

Autor Principal: *Dra. Raquel Morquecho Sánchez

Co-autores: *Dr. José Alberto Pérez García, **Dra. Verónica Morales Sánchez, *Dr. Jorge Zamarripa Rivera, *MC. Juan Carlos Arturo González Castro

*Facultad de Organización Deportiva, Universidad Autónoma de Nuevo León

**Facultad de Psicología, Universidad de Málaga

Bases metodológica, técnicas de análisis de datos y plataforma on-line aplicados en la Gestión Deportiva

Introducción

El servicio de calidad es un medio que las organizaciones tienen para diferenciarse y competir en el mercado, por ello, una adecuada aplicación de gestión de la calidad en las organizaciones deportivas; así como el uso de instrumentos viables, plataformas on-line, metodologías y técnicas de análisis de datos, supone una evolución con respecto a las exigencias del mercado actual.

En México en este ámbito existen escasas investigaciones aplicadas al campo de la gestión deportiva, por tal motivo en el presente capítulo se muestra un breve resumen de las plataformas on-line, las bases metodológicas y las técnicas de análisis de datos utilizadas en el ámbito de la gestión deportiva, esto con el objetivo de promover las estrategias que se vienen empleando para mejorar la calidad del servicio deportivo. En este sentido, Anguera (2003) establece que el marco metodológico sustenta el estudio dentro de la investigación social, siendo esta muy utilizada en el ámbito del deporte por sus diversas aplicaciones; esto se debe a la facilidad que ofrece para obtener respuestas masivas.

Plataforma on-line en la gestión deportiva

En la actualidad la informática e internet son una revolución técnica y conceptual en todas las áreas de conocimiento y especialmente en la Gestión Deportiva. La aparición de plataformas on-line, que utilizando distintas metodologías de investigación (selectiva o experimental), proporcionan los actuales procesos de formación que caracterizan a las comunidades en línea y que tienen como objetivo la enseñanza. (Hernández Mendo, Morales Sánchez y González Ruiz, 2012).

En la investigación de Hernández Mendo et. al 2012, se tiene como objetivo el presentar, parcialmente la fiabilidad de una parte de la plataforma MenPas (www.menpas.com). Se presentan los datos relativos a la fiabilidad del área de calidad/organizacional. Actualmente está compuesta de 16 áreas de trabajo y más de 70 tareas y cuestionarios que tienen como objetivo evaluar desde la atención pasando por los estados de ánimo, liderazgo, socialización, motivación y calidad organizacional.

Los resultados avalan la propuesta metodológica de uso de la plataforma MenPas (www.menpas.com) como una herramienta fiable de evaluación psicosocial on-line en el área de la gestión de la calidad, pues presenta unos datos óptimos en el Alpha de Cronbach (entre 0.838 y 0.961). El uso de plataformas tanto para la gestión como para la investigación proporciona una valiosa herramienta que optimiza los recursos materiales y el tiempo; así como facilitar el acceso a grandes muestras como lo avala el trabajo de Gosling, Vazire, Srivastava, y John (2004).

Metodología Selectiva

Según Anguera (1990) considera que la metodología selectiva se caracteriza por tener un nivel medio de control interno, aspecto que la diferencia de la metodología observacional y de la experimental, pues en la primera la respuesta se recoge mediante un registro sin la

intervención del individuo observado, y en la segunda, el responsable de obtener la información mantiene un alto grado de dominio y control de las variables de medida.

Existen tres criterios a considerar por los investigadores para caracterizar cualquier metodología:

- a) Generalizabilidad con respecto a las poblaciones a las que se desea aplicar la información.
- b) Precisión con respecto a la medición y control de las variables implicadas.
- c) Fiabilidad con respecto a los contextos en los cuales opera.

En el proceso de uso de la metodología selectiva se han distinguido tres momentos de desarrollo según Arnau (1990), Buendía, Colás y Hernández (1998):

- a) Teórico-conceptual: incluye el planteamiento de objetivos y/o problemas e hipótesis de la investigación.
- b) Metodológico: se inicia en el momento de seleccionar la muestra y definir las variables de objeto de estudio.
- c) Estadístico-conceptual: incluye la confección del instrumento, la codificación y análisis de los datos para poder elaborar conclusiones.

De acuerdo con Anguera (2003), deben ser consideradas cinco fases en este proceso de uso de esta metodología:

- a) Planificación del estudio: La característica fundamental de esta etapa es la formulación y delimitación de los objetivos específicos que permitan la formulación de la hipótesis (si las hubiera). Estas acciones deben ser planificadas cuidadosamente. Se recomienda a partir de una adecuada documentación y recopilación de la información de estudios que tengan conexión con el área de estudio.

