

Análisis de la validez del programa de simulación 3D My-School para la detección de alumnos en riesgo de consumo de drogas y acoso escolar*

Analysis of the validity of the My-School 3D simulation program for the detection of students at risk of drug use and bullying

Fecha de recepción: 06 Febrero 2013 | Fecha de aprobación: 01 Diciembre 2017

ADOLFO J. CANGAS^a

Universidad de Almería, España

ORCID: <http://orcid.org/0000-0002-5646-5582>

JOSÉ A. CARMONA

Universidad de Almería, España

ÁLVARO I. LANGER

Universidad Austral de Chile, Chile

JOSÉ GALLEGO

Universidad de Almería, España

ANTONIO SCIOLI

Asrem Azienda Sanitaria Regionale Molise, Italia

^a Autor de correspondencia. Correo electrónico: ajcangas@ual.es

Para citar este artículo: Cangas, A. J., Carmona, J. A., Langer, Á. I., Gallego, J., & Scioli, A. (2018). Análisis de la validez del programa de simulación 3D My-School para la detección de alumnos en riesgo de consumo de drogas y acoso escolar. *Universitas Psychologica*, 17(2), 1-11. doi: <https://doi.org/10.11144/Javeriana.upsy.17-2.avps>

RESUMEN

El presente escrito está dirigido al análisis de las propiedades psicométricas del programa de simulación 3D My-School (MS), el cual fue diseñado para detectar alumnos en riesgo de consumo de drogas y acoso escolar. En concreto, se han llevado a cabo dos estudios empíricos encaminados a evaluar, por un lado, la validez de contenido de las escenas, y por otro la validez de criterio (predictiva) del programa informático. Si atendemos a la evaluación de la validez de contenido, los resultados mostraron la existencia de un acuerdo sustancial entre el total de jueces expertos con respecto a la relevancia, coherencia, claridad y poca tendenciosidad de las situaciones. Con respecto a la validez de criterio, no existieron diferencias entre las respuestas dadas en el programa y las entrevistas, el programa informático MS habría mostrado poseer adecuadas propiedades de validez para detectar alumnos en riesgo de consumo de sustancias y acoso escolar.

Palabras clave

realidad virtual; consumo de drogas; acoso escolar; propiedades psicométricas; evaluación.

ABSTRACT

The present paper analyses the psychometric properties of the My-School (MS) 3D simulation program, which was designed to detect students at risk of drug abuse and bullying. In particular, two empirical studies have been carried out to evaluate, on the one hand, the content validity of the scenes, and on the other, the validity criterion (predictive) of the software. Specifically, in terms of the evaluation of content validity, the results showed the existence of a substantial agreement among the expert judges

regarding the relevance, coherence, clarity and low bias of situations. In relation to the criterion validity, there were no differences between the answers given in the program and the interviews. The MS software thus has adequate validity properties for detecting students at risk of substance use and bullying.

Keywords

virtual reality; drug use; bullying; psychometric properties; assessment.

En educación cada vez es más común recurrir a las nuevas tecnologías para el aprendizaje de diversas habilidades académicas (como pueden ser destrezas matemáticas, velocidad lectora, aprendizaje de idiomas, etc.). De hecho, existen multitud de recursos electrónicos, disponibles a través de la Web, para la ayuda en el aprendizaje de estas materias. Igualmente, cada vez es más frecuente el uso de entornos exploratorios basados en tecnología 3D. Estos pueden tener una serie de ventajas sobre procedimientos tradicionales de enseñanza, como es el caso de la posibilidad de realizar un aprendizaje “multimodal”, o la de facilitar los procesos de transferencia de las habilidades aprendidas de un contexto a otro. Así, el aprendizaje puede darse de una manera más *experiential* y menos limitado que en las aproximaciones tradicionales (Freitas & Neumann, 2009). Por ejemplo, un estudio de Kim, Park, y Baek (2009) mostró que el uso de videojuegos, cuando se realiza junto con determinadas estrategias meta-cognitivas, produce una mejora tanto del aprendizaje académico como del propio desempeño en el juego.

