

**DIRECTRICES
PARA LA ADAPTACIÓN Y
USO DE LOS TESTS**

Iñigo Vallo Ruiz

Índice de contenidos

Resumen

Introducción.....	1
Directrices de adaptación.....	4
Algunos comentarios a las directrices.....	9
Un modelo para evaluar la calidad de los tests utilizados en España	12
Proceso de evaluación.....	15
Descripción del Cuestionario para la Evaluación de los Tests (CET)	16
Cuestionario para la Evaluación de los Tests (CET).....	18
Descripción general del test	18
Valoración de las características del test.....	21
Valoración global del test.....	27
El País Vasco, una reflexión metodológica	29
Diversidad y especificidad de las situaciones de contacto de lenguas y de culturas	34
Los instrumentos de evaluación en el País Vasco.....	35
Adecuación del diseño estadístico.....	37
Un ejemplo concreto de adaptación: Adaptación al español del cuestionario Vecú et Sante Perçue de l'adolescent (VSP-A): Una medida genérica de calidad de vida para adolescentes	41
Resumen.....	41
Introducción	42
Sujetos y métodos	44
El cuestionario VSP-A	44
Adaptación de la versión española del VSP-A	44
Primera fase: traducción al español	46
Segunda fase: reunión y discusión en paneles de adolescentes y consenso del equipo de investigación	47
Tercera fase: retro-traducción al idioma original.....	48
Cuarta fase: factibilidad de la administración y comprensión de la versión pre-test.....	48

Resultados	48
Discusión	50
Directrices Internacionales para el uso de los test	58
Objetivo.....	58
Ámbito de aplicación	58
Destinatarios	59
Aspectos contextuales	60
Conocimientos	61
Uso ético de los tests	62
Utilización adecuada de los tests	64
Apéndice A: Directrices para el establecimiento de políticas sobre el uso de los tests	73
Apéndice B: Directrices para desarrollar contratos entre las partes implicadas en la evaluación	75
Apéndice C: Aplicación de los tests a personas con alguna discapacidad	77
Conclusiones.....	80
Normas mínimas para el uso de los tests	82
Comisión de Tests del Colegio Oficial de Psicólogos.....	83
Referencias bibliográficas	84

Resumen

Las directrices tienen como objetivo fundamental mejorar el uso que los profesionales hacen de los tests, siendo elaboradas originalmente por la Comisión Internacional de Tests (ITC), bajo la dirección del profesor Dave Bartram. La versión en castellano, ha sido realizada en el seno de la Comisión de Tests del Colegio Oficial de Psicólogos (COP). El Colegio, a través de sus representantes en la citada Comisión Internacional de Tests y en la Comisión Europea sobre Tests de la Federación Europea de Asociaciones Profesionales de Psicólogos (EFPPA), ha respaldado las directrices y ha participado activamente en la elaboración de las mismas, consciente de la necesidad de contribuir a mejorar la práctica de los tests en nuestro país. Un medio eficaz para mejorar los tests y su práctica, es ofrecer a los usuarios toda la información posible acerca de su calidad y características. El Colegio Oficial de Psicólogos, tiene la intención de promover un proceso de evaluación de los tests más utilizados en España, para proporcionar así una información técnica y precisa a los usuarios de las pruebas, creando el CET Cuestionario de Evaluación de los Tests. En las sucesivas páginas, se describen las características del procedimiento de evaluación y se presenta el modelo de cuestionario para llevarlo a cabo. En España, la construcción, traducción, adaptación y utilización de instrumentos de medida en el ámbito de las Ciencias Sociales en general, y en el ámbito de la Psicología en particular, se encuentran asociados a una problemática que requiere una respuesta metodológica específica. Por consiguiente, afrontar la discusión en términos de metodología aplicada, es la única solución que permite que los bilingües puedan disponer de instrumentos de medida adecuados a cada una de las realidades que la situación de contacto de lenguas y culturas genera. En Europa, durante la última década, se han desarrollado diversas medidas de calidad de vida relacionada con la salud para uso exclusivo en niños/as y adolescentes. No obstante, existen pocos instrumentos de estas características adaptados en España. Se puede tomar como ejemplo El Vecú et Sante Perçue de L'adolescent (VSP-A), el cual es un instrumento genérico de calidad de vida relacionada con la salud para adolescentes de 11 a 17 años.

Introducción

Los tests constituyen uno de los instrumentos de medida más utilizados por los psicólogos en el ejercicio de la profesión. Como cualquier otro instrumento de medida, en otras áreas científicas, los tests pueden utilizarse de forma correcta o incorrecta, siendo especialmente grave en psicología y/o educación, donde con frecuencia están implicadas decisiones de gran importancia para la vida de las personas. La mayoría de los problemas en torno a los test, provienen de su uso inadecuado, más que del test. En sí, de su construcción, o de sus propiedades técnicas. Y éste uso inadecuado es el responsable de muchas de las críticas que reciben los test, cuestionarios y escalas psicológicas. Pero en realidad, los test, siguen siendo un instrumento imprescindible en la práctica psicológica y educativa en todo el mundo, y lo seguirá siendo en el futuro, por tanto, no hay que ahorrar esfuerzos a la hora de mejorar su uso. Las propias casas editoras de test, son las primeras interesadas en un uso deontológico apropiado, los Colegios Profesionales, suelen tomar algunas al respecto, como restringir el acceso a los tests, en función de su especialización, o registrar oficialmente a los usuarios cualificados.

La globalización mundial que se produce a todos los niveles y esferas de la actividad humana, y las rápidas comunicaciones, potencian que los instrumentos de medida generados en un país, rápidamente puedan utilizarse en otro distinto. Los test, deben adaptarse cuidadosamente en función de las diferencias interculturales entre el idioma/cultura originario y aquellos en los que se pretende utilizar, en el caso de España, hay un alto porcentaje de tests, y escalas que provienen de traducciones y adaptaciones de otros idiomas y culturas. Traducir un test, no es una cuestión de saber dos idiomas. Es algo más complejo, por tanto más que traducción habría que hablar de adaptación. El Colegio Oficial de Psicólogos (COP) participó en un Comité Internacional auspiciado por la Comisión Internacional de Test (ITC), de la cual es miembro, para preparar unas directrices técnicas, para la adaptación de los test, de unos países a otros.

Las directrices que se proponían, en ese documento, tienen como objetivo fundamental mejorar el uso que los profesionales hacen de los tests y han sido elaboradas originalmente por la Comisión Internacional de Tests (ITC), bajo la dirección del profesor Dave Bartram. La versión en castellano, ha sido realizada, en el seno de la Comisión de Tests del Colegio Oficial de Psicólogos (COP). El Colegio, a través de sus representantes en la citada Comisión Internacional de Tests y en la Comisión Europea sobre Tests de la Federación Europea de Asociaciones Profesionales de Psicólogos (EFPPA), que además de respaldar las directrices, ha participado activamente en la elaboración de las mismas, consciente de la necesidad de contribuir a mejorar la práctica de los tests en nuestro país.

Los documentos más utilizados en la elaboración de las directrices, junto con otros en castellano, aparecen citados al final; documentación complementaria en castellano puede consultarse en la sección de la Comisión de Tests incluida en la página web del COP (<http://www.cop.es/tests/>).

Las directrices reflejan principios generales implicados en el uso adecuado de los tests, pero no pretenden uniformar las diferencias legítimas existentes entre la práctica en los diferentes países o áreas profesionales. Existen bastantes razones por las que se necesitan unas directrices para el uso de los tests con carácter internacional. Veamos algunas de ellas.

- a) Los países difieren ampliamente en el grado de control legal que pueden ejercer sobre el uso de los tests y sus consecuencias para las personas evaluadas. Por tanto, la existencia de un conjunto de directrices internacionalmente aceptadas proporcionará a las asociaciones nacionales de psicología y otras organizaciones profesionales una buena base documental para completar directrices ya existentes o proceder a crearlas.
- b) El derecho a comprar y utilizar materiales psicométricos varía mucho de unos países a otros. En algunos países el uso de los tests está restringido a los psicólogos, en otros a quienes estén registrados por los distribuidores de las

pruebas, e incluso en algunos los usuarios pueden obtener el material psicométrico libremente, sin restricciones.

- c) Algunos tests bien conocidos han aparecido en Internet sin la autorización de autores y editores, violando los derechos de copyright, y sin ninguna precaución en relación con la seguridad del test.
- d) En el ámbito de la psicología del trabajo, el aumento de la movilidad internacional de los trabajadores ha incrementado la demanda de tests para utilizarlos con aspirantes de diferentes países. A menudo los tests son aplicados en un país para empresas ubicadas en otros países.
- e) En países como Estados Unidos e Inglaterra, entre otros, se están llevando a cabo trabajos sobre evaluación a distancia vía Internet en las áreas de Trabajo y Educativa. Esto plantea todo un conjunto de problemas en relación con las normas de aplicación, control del proceso de evaluación y seguridad de los resultados.

Por todas estas razones, y otras que se podrían enumerar, parece conveniente disponer de unas directrices internacionales que permita una referencia común a la hora de establecer normas nacionales para la mejora del uso de los tests.

Las presentes directrices recogen el trabajo de especialistas en evaluación psicológica y educativa (psicólogos, psicómetras, editores y constructores de tests) de diferentes países, entre ellos España. La idea no ha sido tanto el inventar unas nuevas directrices, como juntar las hebras comunes que se encuentran en las distintas directrices, códigos de práctica y estándares ya existentes, formando una estructura coherente que resulte comprensible y de uso práctico.

DIRECTRICES DE ADAPTACIÓN

Hay varias razones para adaptar un test, una de ellas es que adaptar un test es más barato y rápido que preparar uno nuevo para un segundo idioma, los tests se adaptan a veces para mejorar la imparcialidad de la evaluación, permitiendo a las personas su uso, junto con las escalas psicológicas en su idioma; los métodos para llevarlo a cabo y establecer equivalencias entre las puntuaciones, es tarea más ardua. Algunos investigadores interculturales han señalado que un alto porcentaje de investigaciones son defectuosas, incluso, no válidas, debido a una adaptación deficiente de los test utilizados.

Con respecto a las fuentes de error, los estándares sobre los tests psicológicos y educativos elaborados en 1985 por las asociaciones americanas American Educational Research Association (AERA), American Psychological Association (APA) y National Council on Measurement in Education (NCME), proporcionan cuidadosas directrices para aquellos psicólogos que seleccionan, construyen, aplican y usan tests psicológicos y educativos. Tres de los estándares son especialmente relevantes en el contexto de la aplicación y el uso:

Estándar 6.2. Cuando se hacen cambios sustanciales en el formato del test, modo de aplicación, instrucciones, idioma o contenido, el usuario debería de revalidar el test para las nuevas condiciones, o tener argumentos que apoyen que no es necesaria o posible una validación adicional.

Estándar 13.4. Cuando se traduce un test de un idioma o dialecto a otro, debe de establecer su fiabilidad y validez en los nuevos grupos lingüísticos a los que se aplique.

Estándar 13.6. Cuando se pretende que las dos versiones de un test en idiomas distintos sean comparables, hay que aportar pruebas acerca de la compatibilidad de los tests.

Estos estándares, proporcionan un marco para analizar las fuentes de error o invalidez, que pueden surgir en la adaptación y pueden organizarse en varias categorías como son, por un lado las diferencias culturales y de idioma, los aspectos técnicos y métodos y por último, la interpretación de los resultados. Ésta es la razón fundamental por la que la Comisión Internacional de Tests, se planteó la tarea de elaborar un documento técnico al respecto, donde hay cuatro grandes áreas implicadas en la confección de un test, en las cuales se pueden tomar decisiones erróneas como son:

- Contexto
- Construcción y adaptación
- Aplicación
- Interpretación de las puntuaciones

- Contexto:

Se refiere al ámbito sociocultural en el que están enmarcadas las personas a las que se aplicará el test. La mayoría de los constructos utilizados en Psicología son altamente dependientes de los aspectos culturales, por tanto, lo primero de lo que hay que cerciorarse es que el constructo a evaluar es equiparable en ambas culturas. Si no se da esa equivalencia razonable carece de sentido la adaptación, y si se hace se estarán midiendo constructos distintos en cada cultura. Hay que aportar datos empíricos de la equivalencia, para lo cual pueden llevarse a cabo estudios sobre el terreno por equipos de personas multiculturales y multilingües para evitar un sesgo etnocéntrico. Por ejemplo, evaluación de variables como inteligencia, actitudes, o personalidad, resultan altamente dependientes de la matriz cultural, por lo que su utilización en estudios comparativos requiere sumo cuidado al adaptar la prueba. Si existen datos acerca de la equivalencia de los constructos, puede pasarse a la fase de la adaptación propiamente dicha, que medirá las conductas concretas para evaluar el constructo.

- Construcción y adaptación del test:

La filosofía general que ha de guiar la adaptación, es que el test mida el mismo constructo de la misma forma, pero ajustándose a las peculiaridades de la nueva población. Naturalmente esto no resulta nada fácil. La primera decisión difícil es la selección de los traductores. Es fundamental que los traductores no sólo conozcan ambos idiomas, sino que estén familiarizados también con ambas culturas. En numerosas situaciones, hay que cambiar los ítems casi completamente, y eso sólo se puede hacer desde un conocimiento profundo de la cultura objetivo, y desde un nuevo entrenamiento riguroso para los aplicadores.

- Interpretación:

En general, interpretar los resultados de cualquier test, requiere un psicólogo especialista en el área del test, con la formación psicométrica apropiada. Y esto sigue siendo válido cuando el test, proviene de una adaptación. Lo que ocurre es que con los tests adaptados hay un peligro adicional, a saber, la tentación de comparar grupos o países en el test y establecer una clasificación de ganadores y perdedores. Como señalan algunos autores (Hambleton y Bollwark, 1991; Wesbury, 1992), los estudios comparativos deberían de usarse para comprender las semejanzas y diferencias entre los grupos analizados, pero nunca para establecer comparaciones sin más. Y no es adecuado establecerlas, porque raramente encontraremos dos comunidades que sean equiparables completamente, en aspectos tan influyentes como la motivación a la hora de hacer las pruebas, currícula escolares, valores culturales, nivel de vida, políticas educativas, oportunidades de acceso a la educación, etc.

Finalmente, para interpretar las puntuaciones el psicólogo debe de disponer de una documentación exhaustiva acerca de cómo se llevó a cabo el proceso de adaptación. El manual del test deberá incluir todo tipo de detalles del proceso adaptativo, que en determinadas circunstancias pueden dar las claves interpretativas de un resultado.

Para tratar de minimizar los errores en las cuatro grandes áreas citadas (contexto, adaptación, aplicación e interpretación), el Comité de la Comisión Internacional de Tests (ITC), propuso recomendar veintidós directrices que se consideran claves para una correcta adaptación o construcción simultánea de los tests psicológicos y educativos.

- Contexto

C.1. Los efectos de las diferencias culturales que no sean relevantes para los objetivos centrales del estudio deberían minimizarse en la medida de lo posible.

C.2. Debería de evaluarse la cuantía del solapamiento de los constructos en las poblaciones de interés.

- Adaptación del test

Los constructores/editores de test deberían:

D.1. Asegurar que el proceso de adaptación tiene en cuenta las diferencias lingüísticas y culturales entre las poblaciones a las que se dirigen las versiones adaptadas del test.

D.2. Proporcionar datos que garanticen que el lenguaje utilizado en las instrucciones, en los propios ítems y en el manual del test, son apropiados para todas las poblaciones culturales e idiomáticas a las que va dirigido el test.

D.3. Aportar evidencias de que las técnicas de evaluación elegidas, los formatos de los ítems, las reglas de los tests, y los procedimientos son familiares a todas las poblaciones a las que van dirigidos.

D.4. Facilitar evidencias de que el contenido de los ítems y los materiales de los estímulos son familiares para todas las poblaciones a las que van dirigidos.

D.5. Aportar una justificación racional sistemática, tanto lingüística como psicológica, para mejorar la precisión del proceso de adaptación, así como reunir datos acerca de la equivalencia de todas las versiones en los distintos idiomas.

D.6. Asegurar que el diseño de recogida de datos permite el uso de técnicas estadísticas apropiadas para establecer la equivalencia entre los ítems correspondientes a las diferentes versiones idiomáticas del test

D.7. Aplicar técnicas estadísticas apropiadas para 1) establecer la equivalencia entre las diferentes versiones de un test, y 2) identificar componentes problemáticos o aspectos del test que puedan ser inadecuados para alguna de las poblaciones a las que va destinado el test.

D.8. Proporcionar información sobre la evaluación de la validez en todas las poblaciones objetivo a las que va dirigido el test adaptado.

D.9. Aportar datos estadísticos sobre la equivalencia de los tests para todas las poblaciones a las que van dirigidos.

D.10. No deben utilizarse preguntas no equivalentes en todas las versiones dirigidas a diferentes poblaciones cuando se prepara una escala común, o cuando se comparan estas poblaciones. Sin embargo, pueden ser útiles para reforzar la validez de contenido de las puntuaciones de cada población por separado.

- Aplicación

A.1. Los constructores y los aplicadores de los tests deberían tratar de prever los tipos de problemas que cabe esperar, y tomar las medidas oportunas para evitarlos mediante la preparación de materiales e instrucciones adecuados.

A.2. Quienes aplican los tests deberían de ser sensibles a cierto número de factores relacionados con los materiales utilizados para los estímulos, los procedimientos de aplicación, y las formas de respuesta, que pueden reducir la validez de las inferencias extraídas de las puntuaciones.

A.3. Aquellos aspectos del entorno que influyen en la aplicación del test deberían de mantenerse lo más parecidos posible para todas las poblaciones a las que va dirigido el test.

A.4. Las instrucciones para la aplicación del test en el idioma fuente y en el objetivo deben minimizar la influencia de fuentes de variación no deseadas.

A.5. El manual del test debería de especificar todos los aspectos del test y de su aplicación que han de revisarse al utilizarlo en un nuevo contexto cultural.

A.6. El aplicador no debe de interferir, debiendo minimizarse su influencia sobre los examinados. Deben de seguirse al pie de la letra las reglas explícitas descritas en el manual del test.

- Interpretación de las puntuaciones

I.1. Cuando se adapta un test para utilizarlo en otra población, debe de facilitarse la documentación sobre los cambios, así como los datos acerca de la equivalencia entre las versiones.

I.2. Las diferencias entre las puntuaciones obtenidas por las muestras a las que se aplicó el test no deben de tomarse sin más directamente. El investigador tiene la responsabilidad de sustanciar las diferencias con otros datos empíricos.

I.3. Las comparaciones entre poblaciones sólo pueden hacerse al nivel de la invarianza que se haya establecido para la escala en la que se expresan las puntuaciones.

I.4. El constructor del test debería de proporcionar información específica acerca de las distintas formas en las que los contextos socioculturales y ecológicos de las poblaciones pueden afectar al rendimiento en el test, y debería sugerir procedimientos para tener en cuenta estos efectos en la interpretación de los resultados.

Algunos comentarios a las directrices

En el documento final que la Comisión Internacional de Tests ha editado con las directrices, cada una de ellas lleva una explicación y fundamentación de su sentido, se indican los pasos a seguir para cumplirla, se señalan los errores más frecuentes y se sugieren referencias y documentación complementarias. Aquí nos limitaremos a hacer algunos comentarios generales. Para una exposición más detallada en castellano véase Hambleton (1996).

- Contexto.

Las dos primeras directrices tratan de asegurar que haya una equivalencia de los constructos medidos en las dos poblaciones de interés. A medida que la distancia cultural entre las dos (o más) poblaciones de interés aumenta, más probabilidades hay de que existan diferencias entre los constructos. Por ejemplo, una variable tan clásica como la inteligencia tiende a ser conceptualizada de muy diverso modo de unas culturas a otras; mientras en Occidente suele asociarse a eficiencia y rapidez,

en algunas sociedades del Este se asocia con actitudes reflexivas y reposadas, más que con rapidez (Lonner, 1990). La primera directriz hace hincapié en la necesidad de minimizar la incidencia de cualquier aspecto ajeno al constructo medido. Por ejemplo, en algunos estudios internacionales recientes (Wainer, 1993) se observó que las diferencias en la motivación podían estar a la base de algunas de las diferencias halladas.