- b) Elaboración del instrumento (si fuera preciso): La encuesta es el instrumento mas utilizado en un conjunto de procedimientos estandarizados de investigación.
- c) Uso del instrumento: En este apartado deben ser considerados los siguientes puntos, como la formación y capacitación de los encuestadores, el plan de muestreo, selección de la muestra y los riesgo de error (errores de no observación, medida y de procedimiento).
- d) Transformación y análisis de datos: El proceso de tratamiento de los datos transcurre desde el momento de la recogida de información hasta que se obtienen los resultados del estudio.
- e) Elaboración del informe: Tras la obtención de los resultados del estudio, se procede a la redacción del informe de investigación que permite la difusión del trabajo realizado.

Técnicas de Análisis de Datos

Análisis de fiabilidad

Existen muchas posibilidades de error que puedan afectar a la calidad de la información obtenida, estos errores pueden ser aleatorios o sistemáticos con lo cual afectarán a problemas de fiabilidad y validez respectivamente.

De acuerdo a Abad, Garrido, Olea y Ponsada (2006) se entiende por fiabilidad a el grado de estabilidad que diferentes subconjuntos de ítems miden un rasgo o comportamiento homogéneo; es decir el grado en que covarian, correlacionan o son consistentes entre sí. Indica la replicabilidad de la medida a través de distintas condiciones, momentos, etc.

El Modelo Clásico o Modelo Lineal Clásico, desarrollado por Spearman, está fundamentado en diversos supuestos a partir de los cuales se definen los conceptos de puntuación verdadera y error. El modelo de Spearman distingue entre el valor real del

atributo que se mide (la puntuación verdadera) y la medida fiable que se obtiene en el proceso de medición (la puntuación observada). Es decir, para poder graduar la precisión de un instrumento, es que, cuanto más preciso es, más se replicarán nuestras observaciones en sucesivas mediciones.

El Modelo Clásico nos permite deducir que parte de la variabilidad en las puntuaciones en un test se debe a la variabilidad en el atributo medido y que parte se explica por la presencia de errores en el proceso de medición. El presente modelo está conformado por tres teorías:

1. Teoría Clásica de los Test (TCT): Thurstone (1947) determina que las medidas de un instrumento deben ser independientes a los objetos medidos, supone que los ítems son equivalentes y que la fiabilidad se reparte de igual manera en todos los ítems. De acuerdo a Hernández Mendo (2006) la TCT considera los casos de un test como una muestra representativa de un universo de ítems, los cuales son considerados indicadores del constructo que medimos; por lo cual es ilógico que la medición de las variables sean independientes del instrumento, así como las propiedades de los instrumentos no son independientes de los sujetos.
2. Teoría del Respuesta al Ítem (TRI): Esta teoría tiene ventajas sobre la TCT, ya que se tiene en cuenta el sesgo del test y hace suposiciones de mayor potencia estadística. La TRI tiene como objetivo obtener mediciones que no varíen en función del instrumento; así como de disponer de instrumentos que sean invariantes respecto a los objetos de estudio. Esta teoría relaciona a los sujetos e ítems de modo interactivo, toma a los ítems como una unidad de análisis, lo cual permite describir las propiedades psicométricas del instrumento.

Hernández Mendo (2006) muestra una principal diferencia entre las TCT y TRI, para la TCT esta se dirige a las propiedades de la puntuación global de un test, mientras que la TRI se centra en las propiedades particulares de cada ítem.

3. Teoría de la Generalizabilidad (TG): Es una teoría conocida como de los errores multifacetas, la cual consiste en obtener la variabilidad real a la variabilidad del error, por medio de estudios apriorísticos, estimando diseños de medida precisos para llevar a cabo planes de optimización.

Podemos calcular la fiabilidad de diferentes maneras:

1. Correlación entre formas paralelas, es cuando se replican las mismas medidas al aplicar una prueba paralela con ítems distintos. Si ambas formas son paralelas, la correlación entre ambas indica su grado de equivalencia. En este sentido, replicabilidad implica que debemos obtener las mismas medidas cuando medimos lo mismo con pruebas equivalentes.
2. La estabilidad temporal de las medidas que proporcionan nuestro instrumento. La replicabilidad implica que debemos obtener las mismas medidas cuando medimos lo mismo en momentos distintos.
3. Finalmente, la consistencia interna, es el grado en que diferentes partes del test miden un rasgo. En este sentido, replicabilidad implica una concordancia entre las puntuaciones cuando medimos lo mismo con distintas partes del test. Esta última forma de calcular la fiabilidad está representada por el coeficiente “alfa de Cronbach”, el cual refleja el grado de covariación de los ítems. Su valor puede variar entre cero y uno, por lo tanto según Muñiz (2001) si los ítems covarían fuertemente asumirá un valor cercano a 1, mientras que si los ítems son linealmente independientes, asumirá valores cercanos a 0.

Visauta (1998) afirma que todo cuestionario compuesto por ítems debe contener un análisis de consistencia interna que detecte hasta qué punto es fiable la realidad evaluada con el cuestionario utilizado. Éste tipo de análisis indica el grado en que distintos ítems son coherentes entre sí, midiendo de esta forma la misma variable o una misma magnitud.