Por otro lado, también existen cada vez más recursos informáticos encaminados a la evaluación de los diferentes problemas psicológicos habituales en los centros educativos, como pueden ser el consumo de sustancias, acoso escolar, trastorno por déficit de atención, fobia escolar, ansiedad ante los exámenes, etc (Gutiérrez-Maldonado, Alsina-Jurnet, Carvallo-Becú, Letosa-Porta, & Magallón-Neri, 2007; Carmona, Cangas, & Langer, 2012; Kyriakakis, Van Rooyen, & Sisemore, 2000; McCabe, Boyd, Young, Crawford, & Pope, 2005; Zoll, Enz, Schaub, Aylett, & Paiva, 2006).

Desde luego, el empleo de los recursos electrónicos en la población juvenil posee varias ventajas en relación a los medios tradicionales de evaluación psicológica. Por un lado, los adolescentes están muy familiarizados con las nuevas tecnologías, que tienen para ellos un gran atractivo e interés. Por otro lado, son procedimientos que se pueden aplicar fácilmente a poblaciones muy amplias, es posible dar retroalimentación inmediata a los participantes, se pueden incluir más recursos técnicos (como luces y sonidos) que ayudan a centrarse más en la tarea, etc. (Bates & Cox, 2008; Buchanan, Johnson, & Goldberg, 2005; Johnson, 2005; Reips, 2000).

Recientemente, en la universidad de Almería se ha desarrollado una herramienta informática que utiliza entornos de simulación 3D para evaluar el comportamiento del alumno ante situaciones de acoso escolar y consumo de sustancias. Diseñada para ser utilizada como herramienta de evaluación tipo *screening* en ordenadores convencionales, el programa de simulación denominado My-school (MS) consta de un total de 17 escenas, las cuales muestran diversas situaciones conflictivas de consumo de sustancias y acoso escolar, ante las cuales el alumno ha de indicar cuál sería su manera de responder al verse inmerso en ese tipo de situaciones de riesgo. Una descripción pormenorizada del programa informático puede ser consultada en Carmona, Cangas, García, Langer, y Zárate (2012) y Carmona, Espínola, Cangas, e Iribarne (2010a, 2010b, 2011).

El programa MS ha mostrado una fiabilidad buena (Carmona et al., 2012), ya que ahora completa sus pruebas psicométricas con medidas de validez. Específicamente, en el presente estudio se han utilizado dos procedimientos habituales en la construcción de instrumentos de evaluación nuevos, como son el estudio de la validez de contenido y la validez de criterio predictiva (López Feal, 1986). Así, a partir del criterio de expertos, se persigue evaluar la pertinencia del contenido de las diferentes situaciones de consumo de sustancias y acoso escolar que pretende evaluar el programa informático; asimismo, a partir de la utilización

de un criterio externo (como es la entrevista) se pretende evaluar la validez predictiva de la herramienta informática. De esta manera, el presente trabajo se divide en dos estudios, cada uno de ellos encaminado a mostrar los resultados de las dos pruebas de validez utilizadas.

Método

Estudio I

Participantes

Al ser la finalidad del presente estudio analizar las propiedades psicométricas relativas a la validez de contenido de los ítems que componen el programa de simulación 3D My-School (MS), se recurrió a una muestra de expertos con experiencia específica en el ámbito de los comportamientos de consumo de drogas y acoso escolar. En total participaron cuatro jueces ($N = 4$), dos expertos en el campo del consumo de drogas ($n = 2$) y dos expertos en el campo de los comportamientos de acoso escolar ($n = 2$), todos ellos pertenecientes a centros de investigación o instituciones sanitarias de reconocido prestigio. Si atendemos a los participantes que colaboraron en este estudio ($N = 4$), un 50 % fueron hombres ($n = 2$), y un 50 % mujeres ($n = 2$). Las características sociodemográficas pueden consultarse en la Tabla 1. Con respecto a los ámbitos específicos de experiencia, dos de los expertos presentaban una amplia experiencia en el campo del consumo de drogas ($n = 2$) y los otros dos expertos en el campo de los comportamientos de acoso escolar ($n = 2$).

