar la evaluación, si es posible, en múltiples escenarios. El efecto de selección se minimiza determinando el grado de ajuste entre las categorías diseñadas y la realidad del evaluador y la del objeto evaluado. El efecto de escenario y de la historia, como amenazas al



efectuar interpretaciones intergrupales de constructos, disminuyen en la medida que la dinámica interactiva y las experiencias históricas se identifican con la mayor claridad.

Confiabilidad Cualitativa

La confiabilidad, bajo la perspectiva cualitativa, es un proceso empírico para verificar el grado de similitud entre el contexto del evaluador y el contexto de la institución educativa evaluada. Es relativa y depende totalmente del grado de superposición o apareamiento de sus características más sobresalientes. A su vez, depende enteramente de la descripción detallada que realice el evaluador sobre el tiempo, lugar y cultura donde ocurre la evaluación, para poder así transferir juicios hacia otros evaluadores.

Según Goetz y LeCompte (1988), existen dos tipos de confiabilidad para la evaluación:

a. Confiabilidad interna: consiste en conocer la congruencia existente entre las inferencias relativas al objeto de evaluación, derivadas éstas de los constructos elaborados por el evaluado sobre el objeto, para lo cual es fundamental la descripción y composición de los acontecimientos sin tomar en cuenta su frecuencia. Usualmente se utilizan las siguientes técnicas e instrumentos:

- descriptores de bajo nivel inferencial (narraciones y relatos concretos, minuciosos y precisos, incluyendo datos discrepantes)
- varios evaluadores (equipos de evaluadores abocados a la evaluación de un mismo aspecto)
- revisión por otros evaluadores (corroboración de los hallazgos por parte de otros evaluadores)
- datos registrados automáticamente (registros de video, audio, fotografías, etc.)

b. Confiabilidad externa: consiste en establecer condiciones empíricas en el proceso de evaluación que puedan ser replicadas por otros evaluadores, no significando esto que los hallazgos sean replicables, debido a que el flujo de información depende del contexto sociocultural donde ocurre la evaluación. El evaluador aumenta la confiabilidad externa de la evaluación dando respuesta a cuatro aspectos:

- status de la institución evaluada
- selección de la evaluación (caracterización de las técnicas e instrumentos que permitan la replicabilidad de la evaluación)
- situaciones y condiciones sociales (contexto físico, social e interpersonal donde se recoge la información relativa a la evaluación)
- premisas y constructos analíticos (especificación adecuada de los constructos, conceptos y definiciones que caracterizan la evaluación)
- métodos y estrategias de recogida y análisis de información (presentación diáfana de la estrategia y metodología de evaluación).



Confiabilidad y validez cualitativas

Tomado de <http://www.ucv.ve/uploads/media/Hidalgo2005.pdf>

Confiabilidad Cualitativa

La confiabilidad depende de procedimientos de observación para describir detalladamente lo que esta ocurriendo en un contexto determinado, tomando en cuenta para ello el tiempo, lugar y contexto objeto de investigación o evaluación, para poder así intercambiar juicios con otros observadores sean estos investigadores o evaluadores.

De allí que la confiabilidad representa el grado de similitud de las respuestas observadas entre el contexto del investigador o evaluador y el investigado o evaluado.

Kirk & Millar (1988), recomiendan a los investigadores cualitativos ir hacia la búsqueda de la consistencia de los hallazgos tomando como base los cuatro procesos de investigación etnográfica: invención, descubrimiento, interpretación y documentación, con la finalidad de poder coordinar la toma de decisiones.

Autores como Goetz y LeCompte (1988), señalan que la confiabilidad representa el nivel de concordancia interpretativa entre diferentes observaciones, evaluadores o jueces del mismo fenómeno. Para estos autores la confiabilidad de una investigación etnográfica depende de la solución a sus problemas de diseño interno y externo. En este sentido, establecen para la evaluación dos tipos de confiabilidad que reconocen como: confiabilidad interna y confiabilidad externa.

Confiabilidad interna

Este tipo de confiabilidad se evidencia cuando varios investigadores, estudiando la misma situación, concuerdan en sus conclusiones. El nivel de consenso entre diferentes observadores de la misma realidad eleva la credibilidad que merecen las estructuras significativas descubiertas en un determinado ambiente, así como la seguridad de que el nivel de congruencia de los fenómenos en estudio es consistente.

Los etnógrafos suelen utilizar varias estrategias para reducir las amenazas que se le presentan a la confiabilidad interna. Asimismo destacan:

1. Usar categorías descriptivas de bajo nivel de inferencia, es decir, lo más concretas y precisas posible, cercanas a la realidad observada.
2. Emplear varios investigadores u observadores, debido a que ello garantiza un mejor equilibrio de las observaciones, los análisis y la interpretación.