Análisis factorial

El Análisis Factorial (AF) es una técnica estadística multivariante que sirve para estudiar las dimensiones que subyacen a las relaciones entre variables. El origen de la técnica se remonta a principios del siglo XX, cuando el psicólogo británico Charles Spearman estudió el patrón de correlaciones entre distintas medidas de rendimiento. El análisis factorial es uno de los procedimientos estadísticos más utilizados en la investigación social, donde está inmersa la gestión deportiva.

Marôco (2010) determina que el análisis factorial es una técnica de modelación lineal general, cuyo objetivo es identificar un conjunto reducido de variables latentes (factores) que expliquen la estructura correlacional observada entre un conjunto de variables manifiestas (ítems).

Los modelos del análisis factorial puede clasificarse en dos tipos: Análisis Factorial Exploratorio (AFE) y Análisis Factorial Confirmatorio (AFC), los cuales desarrollaremos a continuación.

Análisis Factorial Exploratorio (AFE)

Según Lloret, Ferreres, Hernández-Baeza, y Tomás (2014), el Análisis Factorial Exploratorio es una de las técnicas más usadas en el desarrollo, validación y adaptación de instrumentos. En las últimas décadas numerosos estudios han detectado la necesidad de actualizar los criterios clásicos para incorporar aquellos más adecuados para realizar una investigación de calidad.

De acuerdo con Lloret et al 2014, actualmente existen una variedad de software disponible en el mercado SPSS, FACTOR (Lorenzo-Seva y Ferrando, 2006), LISREL (Jöreskog y Sörbom, 2007), y MPlus (Muthén y Muthén, 2007), la utilización los mismos, lo decide el investigador, lo cual permite o limita la aplicación de estos nuevos estándares.

El análisis factorial exploratorio es una técnica de reducción de la dimensionalidad que permite pasar de un conjunto de variables observadas (ítems) a un número de variables latentes (factores); es decir estudia que estructura factorial se ajusta mejor a los datos y no se requiere de previsiones exactas (Abad, Olea, Ponsoda y García, 2011).

Las pruebas que indican la pertinencia del análisis factorial exploratorio:

1. Determinante de la matriz de correlaciones: Es un indicador del grado de intercorrelaciones existente.
2. Prueba KMO (medida de adecuación muestral de Kaiser-Meyer-Olkin): Compara los coeficientes de correlación simple con los coeficientes de correlación parcial, debiendo asumir valores mayores a 0.6.
3. Test de Esfericidad de Bartlett: Comprueba si la matriz de correlaciones es una matriz de identidad, se buscan valores elevados del test con significatividad inferior a 0.05.

Según Morquecho (2014) a continuación se menciona una breve explicación del proceso para llevar a cabo un AFE, en primer lugar se debe decidir cuál es el número de factores que se van a extraer por medio de un método de extracción (máxima verosimilitud (ML), mínimos cuadrados no ponderados (ULS) y factorización de componentes principales (EP)) el cual no referimos a los procedimientos donde se estiman los parámetros del modelo. Desde el punto de vista estadístico el modelo de extracción más utilizado es el (ML); ya

que en este modelo se toma en cuenta una matriz de correlaciones que pretende aproximar las estimaciones de los parámetros.

Una vez decidido el método de extracción, se comienza con la selección del número de factores, el cual tiene como objetivo establecer cuantas dimensiones esta midiendo un test; es decir cuántos factores deben incluirse en la solución factorial. Comenzando con el análisis de la solución de un factor, si el ajuste del modelo es bueno se tendrá una solución unidimensional; en el caso que no sea así, se considerará una solución de dos factores y así sucesivamente hasta que se obtengan los factores necesarios para que ajuste el modelo.

Ya que se establezca el número de factores se obtiene una solución factorial la cual se llama matriz factorial inicial, es aquí donde el investigador toma decisiones y busca la más simple por medio de la elección de un método de rotación (ortogonal) (factores independientes), oblicua (factores correlacionados).

Los métodos disponibles para la rotación que es donde se distribuye la varianza en otros factores son:

1. Ortogonales (rotación varimax, quartimax y equimax)
2. Oblicuos (rotación oblimin directo y promax)
3. Ortogonal (rotación varimax)

Estas decisiones son importantes; ya que este tipo de análisis permite a los investigadores que un estudio sea replicable y que se vuelvan analizar los resultados.

Análisis Factorial Confirmatorio (AFC)

El análisis factorial confirmatorio (AFC) es una herramienta estadística fundamental en psicometría para obtener evidencia sobre la estructura interna de los ítems. En este análisis el investigador plantea hipótesis definidas a priori sobre el número de factores, si hay o no

correlaciones entre los factores, como saturan las variables observadas y si existen correlaciones entre los términos de error específicos (Abad, Olea, Ponsoda y García, 2011).

El Análisis Factorial Confirmatorio (AFC) se pone a prueba si una solución factorial concreta es o no adecuada para los datos, por medio de modelos de ecuaciones estructurales (Jöreskog, 1970).