Tabla 1
Características sociodemográficas de la muestra

Características	Experto 1	Experto 2	Experto 3	Experto 4
Sexo	Mujer	Mujer	Hombre	Hombre
Área profesional	Psiquiatría	Psicología	Educación	Psicología
Nivel académico	Licenciado	Doctor	Doctor	Doctor
Campo de experiencia	Drogas	Acoso escolar	Acoso escolar	Drogas
Ámbito de trabajo	Aplicado	Aplicado	Investigación	Investigación
Tipo de centro de trabajo	Hospital	Hospital	Universidad	Universidad
Tiempo de desempeño	6 meses	20 meses	25 meses	13 meses

Instrumentos

Para la recogida de información de los juicios dados por los expertos, los autores del presente estudio diseñaron una escala *ad hoc*, con un puntaje de cero a cien para medir la pertinencia de los ítems en términos de relevancia, coherencia, claridad y tendenciosidad de las situaciones del programa MS para detectar comportamientos de consumo de drogas y acoso escolar. En base a lo sugerido por Osterlind (1989), se decidió no incluir un sistema de respuestas dicotómicas acuerdo/desacuerdo por no ser suficientemente adecuado para la evaluación del contenido de los ítems con respecto al acuerdo existente entre los evaluadores expertos.

De manera pormenorizada, la plantilla del protocolo de juicios de expertos utilizada incluyó una descripción tanto del propósito del juicio de expertos como del de la prueba objeto de evaluación, en donde se especifica claramente la finalidad para la que sería utilizada la información procedente de sus respuestas. A su vez, también se incluyó en dicha plantilla el objetivo general y los específicos del estudio que se estaba llevando a cabo, además de una conceptualización del universo del contenido referido a la definición de comportamientos de consumo de sustancias y acoso escolar. A continuación, se incluyó una tabla con información acerca de la operacionalización de las variables relevantes a tener en cuenta para el proceso de evaluación de los distintos dominios presentes en el programa MS. Con toda esta información se persiguió unificar los criterios utilizados por los expertos con la finalidad de instruir claramente al juez en la dimensión y el indicador que mide cada ítem o grupo de ellos, algo que es clave para establecer la correcta fiabilidad de los datos obtenidos mediante el juicio de los diferentes expertos implicados en este tipo de procedimientos (Escobar-Pérez & Cuervo-Martínez, 2008). Por último, se especificó una descripción tanto de los indicadores utilizados para la evaluación del programa MS, como del sistema de puntuaciones.

En concreto, si atendemos a las puntuaciones, estas estaban comprendidas según una escala de cero a 100, en donde cero se correspondía con no cumplir con el criterio o indicador, y 100 con cumplir con el criterio.

Procedimiento

Siguiendo la recomendación de Millman y Greene (1989), la elección de los expertos fue realizada con base en el propósito o finalidad para el cuál fue construido el programa informático MS, en donde se tuvo en cuenta, por tanto, que el grupo de expertos elegido representase suficientemente una diversidad de puntos de vista con respecto al contenido de las situaciones que componen dicho programa. Finalmente, los expertos seleccionados procedieron del ámbito de la psicología, psiquiatría y de la educación, asegurándose así la independencia entre los mismos. Por su parte, para la elección del número de expertos a incluir se siguió la recomendación, sugerida por Lynn (1986) (citado en McGartland, Berg-Weger, Tebb, Lee, & Rauch, 2003), de incluir un mínimo de tres jueces expertos. El tiempo total empleado por cada uno de los participantes en contestar a la plantilla estuvo comprendido entre los 30 y 40 minutos de duración total. Cada uno de los participantes respondió a dicha plantilla de manera individualizada.

Resultados

Con la finalidad de evaluar la validez de contenido de las situaciones del programa MS se recurrió al uso del coeficiente de concordancia *w* de Kendall (Kendall & Babington, 1939) de cara a evaluar el grado de acuerdo existente entre los evaluadores con respecto a la valoración del contenido de las diferentes situaciones del programa. Así, tal y como se muestra en la Tabla 2, se obtuvo para el total de las situaciones un índice *w* de Kendall de 0.71 ($\chi^2 = 315.264$; $p = 0$).

Tabla 2

Índice de concordancia entre las valoraciones de los jueces acerca del total de ítems del programa MS

<i>W</i> de Kendall	0.71
Chi-cuadrado de Pearson	315.264
Grados de libertad	111
Significación asintótica bilateral	0

Para cada uno de los ítems se obtuvo una ordenación de cuatro variables (relevancia, coherencia, claridad y tendenciosidad) según una escala ordinal con valores de cero a 100. En primer lugar, se muestran los estadísticos descriptivos obtenidos tras la aplicación del procedimiento no paramétrico referido a los rangos y las medias obtenidas por cuanto a la valoración por parte de los expertos del contenido de las situaciones relacionadas con acoso escolar (Tabla 3) y consumo de drogas (Tabla 4).