3. Pedir la colaboración de los sujetos informantes para confirmar la “objetividad” de las notas o apuntes de campo y asegurarse de que lo visto o registrado por el investigador coincide o es consistente con lo que ven o dicen los sujetos de la investigación.

4. Utilizar todos los medios técnicos disponibles en la actualidad para conservar en vivo la realidad presenciada. Para ello se emplean como técnicas e instrumentos de recolección de información: las narraciones y relatos concretos, equipo de evaluadores u observadores, revisión de los hallazgos por parte de otros evaluadores u observadores, entrevistas, observaciones participantes y no participantes, registros, videos, films, fotografías, cuestionarios, entre otros.

Confiabilidad externa

La confiabilidad externa se logra cuando al replicar un estudio, diferentes investigadores llegan a los mismos resultados. Los autores consideran que el evaluador puede aumentar esta confiabilidad siempre y cuando recurra a estrategias como las siguientes:

1. Precisar el nivel de participación y la posición asumida por el investigador o evaluador en el grupo estudiado.
2. Identificar claramente a los informantes.
3. Especificar el contexto físico, social e interpersonal en los que se recogen los datos.
4. Precisar los métodos de recolección de la información y de su análisis, de tal manera que otros investigadores puedan servirse del reporte original como un manual de operación para repetir el estudio.

Para Guba & Lincoln (1989), el término de confiabilidad y objetividad están insertados en los llamados criterios paralelos que estiman operen en forma análoga a los criterios de rigor que han sido usados dentro del paradigma convencional durante muchos años. Estos criterios funcionan apropiadamente dentro del marco de referencia del positivismo lógico, pero son inadecuados dentro de los enfoques constructivistas.

En este sentido, Guba & Lincoln han propuesto sus equivalentes:

1. Dependencia (dependability), término inexistente en el idioma español, es paralelo al concepto convencional de confiabilidad o consistencia de las observaciones en el tiempo.
2. Confirmabilidad, concepto considerado paralelo a la concepción de objetividad. A pesar de su paralelismo, estos criterios tienen ciertas afinidades con los positivistas convencionales y, además, son fundamentalmente metodológicos, característica de importancia secundaria para el constructivismo.



Por otra parte, estos autores consideran que la dependencia o consistencia está referida a que en el curso de la observación, el investigador que la realiza es contrastado a la vez con un agente interno o externo al estudio, como persona competente que expresa su opinión sobre todo el proceso observado. Sí los resultados se repiten entonces se puede asegurar la confiabilidad.

El uso de los criterios de “consistencia o dependability” y confirmabilidad, representan una medida que puede aumentar la seriedad y credibilidad de la investigación cualitativa. En este sentido, las condiciones antes descritas pueden ser usadas para revisar tanto el proceso como el producto de la investigación.

Validez Cualitativa

La validez en el contexto de la investigación y evaluación cualitativas, están referidas a “la precisión con que los hallazgos obtenidos reproducen efectivamente la realidad empírica y los constructos concebidos caracterizan realmente la experiencia humana” (Hansen, citado en Pérez Serrano, 1998; p. 80).

Guba & Lincoln (1985), señalan algunos criterios de validez paralelos que estiman procedan en forma análoga a la validez interna y externa usada en el paradigma positivista. En este caso se refieren a sus equivalentes: credibilidad y transferibilidad.

Credibilidad

Acepción paralela al de validez interna, en el sentido de que el isomorfismo entre los hallazgos y la realidad es reemplazado por la similitud entre las realidades construidas por los participantes en el proceso y las reconstrucciones del evaluador atribuidas a ellos (triangulación).

Para los autores, la credibilidad se logra cuando el investigador, a través de observaciones y conversaciones con los participantes del estudio, recolecta información que produce hallazgos que son reconocidos por los informantes como una verdadera aproximación sobre lo que ellos piensan y sienten. Para ello, tiene la tarea de captar el mundo del informante de la mejor manera que él lo pueda conocer, creer o concebir por lo que requiere escuchar de manera activa, reflexionar y tener una relación de empatía con el informante.

De acuerdo con esto, se puede decir que el objetivo del investigador es insertarse en el mundo de las personas lo cual exige paciencia, reflexión y evaluación permanente para describir los sujetos de investigación.

Transferibilidad

La transferibilidad puede ser concebida como paralela a la validez externa o generabilidad. Este criterio está referido a la posibilidad de extender los resultados del estudio a otras poblaciones.