Arias (2008) menciona que los requisitos necesarios para poder llevar a cabo un AFC son la normalidad de la muestra, el control de outliers, el tratamiento adecuado de datos perdidos, la ausencia de multicolinealidad, el número mínimo de observaciones, entre otros aspectos.

Para realizar un AFC se requiere tomar decisiones en cuanto a:

1. Realizar una representación grafica del modelo
2. Estudiar si el modelo está identificado; es decir si existe suficiente información para estimar sus parámetros
3. Estimar los parámetros del modelo (para ellos debe seleccionarse un método de estimación (máxima verosimilitud (ML), mínimos cuadrados (ULS))
4. Obtener indicadores del ajuste del modelo (contrastar si los datos empíricos se ajustan a lo predicho por el modelo teórico)

El AFC tiene algunas ventajas, ya que permite:

1. Contrastar directamente el modelo teórico del investigador
2. Estudiar modelos complejos
3. Establecer restricciones en los pesos
4. Reducir el número de parámetros a estimar

Existen diferentes métodos de estimación del AFC, en el caso de que no se cumpla el supuesto de normalidad multivariada, se recomienda utilizar el método de máxima

verosimilitud robusta (MLM) Maximum Likelihood Mean Adjusted (Bentler, 1995). El método más común de estimación es el de máxima verosimilitud (ML) Maximum Likelihood.

Si uno o más de los indicadores es categórico se optará por otros métodos de estimación tales como mínimos cuadrados ponderados (WLS) Weighted Least Squares, mínimos cuadrados ponderados diagonalizados, (DWLS) Diagonal Weighted Least Squares, mínimos cuadrados ponderados robustos (WLSMV) Weighted Least Squares Mean and Variance Adjusted o mínimos cuadrados no ponderados (ULS) Unweighted Least Squares (Arias, 2008).

Conforme a la literatura existen una serie de índices de ajuste (GFI, AGFI, CFI, NNFI) y de error (RMSR, RMSEA) (Bentler y Bonnet, 1980; Hu y Bentler, 1999; Jackson, 2007; Marôco, 2010) los cuales se recomienda considerar para interpretar el AFC.

Morales Sánchez (2003) realiza una breve descripción de cada índice, los cuales se presentan a continuación brevemente:

(GFI) Goodness of FitIndex o Índice de Bondad de Ajuste: Este índice explica la proporción de covariancia observada entre las variables manifestadas, pese a que no existe ningún umbral absoluto de aceptabilidad, los valores oscilan entre 0 (mal ajuste) y 1 (ajuste perfecto), por lo que altos valores indican un mejor ajuste (>0.90) (Tanaka y Huba, 1985).

(AGFI) Adjusted Goodness of FitIndex o Índice Ajustado de Bondad: Este índice es una extensión del índice GFI. Un nivel aceptable y recomendado sería un valor mayor o igual a 0.90.

(CFI) Comparative FitIndex o Índice de Ajuste Comparado: Bentler (1990) propone este índice el cual representa una comparación entre el modelo estimado y el modelo nulo o independiente. Los valores oscilan entre 0 y 1, indicando valores altos una buena calidad de

ajuste. En este sentido, valores superiores a .90 se consideran buenos mientras que aquellos superiores a .95 son muy buenos.

(NNFI) Non-Normed Fit Index o Índice de Ajuste No Normado: Este índice es propuesto por Tucker y Lewis (1973). La cuota superior no es la unidad por lo que valores superiores a 1 tienen a indicar una sobre-parametrización del modelo.

(RMR) Root Mean Residual o Residuo Cuadrático Medio; (RMSR) Root Mean Square Residual: El índice se obtiene a través de la raíz cuadrada de los residuos al cuadrado entre las matrices observadas y estimadas. Se puede considerar como aceptables los valores comprendidos entre 0.05 y 0.08.

(RMSEA) Root Mean Squared Error of Approximation o Error de Aproximación Cuadrático Medio: Este índice muestra la discrepancia en grados de libertad. Según Arbuckle (2008) considera buenos valores <0.05 , razonables o aceptables valores entre 0.05 y 0.08, mientras que serían valores mediocres aquellos comprendidos entre 0.08 y 1; valores >1 deben rechazarse.

Arias (2008) menciona que como complemento al AFC, a los coeficientes y varianzas de error resultantes mediante una solución completamente estandarizada se obtiene el valor de la varianza media extractada, fiabilidad compuesta, validez convergente y validez discriminante; los cuales definimos a continuación:

Varianza media extractada: Mide el porcentaje de varianza capturada por un constructo, es una medida complementaria a la fiabilidad compuesta. Cuando es superior a 0.50 implica que un alto porcentaje de la varianza es explicada por el constructo en comparación con la varianza de error de medida.

Fiabilidad compuesta: Es la fiabilidad conjunta de los indicadores de una variable latente, se utiliza como medida alternativa y es fácilmente calculable a partir de los resultados del AFC, valores superiores a 0.7 son indicadores de una fiabilidad compuesta apropiada.