Tabla 3

Estadísticos descriptivos asignados a los ítems de acoso escolar

Ítem	Contenido	Variable	Rango promedio	Media	Desviación típica	Mínimo	Máximo
1	Victima de acoso en el patio	Relevancia	60.25	87.5	13.229	70	100
		Coherencia	75.13	86.25	24.281	50	100
		Claridad	64.75	81.25	27.801	40	100
		Tendenciosidad	17.25	13.75	13.769	0	30
3	Victima de acoso en el patio	Relevancia	87.88	96.25	4.787	90	100
		Coherencia	87.88	96.25	4.787	90	100
		Claridad	66.25	90	8.165	80	100
		Tendenciosidad	9.25	7.5	9.574	0	20
6	Comportamiento habitual durante el recreo	Relevancia	52.63	82.5	15.546	60	95
		Coherencia	61.25	87.5	9.574	80	100
		Claridad	71.38	91.25	8.539	80	100
		Tendenciosidad	11.25	8.75	8.539	0	20
7	Testigo de acoso en el patio	Relevancia	79.25	95	4.082	90	100
		Coherencia	84	95	5.774	90	100
		Claridad	71.38	91.25	8.539	80	100
		Tendenciosidad	11.5	10	14.142	0	30
14	Acosador de otro compañero	Relevancia	61.25	87.5	9.574	80	100
		Coherencia	70.63	90	8.165	80	100
		Claridad	72.75	92.5	5	90	100
		Tendenciosidad	14	11.25	13.15	0	30
16	Acoso relacional	Relevancia	49	85	7.071	80	95
		Coherencia	70.63	90	8.165	80	100
		Claridad	77.88	93.75	4.787	90	100
		Tendenciosidad	15.75	13.75	17.97	0	40
17.1	Creencia sobre acoso escolar	Relevancia	57.88	88.75	6.292	80	95
		Coherencia	71.63	91.25	8.539	80	100
		Claridad	67.38	87.5	18.93	60	100
		Tendenciosidad	18	15	17.321	0	40
17.2	Creencia sobre acoso escolar	Relevancia	63	90	7.071	80	95
		Coherencia	81	93.75	4.787	90	100
		Claridad	93.38	97.5	5	90	100
		Tendenciosidad	19.88	18.75	22.5	0	50
17.3	Creencia sobre acoso escolar	Relevancia	64.88	90	4.082	85	95
		Coherencia	73.5	91.25	6.292	85	100
		Claridad	93.38	97.5	5	90	100
		Tendenciosidad	14	11.25	13.15	0	30
17.4	Creencia sobre acoso escolar	Relevancia	67.25	91.25	2.5	90	95
		Coherencia	74.13	93.75	4.787	90	100
		Claridad	79.63	95	5.774	90	100
		Tendenciosidad	14	10	8.165	0	20
17.5	Creencia sobre acoso escolar	Relevancia	67.25	91.25	2.5	90	95
		Coherencia	81	93.75	4.787	90	100
		Claridad	71.38	91.25	8.539	80	100
		Tendenciosidad	16.25	12.5	12.583	0	30
17.6	Creencia sobre acoso escolar	Relevancia	53.25	85	12.247	70	95
		Coherencia	71.25	88.75	13.15	70	100
		Claridad	66.38	86.25	17.97	60	100
		Tendenciosidad	15.75	13.75	17.97	0	40
17.7	Creencia sobre acoso escolar	Relevancia	73.5	91.25	6.292	85	100
		Coherencia	73.5	91.25	6.292	85	100
		Claridad	84.75	96.25	4.787	90	100
		Tendenciosidad	16.63	16.25	22.867	0	50
17.8	Creencia sobre acoso escolar	Relevancia	70.63	90	8.165	80	100
		Coherencia	73.5	91.25	6.292	85	100
		Claridad	79.63	95	5.774	90	100
		Tendenciosidad	19.88	20	27.08	0	60
17.9	Creencia sobre acoso escolar	Relevancia	64.13	87.5	9.574	80	100
		Coherencia	67	88.75	8.539	80	100
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	15.13	17.5	28.723	0	60
17.10	Creencia sobre acoso escolar	Relevancia	59.75	88.75	2.5	85	90
		Coherencia	68.5	91.25	2.5	90	95
		Claridad	93.38	97.5	5	90	100
		Tendenciosidad	17.63	18.75	27.801	0	60