Guba & Lincoln señalan que se trata de analizar qué tanto se ajustan los resultados con otro contexto. En la investigación cualitativa la audiencia o el lector del informe son los que determinan si pueden transferir los hallazgos a un contexto diferente del estudio. Para ello se requiere la descripción detallada del lugar y las características de los sujetos donde el fenómeno fue estudiado. En consecuencia, el grado de transferibilidad es la acción de establecer similitud entre los contextos.

De allí que estos autores relacionen a la validez interna, con la medida en que las observaciones y mediciones científicas sean representaciones auténticas de alguna realidad; y a la validez externa con el proceso de generalización, el cual permite aplicar los hallazgos obtenidos a otra situación similar a la investigada.

Otros investigadores conservando los lineamientos de validez interna propuestos por Guba & Lincoln, señalan que esta precisión se podría obtener aplicando ciertos métodos dentro de los cuales destacan como los más usuales: la triangulación, la saturación y el contraste entre colegas.

Triangulación

Al respecto Taylor y Bogdan (1990), consideran que la triangulación está concebida como un modo de proteger las tendencias del investigador, al confrontar y someter a control recíproco los relatos de los diferentes informantes involucrados en la investigación. Elliot (1986) señala que, al comparar diversas apreciaciones deberá observarse dónde difieren, están de acuerdo o en desacuerdo. En este sentido Santos Guerra (1990), declara que la prioridad de la triangulación está en todos aquellos aspectos coincidentes y recomienda explicar por qué se producen los desacuerdos sobre la base de los métodos empleados. Y para Woods (1993), este método tiene una fuerza enorme, siempre y cuando se utilicen tres procedimientos distintos para obtenerla, debido a que con ello se aumenta la probabilidad de exactitud en los resultados.

En este sentido Denzin citado en Pérez Serrano (1998), señala que existen diversas modalidades o tipos de triangulación, entre los que destaca a los que se mencionan a continuación: (a) triangulación de tiempo o momentos, (b) triangulación de sujetos o niveles combinados de triangulación, (c) triangulación teórica, (d) triangulación del investigado y (e) triangulación de métodos.

Triangulación de tiempo o de momentos: Al respecto Santos Guerra (1990), señala que la información recabada es contrastada desde una perspectiva temporal que se inicia con antes, durante y después que se producen los hechos en los sujetos y en los ambientes, con la finalidad de contemplar el fenómeno desde ópticas distintas pero complementarias para analizarlo en profundidad. De allí que este proceso se desarrolle en forma dinámica tomando en cuenta las intenciones, expectativas propósitos y objetivos (antes) en el compromiso, interés, motivación y participación de los sujetos (durante) para la satisfacción, rectificaciones y análisis (después) que ayudan a comprender su significado y a valorar su desarrollo.



Triangulación de sujetos o niveles combinados de triangulación: En este tipo de procedimiento los involucrados en la investigación o evaluación, participan a nivel individual, interactivo o en grupos y a nivel de colectividades, organizacional, cultural o social.

Triangulación teórica: está basada en la preferencia de teorías alternativas sobre la utilización de un solo punto de vista, la del investigador.

Triangulación del investigado: a este tipo de procedimiento se le denomina también, triangulación intersubjetiva, debido a que la verificación se realiza del contraste de la información emitida por un grupo o equipo de investigadores que participan en el estudio, con la finalidad de disentir diferentes puntos de vista sobre la misma realidad.

Triangulación de métodos: se realiza usando una gama diversa de instrumentos o métodos. En este procedimiento subyace un principio básico que consiste en recoger observaciones acerca de una situación o algún aspecto de ella, desde una variedad de perspectivas para después compararlas y contrastarlas.

Pérez Serrano (1998) manifiesta que la triangulación es más efectiva cuando el investigador combina varias fuentes o métodos, debido a que estos permitirán contrastar puntos de vista distintos sobre una misma situación.

Saturación

Este procedimiento consiste en reunir las evidencias suficientes para garantizar la credibilidad de la investigación, es decir, cuando en una investigación los informantes claves coinciden en sus observaciones o apreciaciones, se dice que la información se ha saturado debido a que no aporta nuevos hallazgos.

Contraste entre colegas

Este método de validez también es conocida como triangulación de sujetos, respondente o de negociación; y como se comentó anteriormente la información obtenida es contrastada por el investigador desde diferentes puntos de vista con otros informantes, observadores o personas implicadas en la investigación.

Como puede observarse, el concepto de validez es importante tanto para los estudios cuantitativos como los cualitativos, sin embargo en éstos últimos no representan a un concepto solo, fijo o universal, sino el producto de un constructo que surge de los procesos y de las intenciones de las metodologías empleadas.