Validez Convergente: Se evalúa por medio de los valores de t correspondientes a las saturaciones factoriales. Es el coeficiente de correlación entre medidas del mismo constructo cuando se utilizan distintos procedimientos de medida. Los valores de t superiores a 1.96 proporcionan evidencia de validez de los indicadores utilizados para medir los constructos; es decir los valores significativos de t indican que todos los indicadores evalúan el mismo constructo.

Validez Discriminante: Es el coeficiente de correlación entre medidas de distintos constructos cuando se utiliza el mismo procedimiento de medida, puede determinarse en la medida en que la varianza media extractada de cada variable latente es superior al cuadrado de la correlación entre ellas.

Análisis de variabilidad

El análisis de variabilidad, los gestores deportivos lo pueden aplicar para conseguir un acercamiento a la mejora de la calidad en los servicios deportivos. De acuerdo a Blanco (1993), las etapas para calcular el análisis de variabilidad son las siguientes:

1. Análisis de componentes de variancia
2. Análisis de generalizabilidad
3. Plan de optimización

Análisis de componentes de variancia

Conforme a una revisión bibliográfica autores como Blanco, Castellano y Hernández Mendo (2000); Blanco y Hernández Mendo (1998); Blanco y Losada (2004); Morales

Sánchez (2003); Morales Sánchez, Pérez López, Morquecho Sánchez y Hernández Mendo (2016) y Morquecho Sánchez (2014), han utilizado el análisis de variancia en el área de la actividad física y del deporte para la estimación a partir de la estructura numérica de distintos modelos. Considerando el ajuste de los modelos al Modelo Lineal General por medio de la comparación de la variancia residual de los procedimientos de mínimos cuadrados y de máxima verosimilitud.

Según Blanco Villaseñor y Losada (2004) el procedimiento (GLM) General Linear Model de SAS utiliza el método de mínimos cuadrados (LS) para ajustar modelos lineales generales. Mientras que el procedimiento VARCOMP de SAS calcula estimaciones de los componentes de variancia en un modelo lineal general, el cual tiene cuatro métodos diferentes para la estimación de los componentes de variancia, los cuales son el método Tipo I, que es equivalente a lo que genera GLM en su opción de sumas de cuadrados; el segundo método MIVQUE0 basado en la técnica de mínima norma o mínima varianza, se trata de un procedimiento de componentes de la varianza por mínimos cuadrados corregidos; el tercero es el método de Máxima Verosimilitud (ML), el cual calcula las estimaciones de máxima verosimilitud de los componentes de variancia y por último el cuarto método de Máxima Verosimilitud Restringida (REML) este separa primero la verosimilitud en dos partes, una que contiene los efectos fijos y otra los aleatorios, utiliza algoritmos similares a los de ML, y también sus valores iniciales son los de MIVQUE0. El procedimiento MIXED de SAS ajusta modelos lineales mixtos y calcula pruebas basadas en las sumas de cuadrados Tipo III para los efectos fijos. De acuerdo con Rao (1997) desde la literatura los estimadores de máxima verosimilitud presentan mejores propiedades asintóticas.

Es de vital importancia para el área de gestión deportiva el realizar este tipo de análisis; ya que son la pauta para poder llevar a cabo el análisis de generalizabilidad y el desarrollo de un plan de optimización que contribuirá a realizar mejor el costo-beneficio.

Análisis de Generalizabilidad

Según Hernández-Mendo, Blanco-Villaseñor, Pastrana, Morales-Sánchez, y Ramos-Pérez (2016), la Teoría de la Generalizabilidad (TG) es relativamente joven en cuanto a su uso. Puede considerarse una extensión de la Teoría Clásica de los Tests (TCT), utilizando los procedimientos del análisis de la varianza y de los diseños experimentales (Martínez-Arias, 1995).

Una de las técnicas de análisis de datos utilizadas en el área de la gestión deportiva, es la Teoría de la Generalizabilidad (Cronbach, Gleser, Nanda, y Rajaratnam, 1972; Cronbach, Rajaratnam, y Gleser, 1963), la cual es necesaria para realizar un control que permita una adecuada calidad de los datos; así como una mejor y precisa estimación de los diferentes tamaños muestrales.

Según Blanco; Castellano; Hernández Mendo; Sánchez López y Usabiaga, (2014), concretan algunas de las aplicaciones que la TG en el cálculo de la fiabilidad, validez y estimación muestral. La aplicación de la TG se establece con el objetivo de obtener un adecuado plan de calidad, la estimación de las fuentes de variabilidad y estimar los diseños de medida óptimos para una adecuada evaluación de la calidad, así como promover las estrategias que se vienen empleando para mejorar la calidad del servicio deportivo mexicano.