Tabla 4
Estadísticos descriptivos asignados a los ítems de consumo de drogas

Ítem	Contenido	Variable	Rango promedio	Media	Desviación típica	Mínimo	Máximo
2	Consumo de tabaco	Relevancia	58	85	17.321	60	100
		Coherencia	74.63	92.5	9.574	80	100
		Claridad	66.38	88.75	13.15	70	100
		Tendenciosidad	9.25	7.5	9.574	0	20
4	Consumo de alcohol	Relevancia	66	91.25	8.539	80	100
		Coherencia	84	95	5.774	90	100
		Claridad	71.38	91.25	8.539	80	100
		Tendenciosidad	12.5	10	11.547	0	20
5	Comportamiento habitual en clase	Relevancia	50.38	86.25	4.787	80	90
		Coherencia	58.38	87.5	6.455	80	95
		Claridad	66.25	90	8.165	80	100
		Tendenciosidad	18	15	17.321	0	40
8	Consumo de marihuana	Relevancia	84	95	5.774	90	100
		Coherencia	77.13	92.5	5	90	100
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	9.25	7.5	9.574	0	20
9	Consumo de cocaína	Relevancia	76	90	14.142	70	100
		Coherencia	77.5	92.5	9.574	80	100
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	9.25	7.5	9.574	0	20
10	Relación con el padre	Relevancia	49.38	83.75	11.087	70	95
		Coherencia	65.63	85	17.321	60	100
		Claridad	80	92.5	9.574	80	100
		Tendenciosidad	14	10	8.165	0	20
11	Relación con la madre	Relevancia	49.38	83.75	11.087	70	95
		Coherencia	65.63	85	17.321	60	100
		Claridad	76.75	90	14.142	70	100
		Tendenciosidad	14	10	8.165	0	20
12	Relación con ambos padres	Relevancia	53.63	85	10	70	90
		Coherencia	67.38	87.5	12.583	70	100
		Claridad	63	87.5	12.583	70	100
		Tendenciosidad	20.13	17.5	15	0	30
13	Consumo de drogas en grupo	Relevancia	44	81.25	10.308	70	95
		Coherencia	53.5	83.75	11.087	75	100
		Claridad	64.75	81.25	27.801	40	100
		Tendenciosidad	21.38	30	38.297	0	80
15	Consumo de MDMA	Relevancia	51.25	75	30.277	30	95
		Coherencia	73.5	91.25	6.292	85	100
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	14	11.25	13.15	0	30
17.11	Creencia sobre consumo de drogas	Relevancia	64.13	87.5	9.574	80	100
		Coherencia	64.13	87.5	9.574	80	100
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	11.5	10	14.142	0	30
17.12	Creencia sobre consumo de drogas	Relevancia	65.88	90	7.071	80	95
		Coherencia	65.88	90	7.071	80	95
		Claridad	86.5	95	5.774	90	100
		Tendenciosidad	11.5	10	14.142	0	30

Por cuanto a la dirección de las valoraciones, si nos centramos en cada uno de los valores observados para cada ítem específico, podemos comprobar que la media de las puntuaciones otorgadas por los jueces, referidos a las variables relevancia, coherencia y claridad del contenido de los ítems es alta en todas las variables (véanse tablas 3 y 4). Por su parte, la variable tendenciosidad del contenido de los ítems mostraría, por el contrario, valores bajos de media en todos los ítems analizados.

A su vez, se realizó un análisis individual por separado de cada uno de los ítems, con la finalidad de analizar el grado de concordancia específico existente entre las valoraciones de los expertos con respecto a cada una de las variables analizadas. Para ello, se utilizó de nuevo el estadístico *w* de Kendall (véase Tabla 5).