Confiabilidad y Validez Hermenéuticas



Para comprender y explicar la complejidad que representan los problemas sociales Guba y Lincoln (1989), proponen una evaluación basada en el enfoque hermenéutico-constructivista, en un contexto participativo que conduzca a la consolidación de una agenda de negociación.

Estos autores establecieron una premisa que concibe la existencia de diferentes realidades como construcciones sociales de los sujetos y considera la interacción entre lo observado y el observador. Asimismo emplean métodos participativos para la integración de las experiencias de los distintos actores que da lugar a otras interpretaciones destinadas a que los involucrados en la investigación o evaluación reconozcan la multiplicidad de lo cultural y el papel de lo simbólico en la conducta humana a través de una estrategia democrática y de participación ética que orienta su construcción en una nueva práctica pedagógica, más humana y sensible al otro, apoyada en una comprensión distinta centrada en la hermenéutica del diálogo. Un evaluador que se caracteriza por ser creativo, flexible, solidario, comunicativo y negociador que construye su propia realidad a través de la experiencia y del diálogo intersubjetivo entre quienes participan en el proceso de evaluación. De allí que la evaluación vista desde esta perspectiva promueva la participación de todos sus actores haciendo uso de herramientas como la autoevaluación, coevaluación y heteroevaluación, procesos estos que le permitirán al evaluador contrastar la realidad observada desde diferentes ángulos y así concebir significados individuales, de pares y colectivos.

Confiabilidad Hermenéutica

La confiabilidad desde una perspectiva hermenéutica está orientada hacia el nivel de concordancia interpretativa entre diferentes observadores, evaluadores o jueces del mismo fenómeno, es decir, la confiabilidad será, sobre todo interna, interjueces.

Según García (2002) la confiabilidad hermenéutica se apoya en tres medios racionales que son: la intuición, la interpretación y el entendimiento que le permita al investigador o evaluador, de manera conjunta, el desarrollo un proceso real y veraz de la situación a investigar o evaluar.

En este sentido, aplicar este tipo de confiabilidad en un proceso de evaluación sea este cual fuera, permitirá al evaluador percibir, explicar y darle sentido a su interpretación y a la de sus colaboradores de la manera más fehaciente, en relación con la situación objeto de evaluación.

Validez Hermenéutica

En la evaluación enmarcada en el paradigma cualitativo y específicamente en el enfoque hermenéutico **la validez se obtiene a través de la contrastación e interpretación de los hallazgos hasta lograr un entendimiento y consenso. En este sentido, las formas de participación: autoevaluación, coevaluación y heteroevaluación, juegan un papel importante que impregnan de validez y confiabilidad el método hermenéutico de investigación** el cual se traduce según Pérez Serrano (1998), en:

Validez ecológica



En cuanto a la validez ecológica, el proceso de la autoevaluación (propia del sujeto), coevaluación (por sus pares) y heteroevaluación (por los actores internos y externos) se debe realizar sin crear situaciones artificiales, es decir, en el contexto social natural donde se producen. En este sentido, se deben tener en cuenta las condiciones de la vida y el ambiente de los sujetos investigados, las motivaciones, los intereses, la comprensión y la solidaridad tanto del evaluado como la del evaluador.

Validez comunicativa

Es el intento de asegurar un adecuado proceso de análisis e interpretación mediante la repetición, a través de la nueva comunicación (feedback) con los evaluados. Más importante es la información de los evaluados sobre los resultados, analizar sus reacciones y mediante estas verificar que el análisis e interpretación corresponden a la realidad.

Validez argumentativa

El evaluador hace públicos sus suposiciones previas y valida sus interpretaciones basándose en los conocimientos que comparte con otros quienes participan en el proceso de evaluación. La argumentación es guiada por reglas y debe ser entendible por otros, garantizando la intersubjetividad en la interpretación. La validez argumentativa es valorada por los profesores quienes se convierten en la investigación fenomenológica como evaluadores críticos de lo interpretado por el investigador.

Validez cumulativa

Se realiza mediante un proceso de contrastación entre los resultados de una investigación con otras realizadas. Esta validez tiene cierta afinidad con la validez del experto y la validez comunicativa (diálogo con otros investigadores y/o los investigados) no puede asegurar validez sino sólo plausibilidad.

Validez práctica

Se logra a través de la realización del conocimiento científico en la práctica social. Es mediante esta praxis cuando el conocimiento expresa su verdadero sentido. Este tipo de validez asume un carácter sistemático, y las posibilidades de casualidad respecto a la aplicación del conocimiento se somete al contexto, por lo cual es difícil lograr coincidencia. La pretensión de réplica de los resultados en la práctica es problemática en la investigación social cualitativa, aunque la validez práctica persigue confirmar o falsear una hipótesis.