- Adaptación.

Una vez que los constructores del test han aportado los datos que garantizan que hay una equivalencia apropiada entre los constructos a medir en ambas poblaciones, y por tanto tiene sentido adaptar el test, empieza el proceso real de su construcción. Este proceso une a la problemática psicométrica general para construir un test, la dificultad añadida de plasmar fielmente el test original, si se trata de una adaptación; podría darse el caso también de una doble construcción simultánea en las dos (o más) poblaciones de interés. El meollo central de la construcción/adaptación gira en torno al funcionamiento adecuado del equipo de traductores. Amén de la posibilidad de utilizar diseños técnicos sofisticados, con la utilización de monolingües y bilingües para asegurarse de las equivalencias, es clave que los traductores conozcan las culturas implicadas además de los idiomas (Bracken y Barona, 1991; Brislin, 1986; Hambleton y Kanjee, 1995; Prieto, 1992). Tal vez algunos puristas apuntarán que no se puede dominar un idioma sin hacerlo también con la cultura y costumbres, y seguramente tienen razón, pero hay que asegurarse de ello. Es imprescindible que los traductores, aún sin ser expertos, conozcan las reglas básicas de la construcción de los tests. Por poner un solo ejemplo, los traductores tienen que saber que cuando se utilizan ítems de elección múltiple la longitud de las alternativas tiene que ser aproximadamente la misma. Una buena exposición en castellano sobre la construcción de los ítems puede consultarse en Prieto y Delgado (1996). Las directrices subrayan la necesidad de asegurarse que los contenidos, formatos, estímulos, gráficos, diagramas, figuras, unidades de medida, y signos en general, son igualmente familiares para las culturas implicadas.

Otro aspecto clave subrayado en varias de las directrices, es que los constructores recojan datos de todo tipo, que permitan evaluar la validez y la comparación de las dos versiones. La tecnología estadística es actualmente potente y variada para analizar los datos, pero no hay que olvidar que cuanto más amplias sean las muestras utilizadas con más eficacia funcionan las técnicas estadísticas. Estas técnicas son complementarias de los correspondientes análisis racionales, no los sustituyen. En los últimos años se ha perfeccionado notablemente la tecnología para evaluar el sesgo de las puntuaciones (Camilli y Shepard, 1994; Fidalgo, 1996; Holland y Wainer, 1993), pero sigue siendo de suma importancia el análisis racional por parte de un equipo de expertos en las culturas implicadas. Lo que las técnicas detectan, es que cierto ítem funciona de modo diferente en un grupo que en otro, por eso suelen agruparse bajo el nombre genérico de Funcionamiento Diferencial de los Ítems, será el psicólogo quien tenga que analizar las razones de ese funcionamiento diferencial. En síntesis, hay que garantizar tanto por medios estadísticos como analítico-racionales, que el test, es igualmente válido para las poblaciones de interés.

- Aplicación.

Las seis directrices correspondientes a la aplicación de los tests tratan de subrayar aquellos aspectos claves para una aplicación correcta de los tests. A veces se descuida la aplicación, confiándola a personas sin la debida formación. Esto constituye un grave error, pues una aplicación defectuosa afecta directamente a la validez del test. Si esto ocurre en general, cuando se trata de un test adaptado hay que extremar las precauciones, asegurándose que las condiciones de aplicación no constituyen una fuente de error respecto del test original. El peligro es mayor cuanto mayor es la distancia cultural entre las poblaciones origen y objetivo. El aplicador no debe de interferir ni con su conducta, apariencia, instrucciones, etc., en el proceso estrictamente reglado de la aplicación del test. Los constructores de los tests deben de asegurarse mediante una selección y entrenamiento adecuados que los aplicadores actúan correctamente. Suele ser muy útil llevar a cabo algún estudio piloto para comprobarlo.

- Interpretación de las puntuaciones.

En el paso final correspondiente a la interpretación de las puntuaciones puede echarse a perder toda la labor anterior de construcción y aplicación del test si no se actúa adecuadamente. En primer lugar, las puntuaciones de los tests siempre son datos que deben de utilizarse convergentemente con otros en la toma de decisiones. Ayudará notablemente a la interpretación si se dispone de la documentación apropiada sobre todo el proceso de construcción y aplicación, diseños utilizados, métodos para la evaluación de la equivalencia entre las dos versiones, ítems que se han modificado y por qué, etc. Así y todo se recomienda prudencia extrema a la hora de interpretar las diferencias halladas entre poblaciones distintas. No se pueden interpretar directamente sin más, hay que fundamentar esas diferencias, cuyas posibles causas pueden ser múltiples. Por ejemplo, cuando se compara el rendimiento académico de estudiantes de países desarrollados con otros del tercer mundo, pueden atribuirse simplista e irresponsablemente las diferencias a diferencias reales en las correspondientes aptitudes, cuando en realidad existen toda una serie de factores diferenciales tales como las políticas educativas, los currícula, el acceso a la escolarización, equipamientos y recursos educativos disponibles, idioma materno frente al escolar, etc., por citar sólo algunos ejemplos. Las directrices subrayan la necesidad de una prudencia suma a la hora de interpretar las diferencias. La tecnología psicométrica desarrollada bajo el enfoque de la teoría de respuesta a los ítems (Hambleton et al., 1991; Muñiz, 1996b; Navas, 1996), permite establecer equiparaciones entre las puntuaciones con bastante precisión.

UN MODELO PARA EVALUAR LA CALIDAD DE LOS TESTS UTILIZADOS EN ESPAÑA

En el ejercicio de su profesión los psicólogos utilizan con frecuencia los tests para la obtención de datos. Para que los datos así obtenidos sean de calidad y puedan ayudar al psicólogo a tomar las decisiones adecuadas, los tests utilizados han de reunir las propiedades técnicas oportunas. Un medio eficaz para mejorar los tests y su práctica es ofrecer a los usuarios toda la información posible acerca de su calidad y

características. El Colegio Oficial de Psicólogos tiene la intención de promover un proceso de evaluación de los tests más utilizados en España para proporcionar así una información técnica precisa a los usuarios de las pruebas. Mas abajo se describen las características del procedimiento de evaluación y se presenta un modelo de cuestionario para llevar a cabo la evaluación.

Los tests psicológicos, como se ha comentado, constituyen una de las herramientas más importantes al servicio de la práctica profesional y de la investigación de los psicólogos. Como ocurre con cualquier otra tecnología científica, los tests pueden ser usados de forma correcta o incorrecta. Una exposición de las actitudes de los psicólogos españoles hacia los tests, los problemas más habituales de uso, así como las pruebas más utilizadas, puede consultarse en el trabajo de Muñiz y Fernández-Hermida (2000), Muchos psicólogos ponen de manifiesto la necesidad de mejorar el uso que se hace de los tests en nuestro país. Este objetivo es compartido por buena parte de las asociaciones profesionales de psicólogos, tanto nacionales como internacionales. El Colegio Oficial de Psicólogos (COP) no ha sido ajeno a esta corriente, y en el año 1995 se puso en marcha una Comisión de Tests integrada por psicólogos procedentes del ámbito académico, de las editoras de tests y del propio COP. La Comisión de Tests fue creada para tratar de promocionar y potenciar el uso adecuado de los tests en nuestro país. Para llevar a cabo esta tarea, la Comisión funciona de forma coordinada con otros grupos de trabajo internacionales con fines similares, tales como la Federación Europea de Asociaciones de Psicólogos Profesionales (EFPPA), o la Comisión Internacional de Tests (ITC).

En general, podrían seguirse tres vías para mejorar el uso de los tests: (i) restringir el acceso de los usuarios, permitiendo su empleo solamente a profesionales debidamente capacitados y acreditados, (ii) incrementar la formación de los usuarios y (iii) fomentar la investigación para desarrollar pruebas eficientes (construcción de nuevas pruebas, adaptación adecuada de las desarrolladas en otros países y revisión periódica de la calidad técnica de las ya disponibles). Aunque lo ideal sería actuar en los tres frentes, en ocasiones la primera vía no está exenta de dificultades, puesto que las asociaciones profesionales no disponen de las armas legales para introducir fuertes

restricciones al uso de los instrumentos. Por ello, parece más viable mejorar la práctica profesional a través de la formación de los usuarios, puesto que las deficiencias en los conocimientos técnicos requeridos son consideradas como la principal causa de las prácticas inadecuadas (Anastasi, 1987; Kaufman, 1982; Muñiz et al., 1999; Muñiz y Fernández-Hermida, 2000). Además, la mejora de la formación, favorece la tercera de las vías antes mencionadas.

En algunos países de nuestro entorno se ha seguido el principio de que facilitar información a los usuarios acerca de la calidad técnica de los tests disponibles es uno de los principales recursos formativos. Así lo han entendido, por ejemplo, las asociaciones profesionales de psicólogos de Holanda y del Reino Unido. Ambas asociaciones publican periódicamente monografías en las que se aporta información técnica de los principales tests usados en esos países (Bartram et al., 1992a, 1992b, 1995 y 1997; Evers, 1996). En estas publicaciones se difunden los resultados de evaluaciones anónimas efectuadas por expertos siguiendo un proceso estandarizado de revisión de los tests.

En nuestro país las informaciones acerca de las características de los tests no son exhaustivas, están muy dispersas, no son fácilmente accesibles o, al ser facilitadas por el editor del test, no pueden ser consideradas independientes e imparciales. Por este motivo, los profesionales carecen en muchas ocasiones de orientaciones científicas que les permitan seleccionar el instrumento más apropiado para sus objetivos. Consciente de este problema, la Comisión de Tests del COP ha diseñado un modelo estandarizado de evaluación de tests (CET) con la finalidad de revisar los tests empleados en nuestro país, al objeto de informar a los usuarios sobre su calidad técnica. Puesto que se dispone de datos recientes bastante fiables sobre los tests más usados en los distintos campos profesionales de la psicología (Muñiz y Fernández-Hermida, 2000), es intención de la Comisión de Tests focalizar inicialmente en estas pruebas las revisiones, aunque también sería posible incluir otros tests en la primera fase de evaluación a instancias de editores y autores

Tanto el sistema de evaluación que se describe a continuación, como el modelo de cuestionario presentado, se han inspirado principalmente en los modelos inglés y holandés actualmente en funcionamiento.

Proceso de evaluación

Se presenta a continuación lo que podría ser en líneas generales el proceso a seguir para llevar a cabo la evaluación de los tests. El primer paso que hay que dar es elegir los tests a revisar. La idea es que sea un proceso continuo, hasta llegar a revisar la mayoría, si no todas, las pruebas que hay en el mercado, si bien parece razonable empezar por aquellas pruebas más utilizadas por los profesionales.

Otro aspecto a resolver es la selección de los revisores para la evaluación de cada test. Los revisores habrán de a) ser expertos en metodología psicométrica y/o en el campo profesional de evaluación al que el test va dirigido. La experiencia obtenida en otros países sugiere que uno de los revisores sea experto en psicometría y el otro en la temática que se pretende evaluar mediante el test (Evers, 1996); b) no estar implicados en la construcción o distribución del test, y c) realizar la evaluación de forma anónima e independiente.

Cada revisor recibiría un ejemplar del test (cuadernillo, hoja de respuesta, diskete, etc), un ejemplar de la documentación aportada por el editor o autor (manual, informes técnicos, trabajos científicos, etc.), y el cuestionario para la evaluación, en formato impreso e informatizado.

Si no existe un acuerdo sustancial entre los revisores, se solicitará a un tercer revisor que lleve a cabo la evaluación de la prueba. Antes de dar publicidad a los resultados de la evaluación, se informará al autor y/o al editor del test para que efectúen las observaciones que consideren oportunas, aporten información complementaria, etc. Finalmente se dará publicidad a los resultados.

Este proceso descrito es similar al que se lleva a cabo actualmente en otros países como Holanda, Inglaterra o Estados Unidos, y coincide en su filosofía general con el que utilizan habitualmente las revistas científicas para la revisión de los artículos. Cualquier sistema de evaluación que se utilice ha de tener necesariamente limitaciones, pero el arriba descrito, parece que mejora significativamente el modelo actual en nuestro país, consistente en no hacer nada.

Descripción del Cuestionario para la Evaluación de los Tests (CET)

Para construir la versión final del cuestionario (CET) se han tomado en consideración los comentarios y sugerencias a una versión inicial, recibidas de los miembros de la Comisión de Tests del COP y de diecinueve expertos de los ámbitos académico y profesional que se relacionan al final. El CET ha sido diseñado para analizar fundamentalmente instrumentos psicométricos basados en la Teoría Clásica de los Tests, puesto que por el momento son los de más frecuente uso en la práctica profesional en España. Está estructurado en tres apartados; el primero tiene la finalidad de elaborar una descripción técnica del test. Está integrado por 31 ítems relativos al nombre de la prueba, autor, constructo medido, ámbito de aplicación, etc. En el segundo apartado se incluye la evaluación técnica de las características del instrumento. Los expertos han de juzgar características como la fundamentación teórica, la adaptación/traducción (si el test ha sido construido en otro país), la fiabilidad, la validez, los baremos, etc. Para lograr este objetivo, se han incluido 32 ítems cerrados y 6 abiertos. En la mayor parte de los ítems cerrados se propone un sistema de cinco categorías ordenadas en función de la calidad de la característica evaluada. Este sistema de cinco estrellas es similar al empleado por el Colegio de Psicólogos del Reino Unido (Bartram et al., 1997).

En los ítems abiertos se solicita una justificación razonada de las respuestas a los ítems cerrados y una evaluación de cada característica.

En último apartado, se solicita una valoración global del test y un resumen de los dos primeros apartados, al objeto de resumir toda la información en una ficha técnica. A continuación, se presenta el cuestionario completo.

Un profesional con una buena información sobre las pruebas y una formación adecuada difícilmente utilizará de forma incorrecta los tests. Enmarcado en esa filosofía, en 1995 el Colegio Oficial de Psicólogos (COP) creó una Comisión de Tests con el fin de analizar los problemas implicados en el uso de los tests. En el presente trabajo se describe el proceso diseñado por la Comisión de Tests del COP para llevar a cabo la revisión de los tests editados en España. También se presenta el cuestionario elaborado a tal efecto. En líneas generales, el procedimiento de evaluación propuesto consiste en revisar los tests de forma sistemática por parte de expertos, utilizando el cuestionario de recogida de datos diseñado a tal efecto, inspirado en los modelos holandés e inglés actualmente en funcionamiento. Cabe esperar que la publicación de los resultados de las revisiones de los tests ayude a los profesionales a elegir y evaluar las pruebas que tienen a su disposición en el mercado, al menos ése ha sido el objetivo que ha guiado a la comisión de tests del COP para llevar a cabo esta iniciativa. Señalar, finalmente, que cualquier sugerencia que pueda mejorar el proceso de evaluación expuesto será bien recibida por los autores, que la elevarán a la comisión de tests del COP.

CUESTIONARIO PARA LA EVALUACION DE LOS TESTS (CET)

1. Descripción general del test

1.1. Nombre del test:

1.2. Nombre del test en su versión original (si la versión española es una adaptación):

1.3. Autor/es del test original:

1.4. Autor/es de la adaptación española:

1.5. Editor del test en su versión original:

1.6. Editor de la adaptación española:

1.7. Fecha de publicación del test original:

1.8. Fecha de publicación del test en su adaptación española:

1.9. Fecha de la última revisión del test en su adaptación española:

1.10. Clasifique el área general de la o las variables que pretende medir el test.

- Inteligencia
- Aptitudes
- Habilidades y Rendimiento académico
- Psicomotricidad
- Alteraciones neuropsicológicas
- Personalidad
- Motivación
- Actitudes
- Intereses
- Valores
- Otros (Indique cuál:.....)

1.11. Breve descripción de la variable o variables que pretende medir el test:

1.12. Área de aplicación

- Psicología clínica
- Psicología educativa
- Neuropsicología
- Psicología forense
- Psicología del trabajo y las organizaciones
- Psicología del deporte
- Servicios sociales
- Psicología del Tráfico
- Otros (Indique cuál:.....)

1.13. Formato de los ítems:

- Respuesta libre
- Respuesta dicotómica (si/no, verdadero/falso, etc)
- Elección múltiple
- Tipo Likert
- Adjetivos bipolares
- Otro (Indique cuál:.....)

1.14. Número de ítems:

1.15. Soporte:

- Administración oral
- Papel y lápiz
- Manipulativo
- Informatizado
- Otro (Indique cuál:.....)

1.16. Cualificación requerida para el uso del test de acuerdo con la documentación aportada:

- Ninguna
- Entrenamiento y Acreditación específica*
- Nivel A
- Nivel B
- Nivel C
- Otra (Indique cuál:.....)

*Indique el nombre de la institución que lleva a cabo la acreditación:

1.17. Descripción de las poblaciones a las que el test es aplicable (especifique el rango de edad, nivel educativo, etc., y si el test es aplicable en ciertas poblaciones específicas: minorías étnicas, discapacitados, grupos clínicos, etc.):

1.18. Indique si existen diferentes formas del test y sus características (formas paralelas, versiones abreviadas, versiones informatizadas o impresas, etc). En el caso de que existan versiones informatizadas, describa los requisitos mínimos del hardware y software.

1.19. Procedimiento de corrección:

- Manual mediante plantilla
- Lectora óptica
- Automatizada por ordenador
- Efectuado exclusivamente por la empresa suministradora
- Mediante expertos
- Otro (Indique cuál:.....).

1.20. Puntuaciones: (Describa el procedimiento para obtener las puntuaciones directas).

1.21. Transformación de las puntuaciones:

- Característica no aplicable para este instrumento
- Normalizada
- No normalizada

1.22. Escalas utilizadas:

- Centiles
- Puntuaciones típicas
- Cocientes de desviación
- Eneatipos
- Decatipos
- T
- Otra (Indique cuál:.....)

1.23. Posibilidad de obtener informes automatizados:

- No
- Si*

*Breve descripción:

1.24. El editor ofrece un servicio para la corrección y/o elaboración de informes:

- No
- Si

1.25. Tiempo estimado para la aplicación del test (instrucciones, ejemplos y respuestas a los ítems).

En aplicación individual:.....
En aplicación colectiva:.....

1.26. Documentación aportada por el editor:

- Manual
- Libros o artículos complementarios
- Disketes/CD
- Otra (Indique cuál:.....)

1.27. Precio de un juego completo de la prueba (documentación, test, plantillas de corrección; en el caso de tests informatizados no se incluye el costo del hardware):

1.28. Precio y número de ejemplares del paquete de cuadernillos (tests de papel y lápiz):

1.29. Precio y número de ejemplares del paquete de hojas de respuesta (tests de papel y lápiz):

1.30. Precio de la corrección y/o elaboración de informes por parte del editor:

1.31. Bibliografía básica acerca del test aportada en la documentación:

2. Valoración de las características del test

2.1. Calidad de los materiales del test (objetos, material impreso o software):

- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente (Impresión y presentación de gran calidad, software muy atractivo y eficiente, etc.).

2.2. Calidad de la documentación aportada:

- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente (Descripción muy clara y completa de las características técnicas, fundamentada en abundantes datos y referencias).

2.3. Fundamentación teórica:

- No se aporta información en la documentación
- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente (Descripción muy clara y documentada del constructo que se pretende medir y del procedimiento de medición).

2.4. Adaptación del test (si el test ha sido traducido y adaptado para su aplicación en España):

- Característica no aplicable para este instrumento
- No se aporta información en la documentación
- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente (Descripción precisa del procedimiento de traducción, de la adaptación de los ítems a la cultura española, de los estudios de equivalencia con la versión original, utilización de la normativa de la International Test Commission, etc.).

2.5. Calidad de las instrucciones:

- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente (Claras y precisas. Muy adecuadas para las poblaciones a las que va dirigido el test).