Tabla 5
Grado de concordancia existente entre los jueces expertos para cada una de las situaciones del programa MS

Ítem	Contenido	<i>w</i> de Kendall	χ^2	<i>g</i> ^l	Sig. asintótica
1	Víctima de acoso escolar en el patio	0.691	8.289	3	0.04
2	Consumo de tabaco	0.737	8.842	3	0.031
3	Víctima de acoso escolar en clase	0.75	9	3	0.029
4	Consumo de alcohol	0.729	8.75	3	0.033
5	Comportamiento habitual en clase	0.688	8.25	3	0.041
6	Comportamiento habitual en el patio	0.651	7.816	3	0.05
7	Testigo de acoso	0.713	8.559	3	0.036
8	Consumo de marihuana	0.825	9.9	3	0.019
9	Consumo de cocaína	0.782	9.387	3	0.025
10	Relación con el padre	0.858	10.297	3	0.016
11	Relación con la madre	0.813	9.75	3	0.021
12	Relación con ambos padres	0.825	9.9	3	0.019
13	Consumo de drogas en grupo	0.25	3	3	0.392
14	Acosador de otro compañero	0.75	9	3	0.029
15	Consumo de MDMA	0.886	10.636	3	0.014
16	Acoso relacional (exclusión social)	0.838	10.054	3	0.018
17.1	Creencia sobre acoso escolar	0.688	8.25	3	0.041
17.2	Creencia sobre acoso escolar	0.917	11	3	0.012
17.3	Creencia sobre acoso escolar	0.886	10.636	3	0.014
17.4	Creencia sobre acoso escolar	0.875	10.5	3	0.015
17.5	Creencia sobre acoso escolar	0.735	8.824	3	0.032
17.6	Creencia sobre acoso escolar	0.729	8.75	3	0.033
17.7	Creencia sobre acoso escolar	0.75	9	3	0.029
17.8	Creencia sobre acoso escolar	0.735	8.824	3	0.032
17.9	Creencia sobre acoso escolar	0.879	10.548	3	0.014
17.10	Creencia sobre acoso escolar	0.907	10.886	3	0.012
17.11	Creencia sobre consumo de drogas	0.9	10.8	3	0.013
17.12	Creencia sobre consumo de drogas	0.932	11.182	3	0.011

Pues bien, si atendemos a los resultados obtenidos para cada uno de los ítems, la prueba mostró la existencia de relaciones estadísticamente significativas ($p < 0.05$) entre las valoraciones de los expertos para todos los ítems analizados, con la excepción de un solo caso (ítem 13 referido al consumo de drogas con el grupo de amigos), para el cual la prueba no mostró valores significativos. En conclusión, estos resultados indicarían la existencia de concordancia entre los jueces expertos con respecto a la valoración de cada una de las situaciones cuando estas son analizadas de manera específica o individualizada. Así, los índices *w* de Kendall se encontrarían en un intervalo comprendido entre 0.651 y 0.932.

Estudio II

Participantes

Con la finalidad de calcular la validez predictiva de los datos obtenidos mediante el programa My-School (MS), se utilizó una muestra de estudiantes de Educación Secundaria Obligatoria (ESO) y Bachillerato, procedente de

dos centros escolares elegidos aleatoriamente, situados en la provincia de Almería (España). Como medida de evaluación de la validez predictiva del programa MS se recurrió a la entrevista en profundidad como criterio externo para confirmar las respuestas dadas seis meses antes en el programa MS por una submuestra procedente de 570 participantes, que respondieron al programa con anterioridad. La selección de la muestra se realizó por estratos con asignación óptima y sin reemplazo. Los grupos contemplados se correspondieron con cada una de las variables independientes del estudio, a saber: ser acosado/acosador, consumidor de tabaco, alcohol, marihuana, cocaína, MDMA o éxtasis y no ser consumidores de ninguna sustancia. Bajo estos criterios se conformó una submuestra compuesta por un total de 58 participantes, los cuales correspondieron al 10 % del total de los 570 participantes que contestaron previamente al programa MS. Un resumen de las características sociodemográficas puede consultarse en la Tabla 6.

Tabla 6

Descripción de las características sociodemográficas de la muestra

	N	Rango de edad	Edad media	Hombres	Mujeres
Entrevista en profundidad	58	11-19	14.7 (DE = 2.388)	47 %	53 %

Instrumentos

A continuación, se presentan los instrumentos utilizados en este estudio.