La Teoría de la Generalizabilidad es una teoría de los errores multifacetas que tiene como objetivo desglosar cualquier tipo de medición de la variabilidad real a la de la variabilidad del error; a su vez reconoce explícitamente las múltiples fuentes de error de medida

(participantes, contextos, tratamientos, sesiones) pudiendo estimar cada una de estas fuentes de error así como las diferentes interacciones entre ellas. El error de medida no es más que el efecto de las fluctuaciones debidas al muestreo de niveles particulares en cada una de las facetas (variables) del universo de observaciones posibles. De acuerdo a Morales Sánchez (2009) optimizar dicha medida es adaptar nuestro diseño para reducir al máximo la varianza del muestreo debido a estas facetas.

El análisis de generalizabilidad es un estudio apriorístico que tiene como objetivo la estimación de diseños de medida precisos. Para el logro de este objetivo se realiza una estimación de los componentes de varianza, dado que su magnitud aporta información sobre las fuentes de error que puedan afectar a una medición (Blanco y Hernández, 1998).

Cabe señalar que la medición en las ciencias del comportamiento, tiene relación con la teoría clásica de los tests, en las que están fundamentados, el coeficiente de fiabilidad, validez y precisión.

La TG es considerada, por un lado, como una extensión de la Teoría Clásica de los Tests, utilizando los procedimientos del análisis de la varianza y de los diseños experimentales (Martínez, 1995); y, por otro, como una teoría de los errores multifaceta asumiendo que cualquier situación de medida posee infinitas fuentes de variación o facetas (Cronbach, Gleser, Nanda y Rajaratnam, 1972).

Berk (1979) y Blanco Villaseñor (1993) mencionan, para que se cumpla la teoría se necesita de los componentes del análisis de la variancia, es decir identificar las variaciones de las facetas. El análisis de los componentes permite obtener una adecuada información, en lo que respecta a la contribución del error en un determinado diseño (Blanco Villaseñor, 1991; 1992).

El análisis de los mismos informa sobre qué facetas contribuyen con más error, para ser modificadas posteriormente en los sucesivos diseños. En la segunda fase del desarrollo de un análisis de generalizabilidad, es donde se lleva a cabo la elección de un modelo de estimación apropiado (ya sea de efectos aleatorios o mixtos), el cual está determinado por el modo de muestrear los niveles de cada faceta (Shavelson y Webb 1991).

De acuerdo a Blanco y Morales (2010), las aplicaciones de la TG se han centrado fundamentalmente en el ámbito de la metodología observacional, pero también puede utilizarse para la optimización de los tamaños de muestra ideales, constituyendo un estudio del costo-beneficio que permita mayor flexibilidad y parsimonia, calculando así un adecuado plan de optimización.

La aplicación de esta teoría según Blanco y Anguera (2003) se lleva a cabo en cuatro fases:

1. Definición de las facetas de estudio: Es descriptiva, donde se identifican los datos en un plan de medida, después se eligen las facetas a tener en cuenta y las interrelaciones entre las facetas estudiadas. Se decide el número de niveles muestreados en cada faceta y se utiliza el análisis de la variancia con el fin de calcular la suma de cuadrados y el cuadrado medio de cada fuente de variación del plan utilizado.
2. Análisis de varianza de las puntuaciones obtenidas sobre las facetas de estudio: Se realiza la elección de un modelo de estimación apropiado (ya sea de efectos aleatorios o mixtos) que está determinado por el modo de muestrear los niveles de cada faceta.
3. Cálculo de los componentes de error: Se analiza las propiedades de uno o más planes de medida. Esta fase sirve para precisar la intención de medida y para especificar qué faceta o facetas constituyen el objeto de estudio.

4. Optimización de los coeficientes de generalizabilidad: En esta fase se realizan los planes de optimización. Las informaciones obtenidas en los análisis anteriores se utilizan para identificar la mejor adecuación posible en los procedimientos de medida.

Hernández Mendo et al 2016, en su trabajo presentan el programa SAGT (acrónimo de Software Application for Generalizability Theory), el cual es una herramienta multilenguaje (inglés, español, francés y portugués) para la aplicación de la Teoría de la Generalizabilidad; el cual puede realizar tareas de estimación y optimización de las facetas de estudio de manera automática, incluyendo representaciones gráficas de las estimaciones realizadas, permite la importación y exportación de los datos desde y hacia otras aplicaciones.

Las disponibilidades de las opciones mostradas en el menú principal variarán en función del estado del programa (si se encuentra conectado a la plataforma MenPas www.menpas.com– Plataforma Psicosocial de Evaluación on-line) y del rol del usuario de la aplicación.

La utilización del programa SAGT ha resultado eficaz en el tratamiento y análisis de los datos a diversos niveles, especialmente en la gestión de opciones novedosas (Hernández Mendo et. al, 2016).

El campo de aplicación de la TG en el área de gestión deportiva en México se llevo acabo en un estudio que se realizó en el ámbito universitario, el cual tenía como objetivo evaluar la calidad percibida en organizaciones de servicios deportivos universitarios; se utilizó el análisis de generalizabilidad mostrando en los diferentes modelos estimados, que las escalas miden constructos diferentes y que, utilizando los planes de optimización, es

posible reducir los niveles de las facetas para poder establecer planes de optimización de costo-beneficio para las organizaciones deportivas universitarias, lo cual apoyaría en el proceso de planificación de los gestores deportivos universitarios (Morquecho, 2014).