2.6. Facilidad para comprender la tarea:

- * () Inadecuada
- ** () Adecuada pero con algunas carencias
- *** () Suficiente
- **** () Buena
- ***** () Excelente (Los sujetos de las poblaciones a las que va dirigido el test pueden comprender fácilmente la tarea a realizar).

2.7. Facilidad para registrar las respuestas:

- * () Inadecuada
- ** () Adecuada pero con algunas carencias
- *** () Adecuada
- **** () Buena
- ***** () Excelente (El procedimiento para emitir o registrar las respuestas es muy simple por lo que se evitan los errores en la anotación).

2.8. Calidad de los ítems (aspectos formales):

- * () Inadecuada
- ** () Adecuada pero con algunas carencias
- *** () Adecuada
- **** () Buena
- ***** () Excelente (La redacción y el diseño son muy apropiados).

2.9. Análisis de los ítems

2.9.1 Datos sobre el análisis de los ítems:

- () Característica no aplicable para este instrumento
- () No se aporta información en la documentación
- * () Inadecuados
- ** () Adecuados pero con algunas carencias
- *** () Adecuados
- **** () Buenos
- ***** () Excelentes (Información detallada sobre diversos estudios acerca de las características psicométricas de los ítems: dificultad o variabilidad, discriminación, validez, distractores, etc.).

2.10. Validez

2.10.1. Validez de contenido:

2.10.1.1. Calidad de la representación del contenido o dominio:

- * () Inadecuada
- ** () Adecuada pero con algunas carencias
- *** () Adecuada
- **** () Buena
- ***** () Excelente (En la documentación se presenta una precisa definición del contenido.
Los ítems muestrean adecuadamente todas las facetas del contenido).

2.10.1.2. Consultas a expertos:

- No se aporta información en la documentación
- * No se ha consultado a expertos sobre la representación del contenido
- ** Se ha consultado de manera informal a un pequeño número de expertos
- *** Se ha consultado a un pequeño número de expertos mediante un procedimiento sistematizado ($N < 10$).
- **** Se ha consultado a un número moderado de expertos mediante un procedimiento sistematizado ($10 \leq N \leq 30$).
- ***** Se ha consultado a un amplio número de expertos mediante un procedimiento sistematizado ($N > 30$).

2.10.2. Validez de constructo:

2.10.2.1. Diseños empleados:

- No se aporta información en la documentación
- Correlaciones con otros tests
- Diferencias entre grupos
- Matriz multirasgo-multimétodo
- Análisis factorial exploratorio
- Análisis factorial confirmatorio
- Diseños experimentales
- Otros (Indique cuales:.....).

2.10.2.2. Tamaño de las muestras en la validación de constructo:

- No se aporta información en la documentación
- * Un estudio con una muestra pequeña ($N < 200$)
- ** Un estudio con una muestra moderada ($200 \leq N < 500$)
- *** Un estudio con una muestra grande ($N \geq 500$)
- **** Varios estudios con muestras de tamaño moderado
- ***** Varios estudios con muestras grandes

2.10.2.3. Procedimiento de selección de las muestras*:

- No se aporta información en la documentación
- Incidental
- Aleatorio
- *Describa brevemente el procedimiento de selección.

2.10.2.4. Mediana de las correlaciones del test con otros tests similares:

- No se aporta información en la documentación
- * Inadecuada ($r < 0.25$)
- ** Adecuada pero con algunas carencias ($0.25 \leq r < 0.40$)
- *** Adecuada ($0.40 \leq r < 0.50$)
- **** Buena ($0.50 \leq r < 0.60$)
- ***** Excelente ($r \geq 0.60$)

2.10.2.5. Calidad de los tests empleados como criterio o marcador:

- No se aporta información en la documentación
- * Inadecuada
- ** Adecuada pero con algunas carencias
- *** Adecuada
- **** Buena
- ***** Excelente

2.10.2.6. Datos sobre el sesgo de los ítems:

- Característica no aplicable para este instrumento
- No se aporta información en la documentación
- * Inadecuados
- ** Adecuados pero con algunas carencias
- *** Adecuados
- **** Buenos
- ***** Excelentes (Información detallada sobre diversos estudios acerca del sesgo de los ítems relacionado con el sexo, la lengua materna, etc. Empleo de la metodología apropiada).

2.10.3. Validez predictiva

2.10.3.1. Describa los criterios empleados y las características de las poblaciones:

2.10.3.1.1. Diseño de selección del criterio:

- Concurrente
- Predictivo
- Retrospectivo

2.10.3.2. Tamaño de las muestras en la validación predictiva:

- No se aporta información en la documentación
- * Un estudio con una muestra pequeña ($N < 100$)
- ** Un estudio con una muestra moderada ($100 \leq N < 200$)
- *** Un estudio con una muestra grande y representativa ($N = 200$)
- **** Varios estudios con muestras representativas de tamaño moderado
- ***** Varios estudios con muestras grandes y representativas

2.10.3.3. Procedimiento de selección de las muestras*:

- No se aporta información en la documentación
- Incidental
- Aleatorio
- *Describa brevemente el procedimiento de selección.

2.10.3.4. Mediana de las correlaciones del test con los criterios:

- No se aporta información en la documentación
- * Inadecuada ($r < 0.20$)
- ** Suficiente ($0.20 \leq r < 0.35$)
- *** Buena ($0.35 \leq r < 0.45$)
- **** Muy buena ($0.45 \leq r < 0.55$)
- ***** Excelente ($r \geq 0.55$)

2.10.4. Comentarios sobre la validez en general:

2.11. Fiabilidad

2.11.1. Datos aportados sobre la fiabilidad:

- Un único coeficiente de fiabilidad
- Un único error típico de medida
- Coeficientes de fiabilidad para diferentes grupos de sujetos
- Error típico de medida para diferentes grupos de sujetos

2.11.2. Equivalencia (Formas paralelas):

2.11.2.1. Tamaño de las muestras en los estudios de equivalencia:

- No se aporta información en la documentación
- * Un estudio con una muestra pequeña ($N < 200$)
- ** Un estudio con una muestra moderada ($200 \leq N < 500$)
- *** Un estudio con una muestra grande ($N \geq 500$)
- **** Varios estudios con muestras de tamaño moderado
- ***** Varios estudios con muestras grandes

2.11.2.2. Mediana de los coeficientes de equivalencia:

- No se aporta información en la documentación
- * Inadecuada ($r < 0.50$)
- ** Adecuada pero con algunas carencias ($0.50 \leq r < 0.60$)
- *** Adecuada ($0.60 \leq r < 0.70$)
- **** Buena ($0.70 \leq r < 0.80$)
- ***** Excelente ($r \geq 0.80$)

2.11.3. Consistencia interna

2.11.3.1. Tamaño de las muestras en los estudios de consistencia:

- No se aporta información en la documentación
- * Un estudio con una muestra pequeña ($N < 200$)
- ** Un estudio con una muestra moderada ($200 \leq N < 500$)
- *** Un estudio con una muestra grande ($N \geq 500$)
- **** Varios estudios con muestras de tamaño moderado
- ***** Varios estudios con muestras grandes

2.11.3.2. Mediana de los coeficientes de consistencia:

- No se aporta información en la documentación
- * Inadecuada ($r < 0.60$)
- ** Adecuada pero con algunas carencias ($0.60 \leq r < 0.70$)
- *** Adecuada ($0.70 \leq r < 0.80$)
- **** Buena ($0.80 \leq r < 0.85$)
- ***** Excelente ($r \geq 0.85$)

2.11.4. Estabilidad (Test-Retest)

2.11.4.1. Tamaño de las muestras en los estudios de estabilidad:

- () No se aporta información en la documentación
- * () Un estudio con una muestra pequeña ($N < 100$)
- ** () Un estudio con una muestra moderada ($100 \leq N < 200$)
- *** () Un estudio con una muestra grande ($N \geq 200$)
- **** () Varios estudios con muestras de tamaño moderado
- ***** () Varios estudios con muestras grandes

2.11.4.2. Mediana de los coeficientes de estabilidad:

- () No se aporta información en la documentación
- * () Inadecuada ($r < 0.55$)
- ** () Adecuada pero con algunas carencias ($0.55 \leq r < 0.65$)
- *** () Adecuada ($0.65 \leq r < 0.75$)
- **** () Buena ($0.75 \leq r < 0.80$)
- ***** () Excelente ($r \geq 0.80$)

2.11. 5 Comentarios sobre la fiabilidad en general:

2.12. Normas

2.12.1. Calidad de las normas:

- () No se aporta información en la documentación
- * () Un baremo que no es aplicable a la población objetivo
- ** () Un baremo aplicable a la población objetivo con cierta precaución
- *** () Un baremo adecuado para la población objetivo
- **** () Varios baremos dirigidos a diversos estratos poblacionales
- ***** () Amplio rango de baremos en función de la edad, el sexo, el nivel cultural y otras características relevantes.

2.12.2. Tamaño de las muestras:

- () No se aporta información en la documentación
- * () Pequeño ($N < 150$)
- ** () Suficiente ($150 \leq N < 300$)
- *** () Moderado ($300 \leq N < 600$)
- **** () Grande ($600 \leq N < 1000$)
- ***** () Muy grande ($N \geq 1000$)

2.12.3. Procedimiento de selección de las muestras*:

- () No se aporta información en la documentación
- () Incidental
- () Aleatorio
- *Describe brevemente el procedimiento de selección.

2.12.4. Comentarios sobre las normas

3. Valoración global del test

3.1. Con una extensión máxima de 1000 palabras, exprese su valoración del test resaltando sus aspectos positivos y negativos, así como recomendaciones acerca de su uso en diversas áreas profesionales. Indique asimismo cuáles son las características de la prueba que podrían ser mejoradas, carencias de información en la documentación, etc.

3.2. A modo de resumen, rellene las Tablas 1 y 2. La Tabla 1 incluye algunos datos descriptivos del test. En la Tabla 2 se resume la valoración de las características generales del test. Tome en consideración el promedio de las calificaciones emitidas en los apartados que figuran en la segunda columna de la Tabla 2.

Característica	Descripción
Nombre del test (apartado 1.1)	
Autor (apartado 1.3)	
Autor de la adaptación española (apartado 1.4)	
Fecha última revisión (apartado 1.9)	
Constructo evaluado (apartado 1.11)	
Áreas de aplicación (apartado 1.12)	
Soporte (apartado 1.15)	

Tabla 1. Descripción del test

Característica	Apartados	Valoración
Materiales y documentación	2.1 y 2.2	
Fundamentación teórica	2.3	
Adaptación	2.4	
Análisis de los ítems	2.9	
Validez de contenido	2.10.1	
Validez de constructo	2.10.2	
Análisis del sesgo	2.10.2.6	
Validez predictiva	2.10.3	
Fiabilidad: equivalencia	2.11.2	
Fiabilidad: consistencia interna	2.11.3	
Fiabilidad: estabilidad	2.11.4	
Normas	2.12	

Tabla 2. Valoración del test

EL PAÍS VASCO, UNA REFLEXIÓN METODOLÓGICA

La adaptación de un instrumento de evaluación de una población original a otra considerada objetivo resulta imprescindible, si se pretende utilizar dicho instrumento en esta última población, únicamente, la constatación de la ausencia de diferencias entre ambas poblaciones, es decir, considerar que todos los sujetos pertenecen a la misma población podría eximirnos del proceso de adaptación. Por lo tanto, la adaptación resulta necesaria cuando la población objetivo difiere de la población original en la que se utiliza habitualmente dicho instrumento en cuanto a características tales como la cultura, el país o la lengua. A fin de facilitar las pautas metodológicas que deben contemplarse para realizar una adecuada adaptación del test, los miembros de la Comisión Internacional de Tests han propuesto, como se comentó anteriormente, 22 directrices para llevar a cabo la adaptación de pruebas psicológicas susceptibles de ser utilizadas en diferentes poblaciones. Entre tales directrices, se consideran de especial relevancia los puntos catalogados como 7 y 8, que hacen referencia a los diseños que se emplean con el objetivo de establecer, tanto mediante criterios racionales de carácter lingüístico y psicológico como mediante criterios estadísticos, la equivalencia entre las diferentes versiones lingüísticas del instrumento.

Se han presentado las directrices que se consideran más importantes en el proceso de adaptación de un test, poniendo de manifiesto las carencias que presentan algunas de estas directrices para llevar a cabo la adaptación en determinadas situaciones de contacto de lenguas tales como, por ejemplo, la que caracteriza al País Vasco, donde coexiste una lengua minorizada, el euskera, y una lengua común, el español. Posteriormente se describe lo que se entiende por «situaciones de contacto de lenguas», término que engloba realidades muy heterogéneas y que, en ocasiones, exigen llevar a cabo correcciones en algunos puntos del proceso de adaptación habitual. Y, por último, se realiza una propuesta de diseño que pretende superar algunas de las carencias que presentan los diseños de uso habitual y que, por tal razón, se considera adecuada para una situación de contacto de lenguas como la que caracteriza al País Vasco.

Las directrices para la adaptación de tests o pruebas psicométricas han sido divididas en cuatro grandes áreas, como ya se ha comentado: 1) contexto, 2) construcción y adaptación del test, 3) aplicación y 4) documentación e interpretación de las puntuaciones.

Para comenzar se presentan algunas de las expuestas directrices, incluidas en el área referida a la construcción y adaptación del test, cuyo objetivo radica en asegurar que las distintas versiones del test miden el mismo constructo de la misma forma, es decir, que son equivalentes y que, a su vez, se adaptan adecuadamente a las características particulares del grupo cultural o lingüístico al que van dirigidas. Son las siguientes:

3. Los constructores/editores de tests deberían asegurar que en el proceso de adaptación se han tenido en cuenta las diferencias lingüísticas y culturales existentes entre las poblaciones a las que se dirigen las versiones adaptadas del test.

4. Los constructores/editores de tests deberían demostrar que el lenguaje utilizado en las instrucciones, en los ítems y en el manual del test resultan adecuados para todas las poblaciones culturales y lingüísticas a las que va dirigido.

5. Los constructores/editores de tests deberían demostrar que las técnicas de evaluación escogidas, los formatos de los ítems y las diversas reglas y procedimientos asociados a la aplicación del test les resultan familiares a todas las poblaciones a las que va dirigido.

6. Los constructores/editores de tests deberían demostrar que el contenido de los ítems y los tipos de estímulos incluidos en el test les resultan familiares a todas las poblaciones a las que va dirigido.

7. Los constructores/editores de tests deberían proporcionar criterios racionales y sistemáticos, tanto de carácter lingüístico como psicológico, para incrementar la

precisión en el proceso de adaptación y demostrar la equivalencia entre las distintas versiones lingüísticas del test.

Los diseños basados en criterios racionales que se emplean con mayor frecuencia con el objetivo de establecer la equivalencia entre las diferentes versiones del test e identificar ítems posiblemente sesgados son los distintos diseños de adaptación directa y el diseño de adaptación inversa.

En los diseños de adaptación directa un equipo de traductores traduce la versión original del test al idioma objetivo. A partir de este momento disponemos de dos opciones fundamentales de diseño:

(a) En las pruebas post-adaptación, se administra la versión traducida del test a una muestra de sujetos y se les pide que interpreten el significado de cada ítem. Se considera que ambas versiones del test, son equivalentes, cuando un alto porcentaje de monolingües competentes en el idioma objetivo, percibe el significado de cada ítem de la misma forma que los sujetos monolingües competentes en el idioma original.

(b) Una segunda posibilidad, dentro de los diseños de adaptación directa, consiste en que jueces bilingües establezcan comparaciones entre las versiones de cada ítem del test, en el idioma original y en el idioma objetivo, y decidan si alguna diferencia de traducción puede dar lugar a una no-equivalencia de significado en las dos poblaciones de interés.

En el diseño de adaptación inversa, un equipo de traductores, traduce la versión original del test al idioma objetivo y otro equipo distinto vuelve a traducir esta última versión al idioma original (versión inversamente adaptada). La equivalencia entre ambas versiones del test, es evaluada por un equipo de jueces, que comparan las semejanzas y diferencias existentes entre los ítems, incluidos en la versión original y en la versión inversamente adaptada del test.

Los diseños basados en criterios racionales, no permiten evaluar adecuadamente la equivalencia entre las distintas versiones del test, ya que adolecen de muchos problemas. Dado que una evaluación adecuada de la equivalencia debe partir de un análisis empírico de las respuestas de los propios sujetos, resulta también necesario utilizar diseños y procedimientos de naturaleza estadística.

8. Los constructores/editores de tests deberían asegurar que el diseño de recogida de datos permite la utilización de técnicas estadísticas adecuadas para establecer la equivalencia entre las diferentes versiones del test.

Los diseños basados en criterios estadísticos, varían en función de dos dimensiones: (1) la naturaleza de la muestra o de los sujetos evaluados (monolingües competentes en el idioma original, monolingües competentes en el idioma objetivo o bilingües) y (2) la versión del test que se les administra a los sujetos (original, directamente adaptada o inversamente adaptada).

Cabe destacar tres tipos de diseños estadísticos de uso frecuente para establecer la equivalencia entre las diferentes versiones del test:

- a) Sujetos bilingües reciben las versiones original y adaptada del test. En este diseño se les administran ambas versiones del test a una muestra de sujetos bilingües y, posteriormente, se comparan las puntuaciones obtenidas por dichos sujetos en las distintas versiones de la prueba. La ventaja principal de este diseño, es que permite controlar el sesgo asociado a las posibles diferencias existentes en las aptitudes de los sujetos, que responden a las diferentes versiones del instrumento. Sin embargo, su principal limitación radica en que su validez, se basa en la asunción de que los sujetos, poseen el mismo nivel de competencia en ambos idiomas, lo cual no es muy probable en situaciones como la del País Vasco, debido a la asimetría lingüística a la que se ven sometidos los bilingües. Además, desde una perspectiva más sociocultural, aunque la adaptación de la prueba al euskera se realice adecuadamente, esto no garantiza que el lenguaje estándar utilizado se adecue al euskera de uso real en el sector de la población a la que va dirigido el test.

- b) Monolingües competentes en el idioma original, reciben las versiones original e inversamente adaptada del test. En este diseño, la equivalencia se evalúa comparando las puntuaciones obtenidas, en las versiones original e inversamente adaptada del test, por una muestra de monolingües competentes en el idioma original. Su principal ventaja radica en que dado que son los mismos sujetos quienes reciben las dos versiones del test en el mismo idioma, el sesgo asociado a las posibles diferencias existentes en las aptitudes de los sujetos queda bajo control. Este diseño, permite obtener información acerca de la idoneidad de la traducción, con respecto a la lengua de uso estándar entre los hablantes competentes en el idioma objetivo. No obstante, presenta limitaciones importantes, ya que no permite recoger datos empíricos acerca del rendimiento de los monolingües competentes en el idioma objetivo o del rendimiento de los bilingües en ese mismo idioma y, por ende, acerca del funcionamiento de la versión adaptada del test. Además, no permite descartar la influencia que puede ejercer el hecho de responder a la versión original del test sobre el rendimiento obtenido por los sujetos en la versión inversamente adaptada de la prueba.
- c) Monolingües competentes en el idioma original, reciben la versión original del test y monolingües competentes en el idioma objetivo, reciben la versión adaptada. En este diseño, la equivalencia se evalúa comparando las puntuaciones obtenidas en las versiones original y adaptada del test por dos muestras de monolingües competentes en los idiomas original y objetivo, respectivamente. Su principal ventaja, radica en que dado que ambas muestras responden al test en su idioma materno, los resultados, son generalizables a las respectivas poblaciones monolingües. Sin embargo, tales resultados son vulnerables al sesgo asociado a las posibles diferencias existentes en las aptitudes de los sujetos, que responden a las distintas versiones del instrumento. Además, este diseño no puede aplicarse en determinadas situaciones de contacto de lenguas tales como, por ejemplo, la correspondiente al País Vasco, en la que no existen monolingües vascoparlantes.