Entrevista semiestructurada. Se realizó una entrevista semiestructurada, *ex profeso*, para la presente investigación, con la finalidad de obtener -o no- la corroboración de los resultados obtenidos mediante el programa de simulación 3D MS. La entrevista estuvo compuesta por cuatro partes diferenciadas en función del contenido al que hacía mención. Una primera parte la formaban preguntas con contenidos neutros genéricos no relacionados con el consumo de drogas ni el acoso escolar. Una segunda parte estuvo compuesta por cuestiones relacionadas con acoso escolar, como es el

ser víctima, acosador o testigo de este tipo de situaciones. En tercer lugar se presenta un conjunto de cuestiones relacionadas con la historia de consumo de sustancias del alumno y del grupo de amigos. La cuarta parte de la entrevista hace referencia a cuestiones relacionadas con el ambiente familiar y cuestiones de personalidad.

Programa My-School (MS). El programa de simulación MS utiliza entornos virtuales tridimensionales para recrear determinados contextos de consumo de drogas y acoso escolar con la finalidad de detectar a modo de herramienta tipo *screening* estos tipos de comportamientos de riesgo. Para ello, el programa MS consta de un total de 17 escenas de simulación, las cuales recrean, mediante la utilización de gráficos tridimensionales (al estilo de los actuales videojuegos), diferentes situaciones de consumo de sustancias y acoso escolar con personajes que interactúan entre sí en situaciones conflictivas y ante las cuales el participante ha de elegir cómo comportarse, dadas las diferentes opciones de respuesta que se le ofrecen. Para una descripción detallada del programa, el lector puede remitirse a Carmona et al. (2010a, 2010b, 2011).

Procedimiento

Una vez seleccionada la submuestra compuesta por 58 alumnos, con la finalidad de no comprometer su participación, así como con la intención de no alterar o influir en la sinceridad de las respuestas de los participantes que fueron entrevistados, se les informó, tanto a los propios participantes de las entrevistas como al resto de compañeros no entrevistados, que se trataba de “un estudio dirigido a recabar información sobre ellos con la finalidad de conocer mejor a los jóvenes de estas edades”, informándoles acerca de la voluntariedad y confidencialidad de sus respuestas. Cada una de las entrevistas realizadas fue llevada a cabo de manera individual por un total de dos entrevistadores con experiencia previa tanto en la evaluación como en el tratamiento de estos tipos de comportamientos.

Las entrevistas se realizaron en el interior de despachos privados habilitados para tal fin por los responsables académicos del centro. Los investigadores eran ciegos a las respuestas emitidas previamente por los participantes. Las entrevistas tuvieron una duración aproximada de 20 minutos por alumno, dependiendo dicha duración de la propia dinámica de cada una de las entrevistas realizadas. Previo a la realización de las entrevistas, se les informó acerca del anonimato, la voluntariedad y la confidencialidad de sus respuestas.

Resultados

En primer lugar, se analizó la posible existencia de respuestas congruentes entre los participantes con respecto a la información obtenida a través de las entrevistas y las respuestas dadas seis meses antes en el programa MS, para lo cual se conformaron tablas de contingencia dicotómicas con el total de congruencias observadas. A continuación, y siguiendo con los análisis estadísticos realizados, se utilizó la prueba de McNemar para muestras dependientes (McNemar, 1947), como medida para identificar la posible existencia de diferencias estadísticamente significativas entre las respuestas emitidas en MS y la entrevista. Al ser la finalidad de dicho análisis la posible corroboración de los datos obtenidos mediante ambas mediciones.

En concreto, con respecto a los resultados obtenidos, la comparación realizada mediante la prueba estadística McNemar para muestras dependientes reveló que no existen diferencias estadísticamente significativas entre las respuestas obtenidas mediante el uso del programa de simulación MS y las obtenidas mediante las entrevistas individualizadas ($\chi^2 = 3.2$; $p > 0.05$). Así, la inexistencia de diferencias significativas entre los datos procedentes de ambos tipos de medidas distintas nos indicaría que existiría una correspondencia significativa (validez de criterio predictiva) entre las respuestas que los participantes dieron a través del programa MS y las respuestas dadas por

esos mismos participantes en la vida real medido mediante el uso de la entrevista. Un resumen de dichos resultados puede consultarse en la Tabla 7.