Lo cual coincide con Blanco y Morales (2010), donde las aplicaciones de la TG se han centrado fundamentalmente en el ámbito de la metodología observacional, pero también puede utilizarse en el área de la actividad física y el deporte, para la optimización de los tamaños de muestras ideales, constituyendo un estudio del costo-beneficio que permita mayor flexibilidad y parsimonia, concluyendo que aplicado a la gestión deportiva mexicana y calculando un adecuado plan de optimización contribuirán para la mejora de las organizaciones deportivas públicas y privadas en México.

Este tipo de análisis y el diseño de cada una de las facetas, ha permitido realizar distintas aproximaciones de los diferentes diseños de medida, estimando las facetas y sus interacciones, para calcular un adecuado plan de optimización de la calidad en servicios deportivos.

Referencias

- Abad, F. J, Olea, J, Ponsoda, J, y García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid. Síntesis.
- Abad F. J, Garrido J, Olea J, Ponsoda V. (2006) *Teoría Clásica de los Tests y la Teoría de la Respuesta al ítem*. Madrid: Universidad Autónoma de Madrid.
- Arbuckle, J. L. (2008). *Amos 17 users' guide*. Chicago, IL: SPSS.
- Anguera, M. T. (1990). Metodología observacional. En J. Arnau Gras, M. T. Anguera Argilaga y J. Gómez Benito (Eds.), *Metodología de la Investigación en Ciencias del Comportamiento* (pp. 123-236). Murcia: Secretariado de Publicaciones de la Universidad de Murcia.

- Anguera, M. T. (2003). La metodología selectiva en el deporte. En A. Hernández-Mendo (Coord.), *Psicología del deporte* (Vol. II) - *Metodología* (74-96). Buenos Aires: Tulio Guterman. (www.efdeportes.com).
- Arias, B. (2008). Desarrollo de un ejemplo de análisis factorial confirmatorio con LISREL, AMOS y SAS. *Seminario de Actualización en Investigación sobre Discapacidad SAID*. Mimeografía no publicada.
- Arnau, J. (1990). Metodología experimental. En J. Arnau, M. T. Anguera y J. Gómez, *Metodología de la Investigación en Ciencias del Comportamiento* (9-122). Murcia: Secretariado de Publicaciones de la Universidad de Murcia.
- Bentler, P. M. (1990). Comparative fit indexes in structural modeling. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. (1995). *Structural Equations Program Manual*. Encino, CA: Multivariate Software, INC.
- Bentler, P. M. y Bonnet, D. (1980). Significance and tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Berk, R. A. (1979). Generalizability of behavioral observations. A clarification of interobserver agreement and interobserver reliability. *American Journal of Mental Deficiency*, 83, 460-472.
- Blanco Villaseñor, A. (1991). La Teoría de la Generalizabilidad aplicada a diseños observacionales. *Revista Mexicana de Análisis de la Conducta/Mexican Journal of Behavior Analysis*, 14 (3), 23-64.
- Blanco Villaseñor, A. (1992). Aplicaciones de la Teoría de la Generalizabilidad en la selección de diseños evaluativos. *Bordón*, 43 (4), 431-459.

- Blanco Villaseñor, A. (1993). Fiabilidad, precisión, validez y generalización de los diseños observacionales. En M.T. Anguera (Ed.), *Metodología observacional en la investigación psicológica* (Vol 2: Fundamentación, pp 151-261). Barcelona: PPU.
- Blanco Villaseñor, A. y Anguera, M. T. (2003). Calidad de los datos registrados en el ámbito deportivo. En A. Hernández-Mendo, *Psicología del Deporte (Vol. II): Metodología* (pp.35- 73). Buenos Aires: Efdeportes.com.
- Blanco Villaseñor, A., Castellano, J., Hernández-Mendo, A., Sánchez-López, C. R., y Usabiaga, O. (2014). Aplicación de la TG en el deporte para el estudio de la fiabilidad, validez y estimación de la muestra. *Revista de Psicología del Deporte*, 23(1), 131-137.
- Blanco Villaseñor, A., Castellano, J. y Hernández Mendo, A. (2000). Generalizabilidad de las observaciones de la acción del juego en el fútbol. *Psicothema, España*, 12 (2), 81-86.
- Blanco Villaseñor, A. y Hernández Mendo, A. (1998). Estimación y generalización en un diseño de estructura espacial. En J. Sabucedo, R. García, E. Ares y D. Prada, Medio ambiente y responsabilidad humana (pp. 579-583). A Coruña: Libro de Comunicaciones. VI Congreso de Psicología Ambiental.
- Blanco Villaseñor, A. y Losada, J. L. (2004). Generalización en Diseños Observacionales: Alternativas de estimación y modelización. *Metodología de las Ciencias del Comportamiento, España, Suplemento, 2004*, 79-84.
- Blanco Villaseñor, A. y Morales Sánchez, V. (2010). Teoría de la Generalizabilidad: investigaciones aplicadas. En M. J. Blanca, R. Alarcón y D. López-Montiel

(Coords.) *XXI Congreso de Metodología de las Ciencias Sociales y de la Salud. Libro de Resúmenes (72)*. Málaga: CEES-IFV, UMA-Tecnolex y Asociación Malagueña de Estudios e Investigaciones Sociales.