9. Los constructores/editores de tests, deberían aplicar técnicas estadísticas adecuadas para (1) establecer la equivalencia entre las diferentes versiones del test y (2) identificar componentes problemáticos o determinados aspectos del test que puedan resultar inadecuados para una o varias de las poblaciones a las que va dirigido.

La última directriz a destacar, se incluye en la sección referida a la Documentación e interpretación de las puntuaciones y es la siguiente:

22. La persona que elabora el test, debería proporcionar información específica acerca de la forma en la que los contextos sociocultural y ecológico de las poblaciones a las que se dirige la prueba, pueden afectar al rendimiento obtenido por los sujetos en la misma, y debería sugerir procedimientos para contemplar tales efectos al interpretar los resultados.

Tras presentar las directrices más relevantes y las limitaciones que presentan en determinadas situaciones de contacto de lenguas y de culturas, a la hora de abordar la complejidad de estas situaciones, se hace un especial hincapié en la situación referida al País Vasco.

Diversidad y especificidad de las situaciones de contacto de lenguas y de culturas

Ya se ha dicho anteriormente que las situaciones de contacto de lenguas y de contacto de culturas, en el contexto europeo en el que vivimos, más que una excepción, son la norma, lo habitual. Es más, la propia historia de la Europa que conocemos es una historia de contacto, de intercambio y, en definitiva, de interacción continua. Lo extraño, lo poco probable, es la uniformidad. Decenas de millones de personas viven inmersas en situaciones de contacto de lenguas y/o culturas (Siguán, 1994). En España más del cuarenta por ciento de sus habitantes, viven en comunidades donde coexisten dos lenguas. No debe entenderse que esos dieciséis millones son bilingües, ni mucho menos, pero sí, que a todos ellos, de una forma u otra, les afecta la situación de contacto de lenguas.

La denominación genérica de bilingüismo se aplica a ámbitos y situaciones tan diferentes entre sí que, en muchas ocasiones, se convierte en un término carente de significado. En psicología se debe distinguir, - la bilingüidad - cualidad que caracteriza a las personas que poseen dos lenguas - , del bilingüismo - término polisémico que denota indistintamente el conocimiento de dos lenguas, la situación sociolingüística y otras acepciones de carácter sociopolítico.

Los instrumentos de evaluación en el País Vasco

En el caso de la editorial TEA, la más importante en cuanto a publicación y distribución de tests, hay 529 publicaciones en castellano frente a solamente 5 publicaciones en euskera, a saber, IHE (1993); D-48 dominoa (1994); Gizarte-Giro eskala (1995); Piele, hezkuntza eta askapen emotiboa bultzateko heziketa-programa (1997); y ECL, irakurriaren ulermenaren ebaluaketa, 1 eta 2 mailak (1998).

El número de pruebas adaptadas al euskera es realmente ridículo, no llega al 1%, y, además, las que se han adaptado no son ni las más necesarias ni, con mucho, las más actuales. No es exagerado decir que no existen tests adaptados al euskera. De hecho, en los últimos seis años, únicamente se han adaptado cinco pruebas al euskera y, además, éstas son de uso restringido.

Ante tal situación, cabe plantearse una serie de cuestiones, a saber: ¿Son necesarias las pruebas en euskera? ¿Pueden aplicarse las pruebas adaptadas al castellano a la población vasco parlante? Los psicólogos que trabajan con personas vasco parlantes se han habituado a utilizar, junto a las pruebas adaptadas al castellano, pruebas de cuestionable traducción al euskera, llegando incluso a practicar la entrevista ayudados de intérpretes absolutamente ajenos a la psicología y atreviéndose a diagnosticar y a desarrollar terapias en tales circunstancias. Lo más grave, sin embargo, es que el problema va en aumento, dado que la cantidad de vasco parlantes es cada vez mayor, especialmente entre la población infantil.

A pesar de llevarse a cabo investigaciones con el suficiente rigor metodológico, los resultados obtenidos en ellas no son satisfactorios. Aun siguiendo, con rigor, las directrices de la Comisión Internacional de Tests, no se consiguen pruebas adaptadas al euskera equivalentes a las del idioma original.

Una de las directrices para la adaptación de los tests plantea tener en cuenta el contexto en el que se pretende realizar la adaptación de las pruebas. Se mencionan aspectos referidos a las lenguas - semántica, pragmática - , a las peculiaridades culturales, en definitiva, a los posibles rasgos diferenciales entre una y otra situación. Pero, en la mayoría de los casos, se supone que se trata de situaciones estándar que no interaccionan, es decir, que responden a un patrón común e invariante. Se presupone, sin fundamento conocido, que el contexto y la situación de cada lengua son estables y mutuamente independientes. Esto no siempre es así. Es más, las situaciones de contacto de lenguas son la expresión evidente de la interacción, la variación y la interdependencia.

En el contacto de lenguas/culturas, cada situación es específica y requiere un estudio exhaustivo para poder aprenderla. Resaltan tres ámbitos que deben tenerse en cuenta antes de iniciar el proceso de adaptación de un test.

1. Lingüístico. Referido a la(s) lengua(s), la distancia interlingüística, la existencia de dialectos, estándar común, literatura, desarrollo léxico, etc.
2. Demolingüístico. Número y proporción de monolingües de cada lengua, proporción de bilingües. En síntesis, número, proporción y distribución de los tipos de hablantes
3. Sociolingüístico. Características de las distintas comunidades monolingües competentes en la(s) misma(s) lengua(s). Definición del estatus de cada lengua, alfabetización, educación, medios de comunicación, etc.

Estos tres ámbitos, recogen los aspectos fundamentales que deben contemplarse para describir las distintas situaciones de contacto de lenguas. El proceso de adaptación de un test, desde una comunidad monolingüe de lengua A para otra, también monolingüe, pero de lengua B, solamente es parcialmente aplicable, cuando se trata de una comunidad donde todos - o la mayoría - de los hablantes conocen la lengua A, y, de entre ellos, sólo algunos son bilingües, es decir, además de la lengua A, conocen también la lengua B.

En determinadas situaciones de contacto de lenguas, como la que caracteriza al País Vasco, los diseños propuestos de acuerdo a las directrices, no permiten obtener la información que se recoge en la número 22. En tales casos, es necesario realizar una adecuación del diseño estadístico a la situación peculiar de contacto de lenguas.

Hace 30 años habría sido absolutamente imposible adaptar un solo test al euskera, por la sencilla razón de que no existía un idioma estándar - euskera unificado - común para todos los vascoparlantes. Existían, y existen, más de seis dialectos y multitud de variantes pero, ni siquiera existía un alfabeto común a todos ellos. Hoy en día existe un euskera unificado, aceptado y habitualmente utilizado. Sin embargo, el tiempo transcurrido para la implementación del euskera unificado - euskara batua -, apenas 30 años, es insignificante en el devenir de una lengua milenaria. Es una situación por tanto variable. Este es el principal factor que ha de tenerse en cuenta si se pretende adaptar el diseño estadístico a fin de que éste posibilite realizar de forma adecuada la adaptación de las pruebas psicométricas al euskera.

Adecuación del diseño estadístico

En primer lugar, cabe señalar que no existen monolingües vascoparlantes. Es decir, los sujetos que viven en el País Vasco son o monolingües castellanoparlantes o bilingües castellano/euskera. No se pueden aplicar, por tanto, los diseños que contemplan la adaptación de la prueba con monolingües de una y otra lengua. Descartados esos diseños, necesariamente hay que trabajar con sujetos bilingües (castellano/euskera), por un lado, y monolingües (castellano), por otro.

La adecuación consiste, básicamente, en trabajar con dos muestras aleatorias, una de monolingües y otra de bilingües. En este diseño, los monolingües cumplimentan las versiones original e inversamente adaptada del test, mientras que los bilingües se someten a las versiones original y adaptada al euskera.

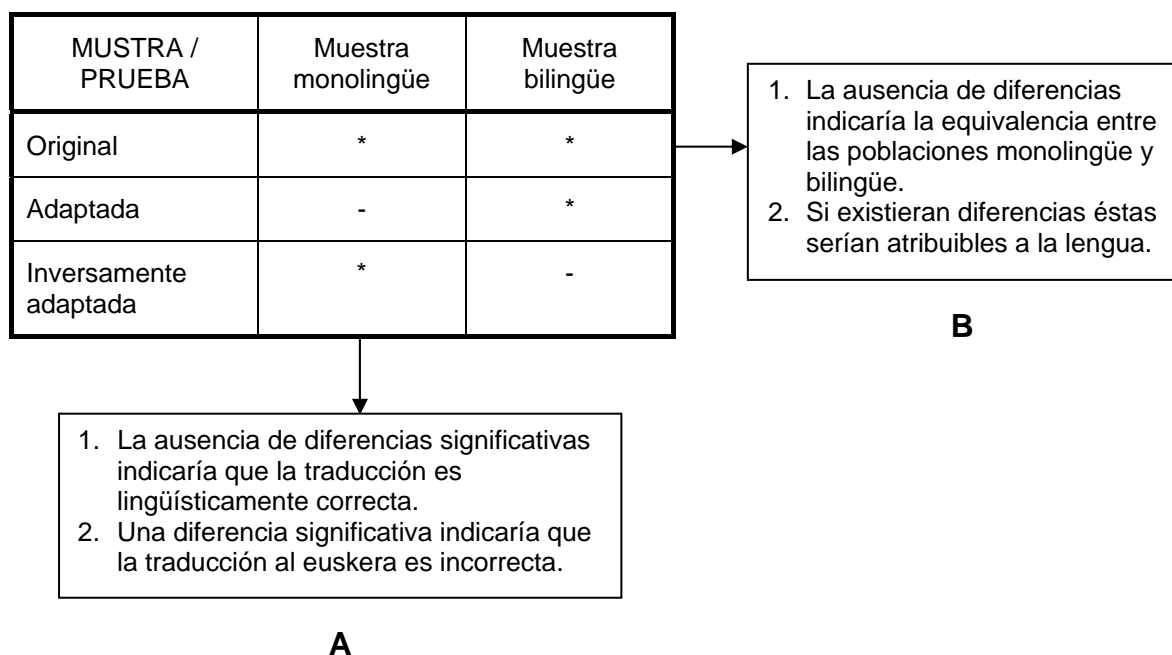
Dicho diseño permite, en primer lugar, contrastar los resultados obtenidos en las pruebas original e inversamente adaptada, es decir, la calidad de la traducción en la muestra monolingüe. En segundo lugar, comparar si existen, diferencias en la prueba original entre ambas muestras -monolingües/bilingües-. Y, por último, examinar si los resultados de los bilingües son diferentes en la versión original del test y en la adaptada al euskera.

En caso de que no existiesen diferencias entre los monolingües y los bilingües en la prueba original, podría alegarse que no tiene sentido adaptar la prueba al euskera cuando puede utilizarse la versión original en la muestra bilingüe. Es, sin duda, un alegato razonable. Sin embargo, siguiendo ese criterio nunca se dispondrían de tests adaptados al euskera. Y, en un proceso de normalización lingüística como el que se vive en el País Vasco, es más que previsible que la estandarización e implementación del uso del euskera unificado en la población vasco parlante, haga necesaria la existencia de pruebas psicométricas en euskera, existan o no diferencias entre las poblaciones monolingüe y bilingüe.

Teniendo en cuenta que todos los vasco parlantes dominan el castellano, la decisión de construir y utilizar pruebas de evaluación psicológica, puede ser una decisión política o puede, por el contrario, responder a criterios deontológicos y metodológicos. ¿Qué psicólogo puede negarle a una persona necesitada de evaluación los medios necesarios para ello, alegando que su lengua o su cultura impiden tal evaluación? Es cierto, que en determinados momentos, razones de índole económica pueden hacer inviable la adaptación de pruebas a multitud de lenguas y contextos. No obstante, ni la política, ni la economía deben constituir el criterio fundamental mediante el que se rigen los psicólogos y los metodólogos. Aquellos que conciben la globalización en términos unificadores y restrictivos, ¿estarían dispuestos a aceptar

que, a partir de ahora, las pruebas de evaluación se construyan, adapten y utilicen únicamente en inglés?

El diseño propuesto, incluiría dos fases previas, que se representan en el siguiente diagrama:



El primer objetivo sería conseguir una traducción lingüísticamente correcta del test. Para ello, además de seguir las directrices para la traducción de tests, hay que asegurarse de la equivalencia entre las versiones original e inversamente adaptada de la prueba. Es decir, debería cumplirse el punto 1 de la fase **A**.

En cuanto a la fase **B** del diagrama, cabe señalar que en una situación de contacto de lenguas/culturas, las diferencias entre los grupos monolingües/bilingües pueden o no existir, no hay ninguna razón a priori que permita afirmar su existencia o su ausencia. La comparación es, por tanto, necesaria. En el caso del País Vasco, a la luz de las investigaciones realizadas, en multitud de pruebas de rendimiento escolar, aptitudes y otro tipo de pruebas, lo observado es que no existen diferencias (Etxeberria, 1999).

Suponiendo que las versiones original e inversamente adaptada sean equivalentes y, en función del resultado obtenido en la comparación entre las muestras de monolingües y bilingües, se procede a la tercera fase, consistente en comparar las versiones original y adaptada al euskera en la muestra bilingüe.

Como es obvio, en el diseño se distinguían cuatro casos:

1. No hay diferencias significativas entre los monolingües y los bilingües en la versión original, y tampoco existen diferencias, en la muestra bilingüe, entre las versiones original y adaptada al euskera. Todas las versiones son equivalentes y la adaptación se ha realizado con éxito.

2. No hay diferencias significativas entre los monolingües y los bilingües en la versión original pero se observan diferencias significativas, en la muestra bilingüe, entre las versiones original y adaptada al euskera. En este caso la versión adaptada al euskera, aún tratándose de una traducción lingüísticamente correcta, al ser significativamente distinta de la versión original en la muestra monolingüe, no sería adecuada respecto a la situación sociolingüística del idioma objetivo y debería llevarse a cabo una readaptación de la prueba atendiendo al uso real del euskera hasta conseguir la equivalencia entre la versión original del test y la readaptada en la muestra bilingüe.

3. Se observan diferencias significativas entre los monolingües y los bilingües en la versión original, pero no entre las versiones original y adaptada en la muestra bilingüe. Si, atendiendo a los resultados obtenidos en la versión original, las diferencias entre las muestras son reales, la versión adaptada al euskera no es adecuada, ya que deberían existir diferencias significativas entre dicha versión y la original en la muestra bilingüe. En este caso, debe revisarse el proceso de adaptación del test.

4. Se observan diferencias significativas entre los monolingües y los bilingües en la versión original y, también, en la muestra bilingüe, entre las versiones original y adaptada. Si los sujetos bilingües puntúan más bajo que los monolingües en la versión

original, se puede pensar que los sujetos bilingües son desequilibrados hacia el euskera, quizá funcionalmente monolingües vascoparlantes, y, si no existiesen diferencias significativas entre el rendimiento de la muestra monolingüe en la versión original y el rendimiento de la muestra bilingüe en la prueba adaptada, se puede concluir que la adaptación del test se ha realizado con éxito. En tal caso, se utilizaría el diseño (c). Si, por el contrario, el rendimiento de los monolingües en la versión original difiriera significativamente del rendimiento de los bilingües en la adaptada, se debería revisar nuevamente todo el proceso de adaptación de la prueba.

En definitiva, el diseño que se ha propuesto, es una combinación entre los diseños (a) y (b) expuestos anteriormente. Este diseño permite detectar los desequilibrios producidos por la situación de contacto de lenguas y adaptar el test, tanto lingüísticamente respecto al idioma estándar de la lengua minorizada, como respecto a la situación sociolingüística referida al uso de ese estándar en el momento de la adaptación de la prueba.

UN EJEMPLO CONCRETO DE ADAPTACIÓN

ADAPTACIÓN AL ESPAÑOL DEL CUESTIONARIO VECÚ ET SANTE PERÇUE DE L'ADOLESCENT (VSP-A): UNA MEDIDA GENÉRICA DE CALIDAD DE VIDA PARA ADOLESCENTES

Resumen

Fundamentos: En la última década se han desarrollado diversas medidas de calidad de vida relacionada con la salud para uso exclusivo en niños/as y adolescentes. No obstante, existen pocos instrumentos de estas características adaptados en España. El Vecú et Santé Perçue de l'adolescent (VSP-A) es un instrumento genérico de calidad de vida relacionada con la salud para adolescentes de 11 a 17 años desarrollado en Francia. El objetivo de este estudio fue adaptar al español el VSP-A, como primera fase para la obtención del cuestionario.

Métodos: Se adaptó la versión del VSP-A de 39 preguntas siguiendo la metodología de traducción directa e inversa incluyendo: 2 Traducciones al español, puntuación del grado de dificultad (0 min-10 máx) y clasificación de equivalencia semántica y cultural, 2 reuniones y discusión en paneles de adolescentes, así como reuniones de consenso del equipo de investigación. Finalmente se realizó una traducción inversa (retro-traducción) al francés y se administró la versión final pre-test en la prueba piloto.

Resultados: La mayoría de las preguntas se clasificaron como equivalentes (24 sobre 39). Tras las reuniones con adolescentes se modificaron algunas preguntas. Tras la retro-traducción. 3 preguntas necesitaron cambios menores.

Conclusiones: La versión española del VSP-A parece semántica y culturalmente equivalente a la versión original en francés y adecuada para adolescentes en España. La sencillez de las preguntas, los comentarios de los adolescentes y la participación de los autores originales en el proceso de adaptación ha permitido obtener una versión pre-test adecuada. La siguiente fase del estudio es la comprobación de la fiabilidad y validez. Se espera que el VSP-A sea de utilidad para medir la calidad de vida relacionada con la salud en encuestas de salud o como instrumento de cribado en colegios o centros de atención primaria en nuestro medio.

Palabras clave: Adolescencia. Calidad de vida. Cuestionario. España. Comparación transcultural.

Introducción

Existe un interés creciente en medir la calidad de vida relacionada con la salud (CVRS), tanto en la investigación y en la práctica clínica, como en la evaluación de servicios sanitarios y en salud pública. Los instrumentos de medida de CVRS recogen la valoración subjetiva del individuo respecto de su estado de salud y su bienestar general y la influencia de éstos sobre la capacidad para llevar a cabo las actividades consideradas importantes por el propio individuo. Las dimensiones que habitualmente

recogen los instrumentos de CVRS son: el funcionamiento físico, psicológico y social; la movilidad y el cuidado personal; y el bienestar emocional. La mayoría de instrumentos de CVRS han sido diseñados para los adultos, si bien en la última década se han desarrollado diversas medidas para uso exclusivo en niños/as y adolescentes, aunque en España existen pocos.

El *Vecú et Santé Perçue de l'Adolescent (VSP-A)* es un instrumento genérico de CVRS, autoadministrado, para adolescentes de 11 a 17 años, que ha sido desarrollado en la Universidad de Marsella (Francia). Este cuestionario fue creado a partir de la opinión de los adolescentes sobre lo que constituía lo más importante sobre su salud percibida. Inicialmente se generaron 118 preguntas a partir de 73 entrevistas en profundidad con adolescentes que recogían sus sensaciones y opiniones sobre la salud, así como el impacto de la salud en su CVRS. Posteriormente se realizaron 2 grupos. Tras analizar las propiedades psicométricas (fiabilidad y validez) se redujo el número de preguntas y se obtuvo una versión con coeficientes de fiabilidad y validez aceptables. La versión original fue diseñada para uso en encuestas de salud de población general, y ha sido administrado a más de 15.000 adolescentes; en Francia el Ministerio de Educación lo ha recomendado como instrumento de cribado en chequeos médicos en los colegios.

La adaptación de un cuestionario a otra cultura, requiere llevar a cabo un proceso sistemático, dirigido a obtener una equivalencia semántica y conceptual respecto del original. Existe un cierto consenso respecto a la metodología más adecuada para la adaptación de medidas de salud, denominada traducción directa e inversa o traducción-retrotraducción. Una parte fundamental del proceso de adaptación incluye la práctica de técnicas cualitativas (reuniones y discusión en paneles o entrevistas cognitivas), para garantizar que el contenido del instrumento sea comprensible y que se incorporen aspectos relevantes para la población a la que va dirigido.

Las características propias del VSP-A como medida genérica de CVRS, así como la falta de este tipo de medidas para adolescentes en España, justifican la

adaptación del cuestionario. El objetivo del presente estudio fue adaptar el VSP-A para su uso en España, como primera fase para la obtención del cuestionario.

Sujetos y método

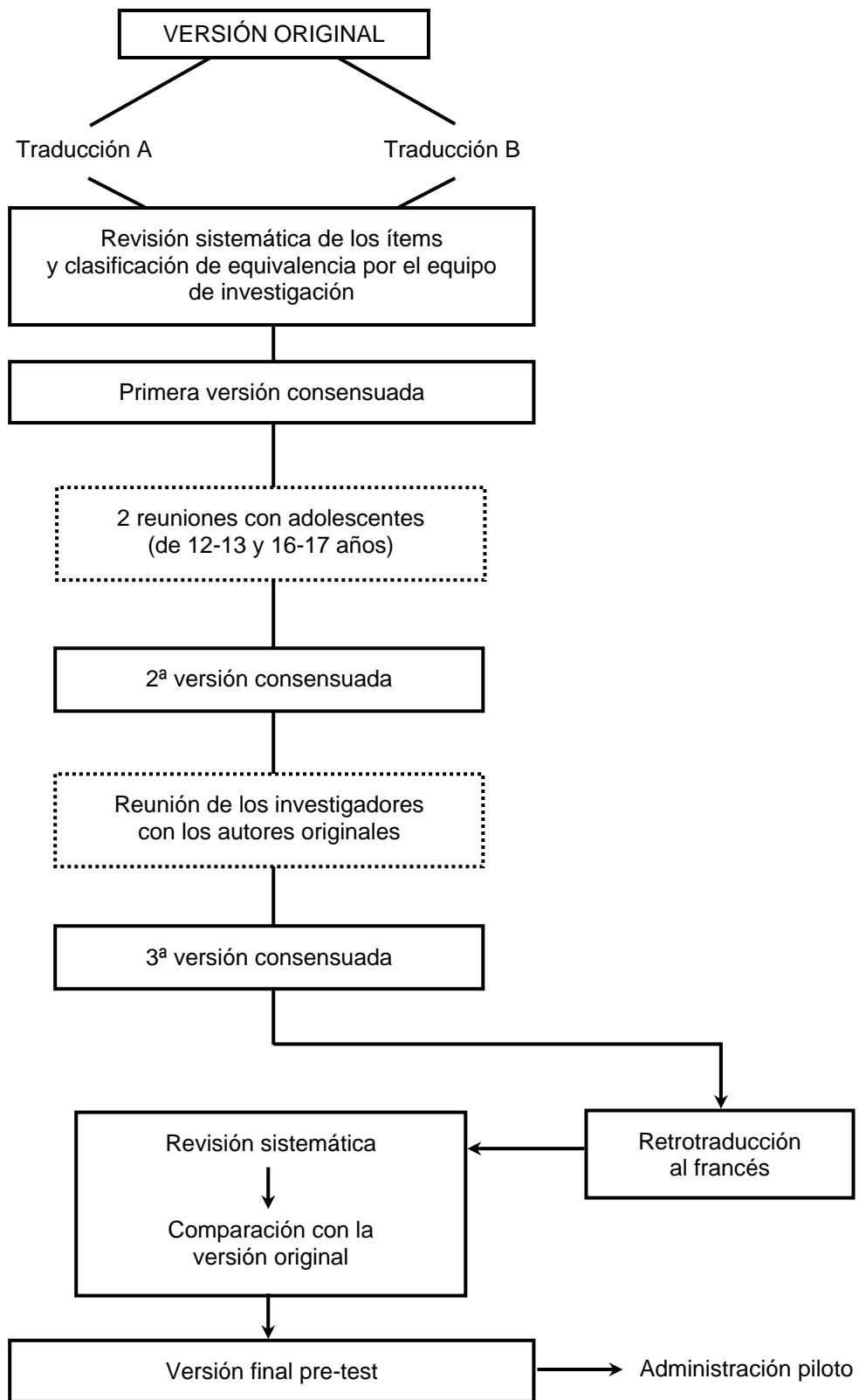
El cuestionario VSP-A

La versión más reciente, que es la adaptada en España, contiene 39 preguntas distribuidas en 11 dimensiones que son: energía; bienestar psicológico; relaciones con amigos/as; actividades de ocio; relaciones con padres; bienestar físico; relaciones con profesores/as; actividades escolares; autoestima; relaciones con el personal sanitario y satisfacción con la vida sentimental y sexual. Las preguntas se puntúan en una escala Likert de 1 a 5. A partir de la puntuación media de las preguntas se genera la puntuación de cada dimensión. Por otro lado, el VSP-A permite obtener una puntuación global única de 0 (peor CVRS) a 100 (mejor CVRS) mediante transformaciones lineales de las 11 dimensiones.

Adaptación de la versión española del VSP-A

Para la adaptación del cuestionario VSP-A se siguió la metodología de traducción directa e inversa (traducción-retrotraducción) recomendada en la literatura. Esta metodología incluyó la traducción al español, la revisión sistemática del contenido del instrumento y discusión por el equipo de investigación, reuniones con grupos de adolescentes, una traducción inversa a la lengua original francesa y administración de la versión pre-test española (figura 1). Los autores originales del cuestionario participaron activamente durante todo el proceso de la adaptación a la versión española, mediante reuniones de consenso.

Figura 1: Esquema del proceso de adaptación transcultural del VSP-A.



Primera fase: traducción al español

Dos traductores profesionales, de lengua materna española y bilingües en francés, tradujeron de forma independiente la versión original francesa del VSP-A al español. El objetivo de esta clase de estudio, fue desarrollar una versión semánticamente equivalente a la original, es decir, que la versión española transmitiera el significado del original, utilizando formulaciones y expresiones de las preguntas que fueran aceptables y culturalmente relevantes para los adolescentes españoles.

El equipo de investigación, junto con los traductores, revisó de forma sistemática las preguntas del cuestionario, las instrucciones y las opciones de respuesta propuestos por ambos traductores, para decidir una primera versión consensuada. Para cuantificar la «adaptabilidad» de su contenido e identificar los problemas iniciales en la adaptación, cada miembro del equipo de investigación, puntuó el grado de dificultad en la adaptación inicial de las preguntas, en una escala de 0 (dificultad mínima) a 10 (máxima dificultad). A partir de dicha puntuación se clasificaron las preguntas en 3 grupos según los criterios siguientes:

(A) Preguntas equivalentes: cuando la dificultad en la traducción fue mínima, ambas versiones traducidas fueron similares, y no se plantearon dudas con respecto al vocabulario y el contenido.

(B) Preguntas con problemas en algunas palabras: cuando ambas versiones traducidas presentaban diferencias entre sí, o el equipo de investigación identificó problemas en alguna/s palabra/s, que incluyeron la necesidad de recoger vocablos o expresiones más adecuadas y naturales en español (B1), o bien incorporar aspectos culturales para conseguir una equivalencia semántica y cultural (por ejemplo, se incluyeron algunas actividades de ocio consideradas más adecuadas para los adolescentes en nuestro medio) (B2).

(C) Preguntas sin equivalencia: se definieron como tal cuando el equipo de investigación clasificó la pregunta como inadecuada en nuestro medio.

A partir de las 2 traducciones en español y de los comentarios del equipo español de investigación, se obtuvo la primera versión consensuada del VSP-A.

Segunda fase: reunión y discusión en paneles de adolescentes y consenso del equipo de investigación.

Se realizaron 2 reuniones con adolescentes procedentes de un colegio de titularidad concertada situado en un barrio de nivel socioeconómico bajo. Cada grupo incluyó 8 adolescentes (4 chicos y 4 chicas) de edades diferentes: 12-13 años y 16-17 años respectivamente. Los objetivos de esta fase fueron valorar la comprensión de la primera versión consensuada, determinar si el vocabulario era adecuado, y evaluar si las preguntas eran comprensibles y relevantes en nuestra cultura. Las sesiones fueron grabadas y transcritas para llevar a cabo un análisis del contenido de las mismas. Cada sesión tuvo una duración aproximada de una hora. Una persona del equipo de investigación moderaba la sesión y otra recogía notas. Se obtuvo el consentimiento previo de los padres para la participación de sus hijos/as. En cada sesión, los adolescentes cumplimentaron el VSP-A y a continuación se realizaron comentarios en grupo al cuestionario, siguiendo un guión semi-estructurado para comprobar la comprensión de las instrucciones, así como de las preguntas y de las opciones de respuesta, del grado de naturalidad de las expresiones y la aplicabilidad del contenido del cuestionario en nuestro medio. Se trataba de recoger las expresiones de uso común, los comentarios y los valores relevantes para los adolescentes. Además se recogieron la edad, el género, la ciudad de nacimiento y la ocupación del padre y de la madre, así como la hora de inicio y de finalización del cuestionario.

Las alternativas propuestas y los comentarios de los adolescentes, fueron valoradas posteriormente por el grupo de investigación, para determinar dónde se deberían introducir cambios en la primera versión consensuada. La introducción de estos cambios, llevó a la segunda versión consensuada del VSP-A en español. Tras una reunión del equipo de investigación, con los autores originales del instrumento para discutir algunas dudas, respecto al contenido de las preguntas y la equivalencia de la

última versión consensuada en español, se obtuvo la tercera versión consensuada del cuestionario.

Tercera fase: retro-traducción al idioma original.

De la tercera versión consensuada del español al francés, se realizó una retro-traducción o traducción inversa, por un traductor de lengua materna francesa y bilingüe en español. Los autores originales, compararon esta versión retrotraducida con la versión original, con la finalidad de detectar diferencias o discrepancias entre ambas y proponer nuevas formulaciones de las preguntas en español, en una reunión de consenso con el equipo de investigación. Estas revisiones de la versión española dieron como resultado la versión pre-test.

Cuarta fase: factibilidad de la administración y comprensión de la versión pre-test.

La versión pre-test, se administró en dos grupos de adolescentes de un colegio público de nivel socioeconómico medio de la ciudad de Barcelona: uno de 12-13 años (4 chicos y 4 chicas) y otro de 16-17 años (4 chicos y 4 chicas). En esta fase del estudio, se valoró la comprensión del vocabulario y la factibilidad de la administración del cuestionario.

Resultados

El rango de puntuaciones de la dificultad en la adaptación osciló entre 0 y 5,5. Se clasificaron 24 de las 39 preguntas del VSP-A como preguntas (A), 14 como preguntas (B) y 1 pregunta como (C). En la tabla 1 se describen las preguntas de tipo B y C. Las preguntas del cuestionario con mayor dificultad de adaptación en la primera fase fueron: «¿te has sentido rodeado/a, ayudado/a por tus amigos/as? (avez-vous ¿té: entouré(e), aidé(e) par vos copains, vos copines?)»; «¿te has sentido comprendido/a, tranquilizado/a por tu amigos/as? (avez-vous été: compris(e), rassuré(e) par vos

copains, vos copines?); «¿has tenido dolor en alguna parte de tu cuerpo? (avez-vous pu: des douleurs, mal quelque part?)».

En las tablas 2 y 3 se muestran ejemplos de los comentarios de los adolescentes, incluyendo los de 12-13 y de 16-17 años en relación a las preguntas formuladas en la primera versión consensuada. Por ejemplo, la expresión «agobiado/a» reflejaba de manera más adecuada cómo se sentían respecto al estrés que «estresado/a», que fue considerada una expresión de personas mayores por el grupo de 16-17 años. La palabra «angustiado/a», no era la más adecuada para expresar cómo se sentían con respecto a su futuro y preferían la palabra «preocupado/a». Se sentían «apoyados por los amigos» en lugar de «rodeados».

En relación a la pregunta: «¿te has sentido satisfecho con tu vida sexual?», los participantes de 16-17 años expresaron que, aunque hablar de sexo... «es algo privado.... refleja sí te sientes más contento o más triste..» «y es necesario incluirlo en este tipo encuestas». Los más jóvenes en cambio (sobre todo las chicas), consideraron que «tenían vida sentimental, no sexual, y que es «una pregunta para mayores».

Tabla 1: Clasificación de equivalencia semántica y cultural de preguntas de la 1ª versión consensuada en español

Pregunta versión francesa	Pregunta de la 1ª versión consensuada	Puntuación mediana de la dificultad en la traducción	Equivalencia
Avez-vous été angoisé(e) ou eu peur en pensant à l'avenir?	¿te has sentido angustiado/a o asustado ante el futuro?	4	(B1)
Avez-vous été entouré(e), aidé(e) par vos copains, vos copines?	¿te has sentido rodeado/a, ayudado/a por tus amigos/as?	4	(B1)
Avez-vous compris(e), rassuré(e) par vos copains, vos copines?	¿te has sentido comprendido/a, tranquilizado/a por tus amigos/as?	5,5	(B1)
Avez-vous eu: des douleurs, mal quelque part?	¿has tenido dolor en alguna parte de tu cuerpo?	3,5	(B1)
Êtes vous allé(e) chez vos copains, vos copines?	¿has ido a casa de tus amigos o amigas?	4	(B1)
Avez-vous eu le moral?	¿has tenido la moral alta?	5	(B1)

Êtes-vous allé(e) jouer dehors avec vos copains, vos copines (vélo, foot, roller...)?	¿has hecho actividades al aire libre con tus amigos o amigas (ir en bici, skating, juegos de pelota...)?	5	(B2)
Avez-vous pu: sortir (aller en ville, faire des achats, aller à la piscine, au cinéma...)?	¿has podido salir (ir al centro, de compras, a la piscina, al cine...)?	3,5	(B2)
Avez-vous été entouré(e), aidé(e) par le personnel soignant (médecin, infirmière, psychologue)?	¿te has sentido rodeado/a, ayudado/a por el personal sanitario (médico/a, enfermero/a, psicólogo/a)?	4,5	(B2)
Avez-vous été satisfait(e) de votre vie sexuelle?	¿te has sentido satisfecho de tu vida sexual?	4	(C)

- (A) Preguntas equivalentes
- (B) Preguntas con problemas en algunas palabras
B1: Necesidad de incorporar expresiones coloquiales o palabras comprensibles
B2: Necesidad de incorporar aspectos culturales
- (C) Preguntas sin equivalencia

En la fase de traducción inversa, se encontraron 2 discrepancias entre la versión original y la retro-traducción al francés: avez-vous été: stressé(e)?; entouré(e), aidé(e) par vos copains, vos copines? Por este motivo, se decidió añadir respectivamente, «estresado/a» (de la primera versión consensuada) al término «agobiado/a»; y «ayudado/a» al término «apoyado/a» para mejorar la equivalencia con la versión original.

Finalmente, después de la administración de la versión pre-test a 2 grupos de adolescentes se modificaron los enunciados de 2 preguntas: se añadió «... o han ido ellos a la tuya?» a la pregunta «¿has ido a casa de tus amigos o amigas?»; y se modificó «¿has tenido dolores?» por la opción propuesta en la 1ª versión consensuada «¿has tenido dolor en alguna parte de tu cuerpo?», ya que ésta última formulación presentó menos problemas de comprensión y recogía mejor el dolor físico.

Discusión

El presente estudio ha permitido obtener la versión española del cuestionario genérico de CVRS para adolescentes VSP-A y constituye la primera fase para la

obtención del instrumento de medida. Esta metodología, con algunas variantes, se ha utilizado en España para la adaptación de diversos instrumentos de CVRS para adultos y adolescentes. La adaptación transcultural de cuestionarios, permite obtener instrumentos de medida de CVRS, comparables a nivel internacional, realizar estudios multicéntricos que incluyan diferentes países, así como aprovechar la experiencia previa obtenida con el cuestionario. Así mismo, es importante seguir una metodología rigurosa y sistemática para garantizar que las cualidades del instrumento de medida que ha sido adaptado, sean similares a las del instrumento original (habilidad, validez y sensibilidad al cambio), así como para facilitar la aplicabilidad del instrumento en la nueva cultura.

Tabla 2: Resultados de los comentarios de los adolescentes sobre la claridad del lenguaje y uso de expresiones cotidianas

Ejemplos de preguntas de la 1ª versión consensuada	Grupo de adolescentes de 12-13 años	Grupo de adolescentes de 16-17 años
¿Te has sentido estresado/a?	«Estresado es lo mismo que desanimado, cansado». «No tienes tiempo para hacerlo todo...».	«Estresado es una palabra para más mayores». Prefieren «agobiado» porque es una palabra más cotidiana.
¿Te has sentido rodeado/a ayudado, por tus amigos/as?	... «los amigos/as te apoyan, te ayudan». «Ayudar es que te resuelvan el problema...». Hay dos ideas en una sola frase.	«Son los padres los que ayudan». «Los amigos/as te apoyan...». Prefieren apoyados, respaldados que rodeados. «Apoyados por los amigos y ayudados por los padres...».
¿Has tenido dolor en alguna parte de tu cuerpo?	«dolor en la barriga...». Los chicos dicen que han tenido dolor «en el tobillo por el fútbol, por hacer deporte...».	La alternativa que proponen es: «has tenido algún dolor en cualquier parte de tu cuerpo...».
¿Te has sentido rodeado/a, ayudado/a por el personal sanitario (médico/a, enfermero/a, psicólogo/a)?	A diferencia de los amigos/as se sienten «ayudados/as por el personal sanitario...», no «rodeados/as, ni apoyados...».	... «más que ayudados se sienten comprendidos/as, aunque depende del caso...».

Tabla 3: Resultados de los comentarios de los adolescentes sobre la adecuación cultural del contenido de las preguntas

Ejemplos de preguntas de la 1ª versión consensuada	Grupo de adolescentes de 12-13 años	Grupo de adolescentes de 16-17 años
¿Te has sentido angustiado/a o asustado/a ante el futuro?	... «angustiado suena muy fuerte». «La palabra asustado/a equivale a estar preocupado. Mejor poner que te preocupa». «Le das vueltas a las cosas; piensas en que las cosas pueden salir mal...». «No es igual que angustiado».	Más que angustiado que suena muy fuerte, se sienten «preocupados... por las notas, la escuela». «Asustados no estamos, más bien preocupados...». Hacen referencia a preocupaciones escolares, y al futuro inmediato.
¿Has hecho actividades al aire libre con tus amigos o amigas (ir en bici, skating, practicar algún deporte...)?	Proponer actividades como «pasear, dar una vuelta, ir en bici...». Una chica dice «ir al cine», pero el grupo comenta que se trata de actividades al aire libre.	Una chica comenta que estas actividades le recuerdan más al verano. «Ir en bici se puede incluir, andar, pasear, sacar al perro, deportes (monopatín, fútbol)...».
¿Te has sentido contento/a con tu vida sexual?	Se ríen. Les hace gracia la pregunta. Tienen vida sentimental, no sexual y creen que es una pregunta para mayores.	Se ríen de la expresión. «El sexo es tabí... son cosas privadas...». «Aunque es importante hacer preguntas así, pondría satisfecho de tu vida sexual porque el sexo te hace gozar...». «Es importante hacer este tipo de pregunta porque te pueden hacer sentir más contento o más triste...».

Durante la adaptación del VSP-A al español, no se encontraron problemas importantes de traducción, ni de adaptación del contenido de las preguntas y la puntuación, de la equivalencia en la primera fase del proceso, indicó que las traducciones eran similares al contenido de las preguntas originales. La formulación sencilla de la mayoría de las preguntas, así como el hecho de que el cuestionario utiliza pocas expresiones coloquiales en la versión original francesa, facilitó su traducción.

Por otra parte, a pesar de las diferencias en el idioma, no se esperaban encontrar diferencias culturales importantes, dada la proximidad geográfica y la pertenencia de ambas poblaciones de adolescentes, tanto la de la versión original como la de la población diana, a culturas mediterráneas.