Buendía, E. L., Colás, B. P. y Hernández, P. F. (1998). *Métodos de Investigación en Psicopedagogía*. Madrid: McGraw-Hill.

Cronbach, L. J., Rajaratnam, N. y Gleser, G. C. (1963). Theory of generalizability: a liberalization of reliability theory. *British Journal of Mathematical and Statistical Psychology*, 16, 137-163.

Cronbach, L. J., Gleser, G. C., Nanda, H. y Rajaratnam, N. (1972). The dependability of behavioral measurements: theory of generalizability for scores and profiles. Nueva York: John Wiley and Sons.

Gosling, S. D., Vazire, S., Srivastava, S., y John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions. *American Psychologist*, 59, 93-104.

Hernández Mendo, A. (2006). Un cuestionario para la evaluación psicológica de la ejecución deportiva: Estudio complementario entre TCT y TRI. *Revista de Psicología del Deporte*, 15(1).

Hernández-Mendo, A., Blanco-Villaseñor, A., Pastrana, J. L., Morales-Sánchez, V., Ramos-Pérez, F. J. (2016). SAGT: Aplicación informática para análisis de generalizabilidad. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 11(1), 77-89.

- Hernández-Mendo, A; Ramos-Pérez, F; y Pastrana, J. (2012). *SAGT: Programa informático para análisis de la teoría de la generalizabilidad*. SAFE CREATIVE. Código: 120491501059.
- Hernández-Mendo, A., Morales-Sánchez, V. y González Ruiz, S. L. (2012). Gestión de la calidad a través de la plataforma MEMPAS. *Cuadernos de Psicología del Deporte*, 12 (1), 147-150.
- Hu, L. T. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55.
- Jackson, D. L. (2007). The effect of the numbers of observations per parameter in misspecified confirmatory factor analytic models. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(1), 48-76.
- Jöreskog, K. G. (1970). A general method for the analysis of covariance structures. *Psichometrika*, 34, 183-202.
- Jöreskog, K. G. y Sörbom, D. (2007). *LISREL 8.80*. [Computer Software]. Lincolnwood, IL: Scientific Software International, Inc.
- Marôco, J. (2010). *Análise de Equacoes Estruturais. Fundamentos teóricos. Software y aplicacoes*. Pero Pinheiro: Rolo y Filhos II, SA.
- Martínez Arias, R. (1995). *Psicometría: teoría de los test psicológicos y educativos*. Madrid: Síntesis.
- Morales Sánchez, V. (2003). *Evaluación psicosocial de la calidad en servicios municipales deportivos: aportaciones desde el análisis de variabilidad*. Universidad de Málaga: Tesis Doctoral.

- Morales Sánchez, V. (2009). Evaluación de la calidad en organizaciones deportivas: análisis de generalizabilidad. *Revista de Psicología General y Aplicada, España*, 62(1-2), 99-109.
- Morales-Sánchez, V., Pérez-López, R., Morquecho-Sánchez, R., & Hernández-Mendo, A. (2016). Generalizabilidad y gestión deportiva. *Cuadernos de Psicología del Deporte*, 16(1), 161-170.
- Morquecho-Sánchez, R. (2014). *Evaluación de la Calidad Percibida en Organizaciones de Servicios Deportivos Universitarios*. Facultad de Organización Deportiva, Universidad Autónoma de Nuevo León: Tesis Doctoral.
- Muñiz, J. (2001). *Teoría clásica de los Test*. Madrid: Pirámide.
- Muthén, L. K. y Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén& Muthén.
- Lorenzo-Seva, U. y Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory Factor Analysis model. *Behavioral Research Methods, Instruments & Computers*, 38(1), 88-91.
- Lloret, S., Ferreres, A., Hernández, A. y Tomás, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30 (3), 1151-1169.
- Rao, C. R. (1997). *Variance components estimation. Mixed models, methodologies and applications*. London: Chapman Hall.
- Shavelson, R.J., & Webb, N.M. (1991). *Generalizability Theory: A Primer*. Newbury Park, CA: Sage Publications.

- Tanaka, J. S. y Huba, G. J. (1985). A fit index for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 38, 197-201.
- Thurstone, L. L. (1947). Multiple factor analysis. Chicago: The University Chicago Press.
- Visauta, B. (1998). *Análisis estadístico con SPSS para Windows. Volumen II: Estadística multivariante*. Madrid: McGraw-Hill.