En el proceso de adaptación de un cuestionario a una nueva cultura, es importante realizar reuniones con la población a la cual va dirigido, para conocer los significados que dan, en este caso los adolescentes, a los contenidos del cuestionario. Las reuniones con adolescentes, permitió redefinir las preguntas con más problemas de adaptación y comprobar la relevancia y la comprensión de los contenidos del cuestionario, como por ejemplo, las preguntas sobre la imagen corporal o sobre la satisfacción con la vida sexual.

En general, la comprensión del cuestionario fue buena, aunque se propusieron algunos cambios para mejorar la claridad del lenguaje e incorporar expresiones más relevantes para adolescentes. Las preguntas referidas a la apariencia física, generaron una reacción de rechazo inicial en algunas de las chicas, aunque reconocieron la importancia del tema para ellas y la necesidad de recoger la autopercepción respecto de la misma. Sólo una pregunta se consideró problemática inicialmente para los adolescentes más jóvenes. Ésta recogía información sobre la satisfacción del adolescente con su vida sexual. Tras los comentarios de los adolescentes, se decidió dejar esta pregunta ya que la mayoría de adolescentes confirmó que podía tener vida sentimental y sexual, aunque algunos de los más jóvenes y algunas chicas comentaron que no tenían «novio/a» ni vida sexual. Por este motivo, se decidió incluir en la segunda versión consensuada, la opción de respuesta «no tengo» en las preguntas sobre satisfacción del adolescente con su vida sentimental y sexual, así como en las preguntas referidas al contacto con el personal sanitario «no me han visitado», para aquellos adolescentes que no tenían vida sentimental/sexual o que no habían tenido contacto con el personal sanitario durante las 4 últimas semanas.

La introducción de esta opción de respuesta para las preguntas relacionadas con la vida sexual y en las preguntas sobre uso de servicios sanitarios, es la principal diferencia de la versión adaptada respecto del original y merece ser comentada. De hecho, en el cuestionario original existe una opción de respuesta parecida [«non concerné(e)», «no me afecta»] en todas las preguntas. Sin embargo, se decidió eliminarla, después de extensas discusiones del equipo investigador, debido a los importantes problemas de comprensión detectados en las reuniones con adolescentes.

En muchas ocasiones, los adolescentes inicialmente seleccionaron la opción «no me afecta» para indicar que no les pasaba nada (por ejemplo, no se habían sentido tristes en las últimas 4 semanas). Por este motivo, se decidió incluir sólo las alternativas de respuesta para especificar que no tenían vida sexual o sentimental o que no habían utilizado los servicios sanitarios en las últimas 4 semanas.

Finalmente, la participación de los autores originales en todo el proceso, ha sido importante para obtener una versión adecuada y equivalente. Tras la retro-traducción de la tercera versión consensuada al francés, en la que los autores originales confirmaron que el proceso de adaptación de la versión española había mantenido el contenido y significado de cada una de las preguntas de la versión original, se produjeron muy pocos cambios, probablemente porque el proceso anterior había sido riguroso y sistemático.

Una limitación del estudio, fue que en ningún momento se intentó averiguar con preguntas abiertas, el concepto que los adolescentes tienen de la calidad de vida, porque se trataba de la adaptación de un cuestionario ya elaborado. Por tanto, es posible que algunos temas relevantes para adolescentes en España, no se hayan incorporado; sin embargo, esto es una limitación de este tipo de proceso de adaptación y sería interesante llevar a cabo estudios futuros, que intentasen investigar en más profundidad la opinión de los adolescentes en España, respecto a su concepto de la calidad de vida.

Por otra parte, el hecho de obtener una versión del VSP-A adaptada al español, no significa que el instrumento tenga las mismas propiedades psicométricas que la versión original. En futuros estudios, será necesario evaluar la fiabilidad (consistencia interna y estabilidad test-retest) y la validez del instrumento (constructo, convergente/discriminante, criterio y estructural), así como su sensibilidad al cambio. También sería conveniente probar la factibilidad de la administración del cuestionario en varias regiones, así como adaptar versiones que incluyan otras lenguas oficiales para uso en comunidades como Cataluña o País Vasco.

Una vez comprobadas sus propiedades psicométricas, el VSP-A será un cuestionario de CVRS útil, sencillo y corto, con un tiempo de administración de 10-15 minutos, para uso en la práctica diaria y en diversos ámbitos de la salud pública. El VSP-A podrá ser utilizado en encuestas de salud o como instrumento de cribado para detectar diferencias sistemáticas entre grupos de adolescentes con diversas características (sociodemográficas o clínicas) y relacionarlas con potenciales diferencias en el estado de salud y en el acceso a los servicios sanitarios. El instrumento también podrá ser útil para establecer necesidades en salud y plantear estrategias para mejorar la CVRS de los adolescentes.

Versión española del cuestionario VSP-A
Tu salud y calidad de vida

Durante las cuatro últimas semanas,

	<i>Nunca Nada</i>	<i>Pocas veces Poco</i>	<i>Algunas veces Moderada- mente</i>	<i>Muchas veces Mucho</i>	<i>Siempre Muchísimo</i>
1. ¿Te has sentido inquieto/a, preocupado/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ¿Te has sentido triste?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ¿Te has sentido estresado/a, agobiado/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ¿Te has desanimado con facilidad?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. ¿Te has sentido preocupado/a por el futuro?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. ¿Te has sentido satisfecho/a con tu vida?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. ¿Te has sentido ayudado/a, apoyado/a por tus amigos/as?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. ¿Te has sentido comprendido/a por tus amigos/as?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Durante las cuatro últimas semanas,

	<i>No tengo</i>	<i>Nunca Nada</i>	<i>Pocas ve- ces Poco</i>	<i>Algunas veces Modera- damente</i>	<i>Muchas veces Mucho</i>	<i>Siempre Muchísimo</i>
9. ¿Te has sentido satisfecho/a con la relación sentimental con tu chico/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. ¿Te has sentido satisfecho/a con tu vida sexual?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Durante las cuatro últimas semanas,

	<i>Nunca Nada</i>	<i>Pocas veces Poco</i>	<i>Algunas veces Moderada- mente</i>	<i>Muchas veces Mucho</i>	<i>Siempre Muchísimo</i>
11. ¿Te has sentido comprendido/a, tranquilizado/a por tus padres?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. ¿Te has sentido satisfecho/a de tus resultados escolares?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. ¿Te has sentido ayudado/a por tus profesores?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. ¿Te has sentido comprendido/a por tus profesores?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. ¿Te has sentido aceptado/a, respetado/a por tus profesores?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Durante las cuatro últimas semanas,

	<i>Nunca Nada</i>	<i>Pocas veces Poco</i>	<i>Algunas veces Modera- damente</i>	<i>Muchas veces Mucho</i>	<i>Siempre Muchísimo</i>
16. ¿Te has sentido acomplejado/a por tu físico, por tu aspecto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. ¿Te has sentido demasiado gordo/a, demasiado delgado/a, demasiado alto/a o demasiado bajo/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. ¿Te ha faltado energía?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. ¿Has estado en buena forma física?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. ¿Te has sentido débil, cansado/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. ¿Te han dado tus padres algún consejo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22. ¿Has ido a casa de tus amigos o amigas o han ido ellos/as a la tuya?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23. ¿Has hecho algún tipo de actividad al aire libre con tus amigos o amigas: ir en bici, practicar algún deporte...?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24. ¿Te has sentido con la moral alta?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25. ¿Te has tomado las cosas por su lado positivo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
26. ¿Has tenido la sensación de que todo iba bien a tu alrededor?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
27. ¿Has tenido confianza en ti mismo/a, te has sentido seguro/a de ti mismo/a?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28. ¿Has tenido buenos resultados en el colegio, en el instituto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
29. ¿Has tenido dolor en alguna parte de tu cuerpo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30. ¿Has podido hablar libremente con tus amigos/as, dar tu opinión?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
31. ¿Has podido reunirse con tu grupo de amigos/as?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
32. ¿Has podido salir: ir al centro, de compras, al cine, dar una vuelta...?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
33. ¿Has podido charlar con tus amigos o amigas?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34. ¿Has podido sincerarte, hablar de tus problemas con tus amigos/as?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
35. ¿Has podido sincerarte, hablar de tus problemas con tus padres?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
36. ¿Has podido hablar libremente con tus padres, dar tu opinión?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Durante las cuatro últimas semanas,

	<i>No me han visitado</i>	<i>Nunca Nada</i>	<i>Pocas veces Poco</i>	<i>Algunas veces Modera- damente</i>	<i>Muchas veces Mucho</i>	<i>Siempre Muchísimo</i>
37. ¿Te has sentido ayudado/a por el personal sanitario (médico/a, enfermero/a, psicólogo/a...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
38. ¿Te has sentido comprendido/a, tranquilizado/a por el personal sanitario (médico/a, enfermero/a, psicólogo/a...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
39. ¿Te has sentido respetado/a por el personal sanitario (médico/a, enfermero/a, psicólogo/a...)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

DIRECTRICES INTERNACIONALES PARA EL USO DE LOS TEST

Objetivo

El objetivo fundamental de las directrices, es mejorar el uso de los tests, describiendo la forma adecuada de utilizarlos.

Un usuario competente, utilizará los tests de forma adecuada, profesional y ética, prestando la debida atención a las necesidades y derechos de las personas implicadas en el proceso de evaluación, y teniendo muy en cuenta las razones para utilizar los tests, así como el contexto en el cual se lleva a cabo su aplicación. Este objetivo se alcanzará, asegurándose de que el usuario de los tests, tiene las competencias y conocimientos necesarios para llevar a cabo el proceso evaluativo.

Ámbito de aplicación

Resulta difícil dar una definición muy precisa de test o de su práctica, pues al hacerlo pueden dejarse fuera algunos procedimientos que debieran incluirse, o, por el contrario, incluir otros que deberían quedar fuera. En estas directrices, los términos test y práctica de los tests se utilizan en sentido amplio. Las directrices son aplicables a muchos procedimientos de evaluación que no suelen denominarse tests, o que incluso tratan de evitar su designación como tests. Más que proponer una definición cerrada, se describe a continuación el ámbito que tratan de cubrir las directrices.

- Los tests incluyen un amplio abanico de procedimientos utilizados en la evaluación psicológica, educativa y ocupacional.
- Los tests incluyen procedimientos para la medición de conductas tanto normales como anormales o disfuncionales.
- Los tests son procedimientos diseñados para ser aplicados bajo condiciones controladas o estandarizadas, y conllevan la utilización de protocolos de puntuación rigurosos.
- Estos procedimientos proporcionan medidas de ejecuciones y conllevan la obtención

de inferencias a partir de muestras de conducta. También pueden incluir procedimientos que proporcionan clasificaciones cualitativas u ordenamientos de las personas.

Cualquier procedimiento utilizado de la forma descrita arriba, puede considerarse como un test, independientemente de su forma de aplicación, el tipo de profesional que lo haya construido, o si requiere contestar a ítems o ejecutar ciertas tareas u operaciones.

Los tests, deben de estar apoyados por datos empíricos, sobre su fiabilidad y validez para medir los objetivos que se proponen. Hay que aportar datos que justifiquen las inferencias que se hacen a partir de las puntuaciones de los tests. Estos datos tienen que estar disponibles para los usuarios de los tests, así como para los profesionales e investigadores que deseen llevar a cabo una evaluación o revisión independiente.

Las directrices que se presentan aquí, son aplicables a todos los procedimientos descritos, se autodenominen o no tests. Deben de tenerse en cuenta para cualquier procedimiento de evaluación, utilizado en situaciones en las que la evaluación, tiene serias implicaciones para las personas, pudiendo causarles daños personales o psicológicos, si no se realiza adecuadamente. Las directrices no se aplican al uso de materiales que pueden tener una semejanza superficial con los tests, pero que los propios participantes, reconocen que se utilizan con fines de diversión o entretenimiento (por ejemplo, los cuestionarios de estilos de vida de los periódicos y revistas).

Destinatarios

Las directrices se aplican al uso de los tests en la práctica profesional, por tanto van dirigidas fundamentalmente a:

- Los compradores y vendedores de materiales psicométricos.
- Los responsables de elegir los tests y el uso que se hará de ellos.

- Quienes aplican, puntúan e interpretan los tests.
- Quienes aconsejan a otros basándose en los resultados de los tests (consultores en ámbitos educativos y laborales, orientadores escolares y profesionales, etc.)
- Las personas encargadas de dar información sobre los resultados de los tests a las personas evaluadas.
- Los constructores de tests.
- Los editores de tests.
- Las personas implicadas en el entrenamiento de los usuarios de los tests.
- Los propios evaluados y las personas cercanas a ellos (padres, cónyuges, etc.)
- Organizaciones profesionales y otras asociaciones interesadas en el uso de los tests
- Legisladores y responsables políticos.
- Usuarios de los tests con fines de investigación.

Si bien las directrices no pretenden cubrir todo tipo de técnica o situación evaluativa, pero muchas de las directrices pueden ser aplicables en situaciones evaluativas que no se identifican estrictamente con la práctica de los tests, tales como los centros de evaluación para la selección de personal, las entrevistas estructuradas y semiestructuradas, orientación escolar y profesional, etc.

Aspectos contextuales

Las directrices son aplicables internacionalmente, pero a la hora de utilizarlas a nivel local en un país determinado, deben de tenerse en cuenta diversos factores o aspectos que pueden modular su aplicación. Entre otros, estos factores contextuales incluyen:

- Diferencias sociales, políticas, institucionales, lingüísticas y culturales.
- La legislación del país en el que se utilizan los tests.
- Directrices y estándares ya existentes establecidos por asociaciones profesionales.
- Diferencias entre la evaluación individual y la colectiva.
- Diferencias en función de la situación evaluativa (educativa, clínica, trabajo, otras).
- Quienes son los receptores de los resultados (los propios evaluados, los padres o

- responsables legales, los constructores de los tests, los empleadores, etc.)
- Diferencias según el uso de los resultados de los tests (por ejemplo, toma de decisiones frente a orientación).
 - Grado en el que la situación permite la posibilidad de comprobar la precisión de las predicciones y su posible modificación a la luz de los resultados posteriores.

Conocimientos

Unos conocimientos psicológicos psicométricos sólidos y una comprensión profunda de todos los aspectos implicados en el proceso evaluativo, constituyen la base fundamental para el uso pertinente de los tests. Los expertos, suelen estar de acuerdo en que la causa más importante del uso inapropiado de los tests, es una formación deficiente de los usuarios. Los conocimientos concretos requeridos en cada caso, variarán en función de la situación y área de aplicación, por tanto, no es fácil describir de forma general, los conocimientos requeridos para la utilización adecuada de los tests en todas las situaciones, depende de cada caso. Por ejemplo, no se requieren los mismos conocimientos para aplicar un test colectivo a un grupo, que para interpretar las puntuaciones, o construir el propio test, por citar un caso obvio. Las directrices no hacen descripciones detalladas de estos conocimientos, ni lo pretenden. No obstante, a la hora de aplicarlas a una situación concreta, habrá que tener muy presente y describir con precisión los conocimientos requeridos para el caso. Esta descripción, ha de cubrir las distintas áreas de conocimientos implicadas en la situación de que se trate, entre ellas deberían de incluirse las siguientes: a) Conocimientos sobre teoría de los tests y propiedades técnicas de los tests, tales como fiabilidad, validez, estandarización, sesgo, análisis de ítems, etc. b) Conocimiento de los tests y principios de la medición para entender adecuadamente los resultados. c) Conocimientos sobre la teoría, modelos y constructos medidos, que permita una elección pertinente de las pruebas e interpretación de los resultados. d) Conocer el abanico de tests disponibles, así como los editores correspondientes. e) Avances técnicos recientes, tales como los tests informatizados, bancos de ítems, etc. Otros muchos conocimientos específicos, han de ponerse en funcionamiento para llevar a cabo una utilización adecuada de los tests, cada una de las directrices puede servir como indicador de la naturaleza de éstos.

1. Uso ético de los tests.

Los usuarios competentes deberían:

1.1 Actuar de forma ética y profesional.

1.1.1. Mantener y promover estándares éticos y profesionales.

1.1.2. Estar al corriente de los debates profesionales y éticos sobre el uso de los tests en su campo de especialización.

1.1.3. Llevar a cabo una política explícita sobre los tests y su uso (véase ejemplo en el Apéndice A).

1.1.4. Asegurarse de que las personas para las que trabajan, o con quienes trabajan, se acogen a estándares éticos y profesionales adecuados.

1.1.5. Actuar con el debido respeto para la sensibilidad de las personas evaluadas y de otras personas o instituciones implicadas.

1.1.6. Presentar la práctica de los tests de forma positiva y equilibrada cuando interactúan con los medios de comunicación.

1.1.7. Evitar las situaciones en las que pueda parecer que existen determinados intereses en los resultados de la evaluación, o en las cuales la evaluación pueda dañar su relación con los clientes.

1.2. Asegurarse de que son competentes para el uso de los tests

1.2.1. Trabajar de acuerdo con los principios científicos.

1.2.2. Establecer y mantener elevados estándares personales de competencia.

1.2.3. Conocer los límites de la propia competencia y no actuar fuera de ellos.

1.2.4. Mantenerse al día de los cambios y avances en relación con el uso de los tests y de su construcción, incluyendo los cambios de normas y legislación, los cuales pueden influir en los tests y su uso.

1.3. Responsabilizarse del uso que hacen de los tests

1.3.1. Ofrecer solo servicios y usar tests para los cuales están preparados.

1.3.2. Aceptar la responsabilidad por los tests elegidos y por las recomendaciones proporcionadas.

- 1.3.3. Dar una información clara y adecuada a los participantes en el proceso evaluativo sobre los principios éticos y las disposiciones legales que regulan el uso de los tests.
- 1.3.4. Asegurarse de que el contrato entre los evaluados y los evaluadores es claro y se ha comprendido (véase un ejemplo de contrato en el Apéndice B).
- 1.3.5. Estar atentos a cualquier consecuencia imprevista del uso de los tests.
- 1.3.6. Esforzarse para evitar cualquier tipo de daño o perjuicio a las personas evaluadas.

1.4. Asegurarse de que los materiales del test están seguros

- 1.4.1. Asegurar un almacenaje seguro y controlar el acceso a los materiales del test.
- 1.4.2. Respetar el copyright y los acuerdos que existan sobre el test, incluyendo cualquier prohibición sobre la copia y transmisión de los materiales, bien sea electrónicamente o de otra forma. Asimismo se respetarán rigurosamente los términos del acceso de otras personas, cualificadas o no, a los materiales.
- 1.4.3. Proteger la integridad del test, no entrenando a las personas con los propios materiales del test, o con otros materiales de prácticas que puedan influir de forma inapropiada en el rendimiento de las personas en el test.
- 1.4.4. Asegurarse de que la tecnología del test no se expone públicamente de tal modo que su utilidad quede deteriorada.

1.5. Asegurarse de que los resultados de los tests se tratan confidencialmente

- 1.5.1. Especificar quienes tendrán acceso a los resultados y definir los niveles de confidencialidad.
- 1.5.2. Explicar los niveles de confidencialidad antes de dar los resultados.
- 1.5.3. Limitar el acceso a los resultados únicamente a quienes tengan la necesidad y derecho a conocerlos.
- 1.5.4. Obtener las autorizaciones pertinentes antes de proporcionar los resultados a otros.
- 1.5.5. Proteger los datos archivados de tal forma que sólo puedan acceder a ellos quienes tengan derecho a hacerlo.

- 1.5.6. Establecer una directrices claras en relación con el tiempo que se van a mantener archivados los datos.
- 1.5.7. Suprimir el nombre y otros datos identificatorios de los resultados si así lo solicita la persona evaluada.
- 1.5.8. Suprimir el nombre y otros datos identificativos de las bases de datos de los resultados, con fines de investigación, elaboración de baremos, u otros tratamientos estadísticos.

2. Utilización adecuada de los tests

2.1. Estimar la utilidad potencial de los tests en una situación evaluativa

Los usuarios competentes deberían:

- 2.1.1. Ofrecer una justificación razonada para el uso de los tests.
- 2.1.2. Asegurarse de que se ha llevado a cabo un análisis riguroso de las necesidades del cliente, categoría diagnóstica, condiciones, o trabajo para el que se utilizará la evaluación.
- 2.1.3. Comprobar que los conocimientos, destrezas, aptitudes, u otras características, que miden los tests correlacionan con las conductas pertinentes en el contexto en el que se van a llevar a cabo las inferencias.
- 2.1.4. Buscar otras fuentes adicionales de información.
- 2.1.5. Sopesar las ventajas e inconvenientes de utilizar tests frente a otras fuentes de información.
- 2.1.6. Asegurarse de que se utilizan todas las fuentes de información colateral disponibles.

2.2. Elegir tests técnicamente correctos y adecuados a cada situación

Los usuarios competentes deberían:

- 2.2.1. Examinar toda la información disponible sobre los tests potencialmente adecuados antes de elegir un test concreto.
- 2.2.2. Comprobar que la documentación técnica sobre el test proporciona suficiente información para evaluar los siguientes aspectos:
 - a. Amplitud y representatividad del contenido del test, adecuación de los grupos normativos utilizados, nivel de dificultad de los contenidos, etc.

- b. Precisión de la medición y fiabilidad para las poblaciones pertinentes.
 - c. Validez para las poblaciones pertinentes y su aplicabilidad para el uso que se hace del test.
 - d. Ausencia de sesgo para los grupos con los que se utilizará.
 - e. Aceptación por parte de quienes están implicados en su uso, incluyendo la pertinencia y validez aparente percibidas.
 - f. Aspectos prácticos, tales como tiempo requerido, coste, o recursos que se necesitan.
- 2.2.3. Evitar el uso de tests que tengan una documentación técnica inadecuada o poco clara.
- 2.2.4. Utilizar tests sólo para aquellos objetivos para los cuales se dispone de una validez empírica adecuada y pertinente.
- 2.2.5. No aceptar un test basándose únicamente en su validez aparente, recomendaciones de otros usuarios, o consejos de quienes tienen intereses comerciales.
- 2.2.6. Responder a las preguntas de las personas implicadas (personas evaluadas, padres, supervisores, representantes legales, etc.), dándoles suficiente información para que entiendan por qué se eligió el test.

2.3. Prestar atención a los aspectos relacionados con el sesgo de los tests

Cuando los tests se van a utilizar con personas de diferentes grupos (por ejemplo: género, cultura, educación, etnia, origen, o edad, entre otros), los usuarios competentes de los tests harán todos los esfuerzos posibles para asegurarse de que:

- 2.3.1. Los tests son imparciales y adecuados para todos los grupos evaluados.
- 2.3.2. Los constructos que se están midiendo son relevantes para cada uno de los grupos evaluados.
- 2.3.3. Existen datos disponibles sobre las diferencias de rendimiento de los grupos en el test.
- 2.3.4. Hay datos disponibles sobre el Funcionamiento Diferencial de los Ítems cuando ello es pertinente.
- 2.3.5. Hay datos sobre la validez que apoyan el uso del test en diferentes grupos.

- 2.3.6. Se minimizan los efectos de las diferencias grupales no relacionadas con el objetivo de la medición.
- 2.3.7. Las directrices sobre la imparcialidad de los tests se interpretan dentro del marco de la legislación al respecto existente en cada país.

Cuando se utilizan los tests en más de un idioma (idiomas distintos, dialectos, lenguaje de signos, etc.) los usuarios competentes harán todos los esfuerzos posibles para asegurarse de que:

- 2.3.8. Las versiones de los distintos idiomas o dialectos hayan sido elaboradas utilizando una metodología rigurosa.
- 2.3.9. Los constructores hayan sido sensibles a los aspectos de contenido, culturales e idiomáticos.
- 2.3.10. Quienes aplican los tests sean capaces de comunicarse perfectamente en el idioma en el que se aplica el test.
- 2.3.11. El dominio de la lengua (en la que se aplicará el test) de las personas evaluadas sea comprobado sistemáticamente, utilizándose la versión más adecuada, o una bilingüe si fuese necesario.

Cuando se utilizan los tests con personas que tienen alguna discapacidad, los usuarios competentes harán todo lo que sea posible para asegurarse de que:

- 2.3.12. Se ha recabado consejo de los expertos acerca de los efectos de la discapacidad sobre el rendimiento en el test.
- 2.3.13. Se han consultado las personas a evaluar y se ha dado un tratamiento adecuado a sus necesidades y deseos.
- 2.3.14. Se han llevado a cabo los ajustes oportunos cuando se evalúa a personas con discapacidades auditivas, visuales, motoras, dislexia, u otras.
- 2.3.15. Se ha contemplado la posibilidad de utilizar procedimientos de evaluación alternativos en vez de modificaciones o ajustes de los tests.
- 2.3.16. Se ha solicitado consejo a expertos en el caso de que el grado de modificación requerido por el test esté más allá de la experiencia y conocimientos del usuario.

2.3.17. Las modificaciones, cuando sean necesarias, se ajustan a la naturaleza de la discapacidad y se han diseñado para minimizar el impacto sobre la validez de las puntuaciones.

2.3.18. La información relativa a cualquier ajuste o modificación hechos en el test o en su aplicación se comunica a quienes interpretan o utilizan las puntuaciones del test, para así facilitar una interpretación apropiada de las puntuaciones.

2.4. Hacer los preparativos necesarios para la aplicación del test

El usuario competente hará todo lo posible para:

2.4.1. Proporcionar en el momento oportuno una información clara a las personas implicadas en la evaluación acerca de la finalidad del uso de los tests, la mejor forma de prepararse para la sesión de tests y los procedimientos a seguir.

2.4.2. Aconsejar a los evaluados acerca de los idiomas y dialectos para los que es apropiado el test.

2.4.3. Proporcionar información a quienes van a hacer el test sobre el tipo de práctica permitida, así como la documentación donde pueden encontrar ejemplares y materiales oportunos.

2.4.4. Explicar claramente a los evaluados sus derechos y deberes (véase Apéndice B).

2.4.5. Obtener el consentimiento explícito de los evaluados o de sus responsables o representantes legales antes de aplicar los tests.

2.4.6. Cuando los tests sean opcionales, explicar a las personas implicadas las consecuencias de hacerlos o no, para que puedan tomar una decisión con fundamento.

2.4.7. Hacer los necesarios ajustes prácticos para asegurarse de que:

a. Los preparativos coinciden con los establecidos en el manual del test

b. Los locales y otras facilidades para aplicar los tests se han reservado con antelación, el entorno físico es accesible, seguro, tranquilo, libre de distracciones, y se ajusta a las necesidades.

- c. Hay suficiente material disponible y se ha comprobado que no han quedado señales de usuarios previos en los cuadernillos o en las hojas de respuestas.
 - d. Las personas implicadas en la aplicación de los tests son competentes.
 - e. Se han hecho los ajustes oportunos para aplicar los tests a las personas con alguna discapacidad (véase Apéndice C).
- 2.4.8. Prever posibles problemas y solventarlos mediante la preparación de materiales e instrucciones.

2.5. Aplicar los tests adecuadamente

Los usuarios competentes deberían:

- 2.5.1. Establecer una buena relación con las personas evaluadas, dándoles la bienvenida y dirigiéndose a ellas de forma positiva.
- 2.5.2. Tratar de reducir la ansiedad de las personas a las que se va a evaluar, evitando crear o reforzar ansiedad innecesaria.
- 2.5.3. Eliminar fuentes potenciales de distracción, tales como alarmas de relojes de pulsera, teléfonos móviles, buscas, etc.
- 2.5.4. Asegurarse de que todas las personas disponen de los materiales necesarios para responder al test antes de comenzar éste.
- 2.5.5. Supervisar convenientemente la aplicación de los tests.
- 2.5.6. Dar las instrucciones en la lengua dominante de las personas evaluadas siempre que sea posible, incluso cuando el test está diseñado para proporcionar datos sobre el conocimiento o dominio de una lengua distinta de la dominante.
- 2.5.7. Ajustarse estrictamente a las instrucciones del manual del test, haciendo los ajustes pertinentes para las personas con alguna discapacidad.
- 2.5.8. Leer las instrucciones pausada y claramente.
- 2.5.9. Dar el tiempo adecuado para hacer los ejemplos.
- 2.5.10. Observar y anotar las posibles desviaciones de los procedimientos estándar del test.
- 2.5.11. Registrar los tiempos de respuesta con precisión cuando se requiera.
- 2.5.12. Asegurarse de que están todos los materiales al final de cada sesión.

- 2.5.13. Realizar la aplicación de modo que permita una supervisión adecuada y una comprobación de la identidad de cada una de las personas evaluadas.
- 2.5.14. Permitir a los ayudantes hacerse cargo de la aplicación sólo si han sido entrenados adecuadamente.
- 2.5.15. Asegurarse de que durante la sesión aplicación no se deja desatendidas o sujetas a distracción a las personas evaluadas.
- 2.5.16. Proporcionar una asistencia adecuada a las personas evaluadas que muestran signos excesivos de ansiedad o desazón.

2.6. Puntuar y analizar los resultados de los tests con precisión

Los usuarios competentes deberían:

- 2.6.1. Seguir al pie de la letra los procedimientos estandarizados de puntuación.
- 2.6.2. Asegurarse de la precisión al asignar las puntuaciones, especialmente en aquellos casos en los que entra en juego el juicio de los evaluadores. Para cerciorarse de la precisión puede volver a puntuarse una muestra de las personas evaluadas, comprobando así la coincidencia entre las puntuaciones
- 2.6.3. Llevar a cabo las transformaciones de las puntuaciones directas a otros tipos de escalas pertinentes.
- 2.6.4. Elegir los tipos de escala más convenientes de acuerdo con el uso que se vaya a hacer de las puntuaciones del test.
- 2.6.5. Comprobar la precisión de las transformaciones de las escalas, así como la de cualquier tipo de análisis o tratamiento que se haga con los datos.
- 2.6.6. Asegurarse de que no se sacan conclusiones erróneas debido a la utilización de baremos desfasados, o inadecuados para las personas evaluadas.
- 2.6.7. Calcular las puntuaciones compuestas cuando proceda, utilizando las fórmulas y ecuaciones propuestas en el manual del test.
- 2.6.8. Inspeccionar los resultados para detectar posibles errores o anomalías en las puntuaciones.
- 2.6.9. Describir e identificar con precisión los resultados, normas, tipos de escalas, fórmulas, etc. utilizados.

2.7. Interpretar los resultados adecuadamente

Los usuarios competentes deberían:

- 2.7.1. Tener una buena comprensión profesional de las bases teóricas y conceptuales del test, de la documentación técnica y de las directrices para el uso e interpretación de las puntuaciones.
- 2.7.2. Tener una buena comprensión profesional de las escalas utilizadas, de las normas y baremos, así como de las limitaciones de las puntuaciones.
- 2.7.3. Tratar de minimizar cualquier sesgo que pueda existir hacia las personas evaluadas en la interpretación de las puntuaciones del test.
- 2.7.4. Utilizar normas o grupos de comparación apropiados cuando estén disponibles.
- 2.7.5. Interpretar los resultados a la luz de la información disponible sobre la persona evaluada (edad, género, escolaridad, cultura, etc.), teniendo en cuenta las limitaciones técnicas del test, el contexto de la evaluación, y las necesidades de las personas o instituciones con intereses legítimos en el resultado del proceso evaluativo.
- 2.7.6. Evitar la generalización de los resultados de un test a rasgos o características de la persona que no han sido medidos por el test.
- 2.7.7. Tener en cuenta la fiabilidad y el error de medida de cada escala, así como otros factores que puedan alterar artificialmente los resultados a la hora de interpretar las puntuaciones.
- 2.7.8. Tener muy en cuenta los datos disponibles sobre la validez del constructo medido en relación con las características de los grupos evaluados, tales como cultura, edad, clase social, género, etc.
- 2.7.9. Utilizar puntos de corte en la interpretación de las puntuaciones sólo cuando se disponga de datos empíricos sobre su validez.
- 2.7.10. Ser conscientes de los estereotipos sociales que pueden existir sobre las personas evaluadas (en relación con su cultura, edad, clase social, género, etc.), evitando interpretar los tests de forma que se perpetúen dichos estereotipos.
- 2.7.11. Tener en cuenta cualquier variación individual o colectiva que se haya hecho respecto al procedimiento estándar en la aplicación de las pruebas.

2.7.12. Tomar en consideración cualquier experiencia previa que la persona evaluada haya tenido con el test, en el caso de que se disponga de datos sobre los efectos de dicha experiencia sobre el rendimiento en la prueba.

2.8. Comunicar los resultados de forma clara y precisa

Los usuarios competentes deberían:

- 2.8.1. Identificar las personas o instituciones pertinentes que pueden recibir los resultados de los tests.
- 2.8.2. Elaborar informes orales o escritos para los receptores de los resultados, siempre con el consentimiento explícito de las personas evaluadas o de sus representantes legales.
- 2.8.3. Asegurarse de que el nivel técnico de los contenidos de los informes es adecuado para su comprensión por los receptores.
- 2.8.4. Dejar muy claro en los informes que los resultados de los tests son confidenciales, y especificar el tiempo que se mantendrán archivados los resultados.
- 2.8.5. Dejar claro que los datos de los tests representan una sola fuente de información que debe analizarse conjuntamente con otras fuentes.
- 2.8.6. Explicar el peso que debe darse a las puntuaciones de los tests en relación con otras fuentes de información acerca de las personas evaluadas.
- 2.8.7. Proporcionar la información sobre los resultados en un lenguaje comprensible para el receptor, de modo que se minimice la posibilidad de interpretaciones incorrectas.
- 2.8.8. Utilizar una forma y estructura para el informe que encaje en el contexto de la evaluación.
- 2.8.9. Si procede, proporcionar información a quienes toman las decisiones acerca de como pueden usar los resultados de los tests para mejorar sus decisiones.
- 2.8.10. Explicar y fundamentar la utilización que se hace de los resultados en los tests para la clasificación de las personas en categorías con fines diagnósticos, u otros.
- 2.8.11. Incluir un resumen claro en los informes escritos, y, cuando sea pertinente, recomendaciones concretas.

2.8.12. Dar información a las personas evaluadas de forma constructiva y positiva.

2.9. Revisión de la adecuación del test y de su uso

Los usuarios competentes deberían:

2.9.1. Seguir y revisar periódicamente los posibles cambios en la población de personas evaluadas, así como los criterios utilizados para la validez.

2.9.2. Estar atentos a posibles impactos negativos de los tests.

2.9.3. Ser conscientes de la necesidad de reanalizar la utilización de un test si se ha llevado a cabo algún cambio en su forma, contenidos, o forma de aplicación.

2.9.4. Tener presente la necesidad de reconsiderar la validez del test si se ha cambiado la finalidad para la que se utilizaba.

2.9.5. Siempre que sea posible, validar los tests para los usos para los que fueron elaborados y participar en los trabajos rigurosos de validación que se lleven a cabo.

2.9.6. Ayudar en la medida de sus posibilidades a mantener al día la información sobre los baremos, fiabilidad y validez del test, proporcionando los datos pertinentes a los constructores, editores o investigadores.

Apéndice A: Directrices para el establecimiento de políticas sobre el uso de los tests

Las directrices que siguen, se refieren a la necesidad de las organizaciones de considerar su política sobre la práctica de los tests de forma sistemática, asegurándose de que todas las personas implicadas en la evaluación, tienen clara esta política. La necesidad de una política explícita sobre la práctica de los tests, no atañe sólo a las grandes organizaciones. Las empresas de tamaño medio o pequeño, al igual que las grandes, deben de prestar atención a la política sobre la práctica de los tests, de la misma forma que lo hacen en relación con la salud y seguridad laboral, igualdad de oportunidades, discapacidades, u otras áreas relacionadas con una práctica adecuada en el tratamiento y atención al personal.

Aunque las siguientes consideraciones pueden necesitar una adaptación para su utilización por parte de usuarios individuales de los tests, cuando actúan como profesionales, sigue siendo importante que tengan muy clara su propia política y que puedan comunicarla a otros.

Los objetivos de una política sobre el uso de los tests serían:

- Asegurar que se alcanzan las metas personales y organizacionales.
- Evitar el posible uso inadecuado de los tests.
- Demostrar explícitamente el compromiso con una práctica adecuada.
- Asegurar que el uso de los tests se ajusta a los fines establecidos.
- Asegurar que los tests se utilizan de forma no discriminativa.
- Asegurar que las evaluaciones se basan en información comprensiva y pertinente.
- Asegurar que los tests son utilizados por personas cualificadas para ello.

Una política sobre el uso de los tests deberá abordar los siguientes aspectos:

- Uso adecuado de los tests.
- Seguridad de los materiales y puntuaciones.
- Quién puede aplicar, puntuar e interpretar los resultados.

- Cualificaciones sobre las personas que van a usar los tests.
- Entrenamiento de los usuarios.
- Preparación de las personas a evaluar.
- Acceso a los materiales y seguridad de los tests.
- Acceso a los resultados y confidencialidad de las puntuaciones.
- Información sobre los resultados a las personas evaluadas.
- Responsabilidades hacia los evaluados, antes, durante y después de la sesión de tests.
- Responsabilidades individuales de los usuarios de los tests.

Cualquier política que se establezca sobre tests, tiene que ser revisada periódicamente, actualizándola de acuerdo con los avances producidos en los tests o en su utilización. Hay que permitir el acceso e informar, sobre la política establecida a las partes implicadas en la utilización de los tests. La responsabilidad de la política de una organización, debe recaer en un usuario cualificado, que tenga autoridad para asegurar su implementación y cumplimiento.

Apéndice B: Directrices para desarrollar contratos entre las partes implicadas en la evaluación

Los contratos entre los usuarios de los tests y las personas evaluadas, deben de estar de acuerdo con los criterios de una práctica adecuada, la legislación correspondiente y la política establecida sobre el uso de los tests. Los puntos que siguen, se ofrecen, como un ejemplo del tipo de aspectos que un contrato de este tipo debe abarcar. Los detalles variarán en función del contexto de la evaluación (ocupacional, educativo, clínico, jurídico, etc.), y de las leyes y regulaciones locales y nacionales.

Los contratos entre los usuarios, las personas evaluadas y otras partes implicadas, aparecen con frecuencia implícitos, no formalizados explícitamente, al menos en parte. Dejando claros los objetivos, el papel y las responsabilidades de cada una de las partes, se ayuda a evitar los malentendidos, los perjuicios y la necesidad de acudir a los tribunales.

El usuario se compromete a:

1. Informar a las personas evaluadas de sus derechos, en cuanto a cómo se utilizarán sus puntuaciones de los tests, así como sus derechos de acceso a ellas.
2. Advertir con anterioridad suficiente sobre cualquier carga económica que conlleve el proceso de evaluación, quién es el responsable de los pagos y fechas en las que deben de hacerse efectivos.
3. Tratar a las personas evaluadas con cortesía, imparcialidad y respeto, independientemente de su raza, género, edad, discapacidad, etc.
4. Utilizar tests de calidad probada, adecuados a las personas evaluadas y a la situación de evaluación.
5. Informar antes de la aplicación de los tests, sobre la finalidad de la evaluación, el tipo de test, a quién se enviarán los resultados y el uso que hará de ellos.
6. Avisar con antelación donde se aplicarán los tests, cuando se darán los resultados, y si pueden o no obtener una copia del test una vez hecho, de su hoja de respuestas, o de sus puntuaciones.

7. Utilizar personas cualificadas para la aplicación de los tests y la interpretación de los resultados.
8. Asegurarse de que las personas a evaluar, saben si el test es opcional o no, y, cuándo lo sea, asegurarse de que conocen las consecuencias de hacerlo o no hacerlo.
9. Asegurarse de que las personas que van a hacer el test, conocen las condiciones (si las hubiera) en las que podrían volver a hacer el test, ser calificados de nuevo, o anulárseles la prueba.
10. Asegurarse de que las personas evaluadas, saben que se les explicarán los resultados después de hacer el test, tan pronto como sea posible y en términos que les resulten comprensibles.
11. Asegurarse de que las personas evaluadas, saben que sus resultados son confidenciales hasta el punto que permitan la ley y una práctica profesional adecuada.
12. Informar a las personas evaluadas, sobre quienes tendrán acceso a los resultados y las condiciones bajo las cuales se proporcionarán.
13. Asegurarse de que las personas evaluadas, conocen los procedimientos para elevar quejas o comunicar algún problema.

El usuario comunicará a las personas evaluadas que se espera de ellas:

14. Un trato cortés y respetuoso con el resto de las personas durante el proceso de evaluación.
15. Preguntar antes del test, si no están seguras acerca de por qué se aplica el test, cómo se aplicará, qué tienen que hacer exactamente, y que se hará con los resultados.
16. Informar sobre cualquier circunstancia que consideren que puede invalidar los resultados del test, o que deseen que se tenga en cuenta.
17. Seguir las instrucciones de la persona que aplica el test.
18. Ser consciente de las consecuencias que se siguen de no hacer el test, si deciden no hacerlo, y estar preparado para aceptar dichas consecuencias.
19. Si se requiere abonar alguna cantidad para hacer la prueba, los pagos se harán en la fecha acordada.

Apéndice C

Aplicación de los tests a personas con alguna discapacidad

Cuando se modifica la forma de aplicación de un test para aplicarlo personas con alguna discapacidad, hay que proceder con sumo cuidado y experiencia. Como siempre, si existe alguna legislación local o nacional específica debe respetarse, así como el derecho a la privacidad de la persona evaluada. Cuando se recoge la información sobre los tipos y niveles de discapacidad, hay que limitarse a la información relativa a la capacidad de cada persona para llevar a cabo las tareas requeridas por el test. Hay que poner especial cuidado cuando los tests se utilizan en el ámbito ocupacional, bien sea para selección, promoción, u otros fines.

No hay una regla sencilla para aplicar un test correctamente, que sea generalizable a todas las personas con algún tipo de discapacidad. Queda a juicio del profesional si es mejor utilizar una forma alternativa de evaluación o modificar el test, o su forma de aplicación. En la práctica, no suele ser posible baremar los tests modificados en muestras amplias de personas con discapacidades equivalentes, de modo que se pueda asegurar la comparabilidad del test con la versión estándar. No obstante, cuando haya datos al respecto, por ejemplo, en el caso de los efectos producidos por la modificación del tiempo de aplicación, el uso de Braille, o versiones auditivas de los tests en cassette, deben de utilizarse como guía para las modificaciones que se lleven a cabo. Aunque una estricta estandarización de la versión modificada, tal vez no sea posible, siempre que lo sea, deben de llevarse a cabo estudios piloto con muestras pequeñas.

Dada la escasez de información sobre el rendimiento en los tests (modificados o no) de las personas con alguna discapacidad, suele ser más adecuado utilizar los resultados de forma más cualitativa. Puede tomarse como una indicación de la característica evaluada, complementándola con información recogida por otros métodos.

En el caso de aplicaciones individuales, el evaluador suele poder ajustar los procedimientos evaluativos a las capacidades de la persona evaluada. Sin embargo, en

el caso de aplicaciones colectivas, aparecen problemas especiales, como ocurre por ejemplo en selección de personal. En estos casos, puede haber dificultades prácticas para modificar la forma de aplicación para determinadas personas dentro del marco de una aplicación colectiva. Además, los tratamientos diferenciados, podrían ser vistos por algunos como discriminatorios. Por ejemplo, si se concede más tiempo para hacer el test, las personas discapacitadas podrían tener la sensación de que son tratadas de forma diferente, mientras que quienes no tienen ninguna discapacidad, pueden pensar que el tiempo extra concedido proporciona a los otros una ventaja.

Consejos sobre necesidades especiales, pueden obtenerse habitualmente de las organizaciones de discapacitados, así como de las propias personas con discapacidades a las que se van aplicar los tests. Suele ser de gran ayuda dirigirse directamente a las personas y preguntarles si hay alguna circunstancia que consideren que debe de tenerse en cuenta. En muchos casos, esta consulta, va a permitir llevar a cabo las modificaciones oportunas en el contexto de aplicación de la prueba, sin tener que modificar el propio test.

Las siguientes reglas generales, pueden utilizarse como guía a la hora de decidir si se modifica el proceso evaluativo y cómo hacerlo.

1. Si la discapacidad no afecta al rendimiento en el test, no es necesario hacer ajustes en la prueba.
2. Cuando la discapacidad influye en el rendimiento en el test, hay que distinguir si esta influencia es incidental o forma parte del constructo medido. Si es incidental, hay que modificar el test, pero si forma parte del constructo medido no. Por ejemplo, una persona con artritis en las manos, tendrá problemas con un test de velocidad que conlleve escribir. Si la capacidad para escribir rápidamente formase parte del constructo medido, entonces no debería de cambiarse el test. Sin embargo, si la finalidad fuese, por ejemplo, medir la rapidez de percepción visual, habría que buscar una forma de respuesta a la prueba más adecuada, pues la artritis estaría influyendo negativamente en el constructo medido.

3. Cuando una discapacidad ajena al constructo medido influye en el rendimiento del test, deben de llevarse a cabo los ajustes convenientes en la prueba.
4. Los usuarios, deben de consultar siempre el manual del test, y los editores, buscar información sobre las modificaciones de las pruebas y sobre posibles formatos y procedimientos alternativos.
5. Los usuarios de tests, deben de consultar a las organizaciones correspondientes de discapacitados, acerca de las implicaciones de una discapacidad concreta, la documentación sobre la discapacidad, y el tipo de modificaciones que podrían ser convenientes.
6. Cualquier tipo de modificación que se haga en el test o en el proceso de aplicación, debe de estar rigurosamente documentada, exponiendo claramente las razones que justifican la modificación.

CONCLUSIONES

La mayoría de los expertos, así como las organizaciones profesionales nacionales e internacionales, coinciden en señalar que una de las medidas más eficaces para mejorar la utilización que se hace de los tests, es la de proporcionar una buena información y formación a los usuarios. Un profesional con una buena información sobre las pruebas y una formación adecuada, difícilmente utilizará de forma incorrecta los tests. Enmarcado en esa filosofía, en 1995 el Colegio Oficial de Psicólogos (COP) creó una Comisión de Tests con el fin de analizar los problemas implicados en el uso de estos. En el presente trabajo se describe el proceso diseñado por la Comisión de Tests del COP para llevar a cabo la revisión de los tests editados en España. También se presenta el cuestionario elaborado a tal efecto, un estudio de adaptación a dialectos y un caso concreto de adaptación del francés al castellano. En líneas generales, el procedimiento de evaluación propuesto, consiste en revisar los tests de forma sistemática por parte de expertos, utilizando el cuestionario de recogida de datos diseñado para tal fin, inspirado en los modelos holandés e inglés actualmente en funcionamiento. Cabe esperar que la publicación de los resultados de las revisiones de los tests, ayude a los profesionales a elegir y evaluar las pruebas que tienen a su disposición en el mercado, al menos, ése ha sido el objetivo que ha guiado a la comisión de tests del COP para llevar a cabo esta iniciativa. Las directrices de la ITC ofrecidas, podrían sufrir alguna modificación menor, por otra parte, esta Comisión, espera que las directrices sean útiles para muchas organizaciones implicadas en el proceso de adaptación y contribuyan a mejorar la validez de las investigaciones interidioma e intercultural.

En la actualidad, se están sopesando algunas ideas para difundir las directrices, una posibilidad sería, recoger el conjunto de pasos rigurosos a seguir por los usuarios cuando adaptan un test y establecer las equivalencias correspondientes, se trataría de integrar las 22 directrices, en un conjunto de pasos ordenados, para llevar a cabo la adaptación de los tests, que pueden incluir actividades como determinar la necesidad del proyecto de adaptación del test, y establecer la equivalencia del constructo de interés en los múltiples idiomas y/o culturas, seleccionar los traductores y escoger un

diseño de traducción, llevar a cabo el diseño y hacer revisiones pertinentes del test adaptado, determinar el uso que se espera del test y diseñar estudios para recoger datos sobre la validez, diseñar, si fuese necesario, un estudio para ubicar y/o definir las puntuaciones en una escala común, documentar el proceso y la evidencia sobre la validez.

NORMAS MÍNIMAS PARA EL USO DE LOS TESTS

1. Evitar errores al puntuar y registrar los resultados.
2. Abstenerse de etiquetar a las personas con términos despectivos basándose en las puntuaciones de los tests.
3. Mantener la seguridad de las plantillas y resto de materiales.
4. Asegurarse que todos los examinandos siguen las instrucciones.
5. Aplicar los tests en unas condiciones que permitan a los examinados un rendimiento óptimo.
6. Abstenerse de entrenar a las personas en los ítems del test.
7. Estar dispuesto a interpretar las puntuaciones y aconsejar a los examinados en sesiones diseñadas para ello.
8. No hacer fotocopias del material psicotécnico.
9. Abstenerse de utilizar hojas de respuesta caseras que pueden no ajustarse con precisión a la plantilla.
10. Establecer una buena relación con los examinados.
11. Abstenerse de responder preguntas de los examinados con mayor detalle que lo permitido por el manual del test.
12. No asumir que una norma para un trabajo vale sin más para otro diferente, y no asumir que las normas válidas para un grupo son automáticamente aplicables a otro distinto.

(Adaptado de: Moreland, K.L., Eyde, L.D., Robertson, G.J., Primoff, E.S. y Most, R.B. (1995). Assessment of test user qualifications. *American Psychologist*, 5, 14-23).

Comisión de Tests del Colegio Oficial de Psicólogos

José Muñiz (Coordinador)
Universidad de Oviedo

M^a Teresa Anguera
Universidad de Barcelona

Rocío Fernández Ballesteros
Universidad Autónoma de Madrid

José Ramón Fernández Hermida
Universidad de Oviedo

Manuel García Pérez
ALBOR-COHS

Miguel Martínez
EOS

Rosario Martínez Arias
Universidad Complutense de Madrid

Eduardo Montes
COP

Gerardo Prieto
Universidad de Salamanca

Carlos Rodríguez Sutil
Universidad Complutense de Madrid

Nicolás Seisdedos
TEA.

REFERENCIAS BIBLIOGRÁFICAS

Aaronson N, Alonso J, Burnam A, Lohr K.N, Patrick DL, Stein R. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Qual Lif Res* 2002; 11:193-205

Alonso J, Antó JM. Instrumentos de medida de calidad de vida relacionada con la salud: características generales y proceso de adaptación transcultural. *Quaderns CAPS* 1990; 14: 16-24.

Alonso J. La medida de la calidad de vida relacionada con la salud en la investigación y en la práctica clínica. *Gac Sanit* 2000; 14:163-67.

Alonso J, Prieto L, Antó JM. La versión española del «SF-36 Health Survey» (Cuestionario de salud SF-36): un instrumento de medida de resultados clínicos. *Mcd Clin (Barc)* 1995; 104:771-6.

Alonso J, Antó JM, Moreno C. Spanish version of the Nottingham health profile: translation and preliminary validity. *A.IP11* 1990; 80:704-8.

Anastasi, A. (1987). What test users should know about the interpretation os test scores. Keynote address at Joint Committee on Testing Practices Second Test Publishers Conference, Rockville, Maryland. (Citado de Fremer, 1996).

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for Educational and Psychological Testing*. Washington DC: American Psychological Association.

Bartram, D., Lindley, P.A.y Foster, J.M. (1992 a). *A Review of Psychometric Tests for Assessment in Vocational Training*. Leicester: BPS Books

Bartram, D. (1995). The Development of Standards for the Use of Psychological Tests in Occupational Settings: The Competence Approach. *The Psychologist*, May, 219-223.

Bartram, D. (1996). Test Qualifications and Test Use in the UK: The Competence Approach. *European Journal of Psychological Assessment*, 12, 62-71..

Bartram, D., Anderson, N., Kellett, D., Lindley, P.A.& Robertson, I. (Eds.). (1995). *Review of Personality Assessment Instruments (Level B) for use in occupational settings*. Leicester: BPS Books.

Bartram, D., Burke, E., Kandola, R., Lindley, P., Marshall, L. & Rasch, P. (Eds.). (1997). *Review of Tests of Ability and Aptitude (Level A) for use in occupational settings*. Leicester: BPS Books.

Bartram, D., Lindley, P.A.y Marshall, L. (1992 b). *Update to the Review of Psychometric Tests for Assessment in Vocational Training*. Leicester: BPS Books.

Beaton DE, Bombardier C, Cruillemin F, Bosi-Ferraz M. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000; 25:3186-91

Bullinger M, Alonso J, Apolonc G, Leplége A, Sullivan M. et al. Translating health status questionnaires and evaluating their quality: the IQOLA approach. *J Clin Epidemiol* 1998; 51:913-23

Canadian Psychological Association. (1987). *Guidelines for Educational and Psychological Testing*. Ottawa: Canadian Psychological Association.

Colegio Oficial de Psicólogos (1987). *Código Deontológico*. Madrid: Colegio Oficial de Psicólogos.

Colegio Oficial de Psicólogos (1999). *Página web del COP, Comisión de Tests*: <http://www.cop.es/test>

Eiser C, Morsc R. A review of measures of quality of life for children with chronic illness. *Ardí Dis Child* 2001; 84:205-211

Etxeberria, F. (1999). *Bilingüismo y educación en el país del euskera*. Donostia: Erein.

Evers, A. (1996). Regulations concerning test qualifications and test use in The Netherlands. *European Journal of Psychological Assessment*, 12, 153-159.
Kauffman, A.S. (1982). *Psicometría razonada con el WISC-R*. México: El Manual Moderno

Eyde, L. D., Moreland, K. L. & Robertson, G. J. (1988). *Test User Qualifications: A Data-based Approach to Promoting Good Test Use*. Report for the Test User Qualifications Working Group. Washington DC: American Psychological Association.

Eyde, L. D., Robertson, G. J., Krug, S. E. et al (1993). *Responsible Test Use: Case Studies For Assessing Human Behaviour*. Washington DC: American Psychological Association.

Fayers MP, Machin D. *Quality of life: assessment, analysis and interpretation*. Chichester: John Wiley Sons; 2000
Fernández Ballesteros, R. (1993). Evaluación psicológica en sus contextos de aplicación. *Revista de Historia de la Psicología*, 14, 97-114.

Ferrer M, Alonso J, Prieto L, Plaza V, Monso E, Marrades R, et al. Validity and reliability of the St. George's Respiratory Questionnaire after adaptation to a different language and culture: the Spanish example. *Eur Respir J* 1996; 9: 1160-6.

Franca Tarragó, O. (1996). *Ética para psicólogos*. Bilbao: Desclée de Brouwer

Fremer, J., Diamond, E. E. & Camara, W. J. (1989). Developing a Code of Fair Testing Practices in Education. *American Psychologist*, 44, 1062-1067.

Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J* 46:1417-32.

Guyatt Gil, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Int Med* 1993; 118:622-9

Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-244.

Hamers, J. y Blanc, M. (1983). *Bilinguisme et bilingualité*. Bruselas: Pierre Mardaga.

Hamers, J. y Blanc, M. (1989). *Bilinguality and bilingualisme*. Cambridge: Cambridge University Press

Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health related quality of life measures. *Qual Life"Res* 1993; 2:441-9

Hennesy CH, Moriarty DG, Zach MM, Scherr P, Brackbill R. Measuring health-related quality of life for public health surveillance. *Public Health Rcp* 1994; 109:665-772.

Herdman M, Fox-Rushby, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res* 1998; 7:323-335

Isasi, X. (1991). *Hizkuntzen arteko ukipen egoera. Ikuspegi metodologikoa*. Bilbo: Servicio editorial de la Universidad del País Vasco/Euskal Herriko Unibertsitatea

Joint Committee on Testing Practices. (1988). *Code of Fair Testing Practices in Education*. Washington DC: Joint Committee on Testing Practices.

Kendall, I., Jenkinson, J., De Lemos, M. & Clancy, D. (1997). *Supplement to Guidelines for the use of Psychological Tests*. Australian Psychological Society.

Moreland, K. L., Eyde, L. D., Robertson, G. J., Primoff, E. S. & Most, R. B. (1995). *Assessment of Test User Qualifications: A Research-Based Measurement Procedure*. *American Psychologist*, 50, 14-23.

Muñiz, J. (1997). Aspectos éticos y deontológicos de la evaluación psicológica. En A. Cordero (Ed.), *La evaluación psicológica en el año 2000*. Madrid: TEA Ediciones.

Muñiz, J. y Fernández-Hermida, J.R. (2000). La utilización de los tests en España. *Papeles del Psicólogo*, 76, 41-49.

Muñiz, J. y Hambleton, R. K. (1996). Directrices para la traducción y adaptación de los tests. *Papeles del Psicólogo*, 66, 63-70.

Muñiz, J. (1998). *Teoría clásica de los Tests*. Madrid. Pirámide

Muñiz, J., Prieto, G., Almeida, L. y Bartram, D. (1999). Test use in Spain, Portugal and Latin American countries. *European Journal of Psychological Assessment*, 15, 151-157

Pinto-Prades JL. Calidad de vida y asignación de recursos sanitarios. *Gac Sanit* 2000; 14:168-74

Rajmil L, Estrada MU, Herdman M, Serra-Sutton V, Alonso J. Calidad de vida relacionada con la salud (CVRS) en la infancia y en la adolescencia: revisión de la literatura y de los instrumentos adaptados en España. *Gac Sanit* 2001; 15:34-43.

Rajmil L, Serra-Sutton V, Alonso J, Starfield B, Riley A, Vázquez JR. The Spanish version of the Child Health and Illness Profile, Adolescent Edition (C1IIP-AL

Rodríguez Sutil, C. (1996). La ética de la devolución en el psicodiagnóstico clínico. *Papeles del Psicólogo*, 66, 91-94.

Sanjuás C, Alonso J, Sanchís J, Casan P, Broquetas JM, Ferrie PJ, Juniper EF, Antó JM. Cuestionario de calidad de vida en pacientes con asma: la versión española del Asthma Quality of Life Questionnaire. *Archivos de Bronconeumología* 1995; 31: 219-26.

Schafer, W. D. (1992). *Responsibilities of Users of Standardized Tests: RUST Statement Revised*. Alexandria, VA: American Association for Counseling and Development.

Shumaker S, Naughton M. The International assessment of health-related quality of life: a theoretical perspective. En: Shumaker S, Berton R, editors. *The International assessment of health-related quality of life: theory, translation, measurement and analysis*. Oxford: Rapid Communications; 1995.

Siguán, M.(1994). *Conocimiento y uso de las lenguas en España*. Madrid: CIS

Siméoni MC, Auquier P, Antoniotte S, Sapin C, San Marco JL. Validation of a French health-related quality of life instrument for adolescents: The VSP-A. *Qual Life Res* 2000; 9:393-403.

Siméoni MC, Sapin C, Antoniotti S, Auquier P. Health-related quality of life reponed by French adolescents: a predictive approach of health status? *J Adolcsc Health* 2001; 28:288-94.

Siméoni MC, Spain C, Antoniotti S, Michel G, Auquier P. Validation of the VSP-AM: A French HRQL questionnaire for ill adolescents. *Qual Life Res* 2000; 9:298.

Siméoni MC, Sapin C, Clement A, El Kiammar M, Auquier P. The core VSP-A and its specific module: a generic HRQL instrument for ill and healthy adolescents. *Qual Life Res* 2001; 10:267.

Tauler E, Vilagut G, Grau G, González A, Sánchez E, Figueras G, Valí O, Ferrer M, Alonso J. The Spanish version of the Paediatric Asthma Quality of Life Questionnaire (PAQLQ): Metric characteristics and equivalence with the original version. *Qual Lif Res* 2001; 10:81-91

Van de Vijver, F. & Hambleton, R. (1996). Translating tests: some practical guidelines. *European Psychologist*, 1, 89-99.