Tabla 7

Correspondencias entre las respuestas del programa informático MS y las respuestas de las entrevistas

Consumo de sustancias	My-School Participantes (n)	Entrevista	
		Correspondencia con My-School	
		Sí	No
Tabaco	7	6	1
Alcohol	11	10	1
Marihuana	9	8	1
MDMA	2	1	1
Cocaína	4	3	1
No consumidor	8	8	0
Subtotal drogas	41	36	5
Acoso escolar			
Acosado	6	6	0
Acosador	4	4	0
No acosado	7	7	0
Subtotal acoso	17	17	0
Total participantes	58	53	5

Tal y como puede observarse en la Tabla 7, con respecto a los participantes relacionados con alguna de las condiciones de acoso escolar, se observa que existe correspondencia entre MS y las entrevistas en el total de alumnos que participaron ($n = 17$), en donde no existe, por tanto, ningún participante para el que no se corroborase en las entrevistas, seis meses después, sus respuestas del programa MS ($n = 0$). Por su parte, con respecto al consumo de sustancias, para cada una de las evaluadas (tabaco, alcohol, marihuana, cocaína y MDMA) existiría un alumno para el que no se habría encontrado que exista correspondencia entre sus respuestas en MS con respecto a esas sustancias en concreto y las respuestas dadas en la entrevista ($n = 5$). No obstante, para el resto de participantes ($n = 36$) sí existiría correspondencia entre sus respuestas de consumo de sustancias obtenidas mediante el programa MS y las registradas en la entrevista seis meses después.

Si atendemos a los porcentajes de alumnos que en sus respuestas se habrían mostrado coincidentes en ambas pruebas, el 100 % de los de ellos, bien como implicados en situaciones de acoso escolar o bien como no acosados, habrían mostrado respuestas similares tanto en el programa MS como en las entrevistas. El porcentaje de alumnos que se habrían mostrado

incongruentes con sus respuestas en ambas medidas fue 0. Con respecto al consumo de sustancias, un 87.8 % de los alumnos habría mostrado respuestas coincidentes en ambas medidas de evaluación. Sin embargo, el 12.2 % de alumnos restante no habría mostrado dicha coincidencia entre la entrevista y sus respuestas de consumo de drogas en el programa MS.

En definitiva, y si tenemos en cuenta los datos de manera global, los análisis muestran la inexistencia de diferencias estadísticamente significativas entre las respuestas mostradas por el programa de simulación 3D MS y las entrevistas, corroborándose así la correspondencia entre las respuestas de consumo de sustancias y acoso escolar obtenidas mediante el programa MS y las respuestas dadas por esos mismos participantes en las entrevistas seis meses después.

Discusión

El presente artículo ha buscado analizar las propiedades psicométricas referidas a la validez del programa de simulación 3D My-School (MS) para detectar alumnos consumidores de sustancias o implicados en acoso escolar. Para ello, en primer lugar, se llevó a cabo el análisis de la validez del contenido de las situaciones que componen el programa MS, utilizando como criterio externo jueces expertos en el campo de los comportamientos de consumo de drogas y acoso escolar. En segundo lugar, se analizó la validez de criterio predictiva de dicha herramienta, para lo cual, como medida o criterio externo, se procedió a la realización de entrevistas individualizadas, las cuales fueron realizadas seis meses después de la aplicación del programa MS en esa misma muestra de participantes.

Pues bien, si atendemos al primer estudio dirigido al análisis de la validez del contenido de las situaciones del programa, la obtención de un índice w de Kendall de 0.71 para el total de ítems o situaciones del programa MS ($p = 0$) nos indicaría la existencia de una concordancia media-alta o acuerdo sustancial entre los cuatro jueces expertos con respecto a las variables

estudiadas (Landis & Koch, 1977). Es relevante destacar la pertinencia de analizar la dirección o el significado de las valoraciones realizadas por los expertos respecto a las variables relevancia, coherencia, claridad y tendenciosidad para cada uno de los ítems o situaciones individualmente. A este respecto, y tal y como señalan Siegel y Castellan (1995), es importante analizar si la concordancia entre los jueces, medida mediante el estadístico w de Kendall, responde a que estén de acuerdo acerca de la idoneidad del ítem que estén clasificando o, por el contrario, exista concordancia con respecto a la baja idoneidad del ítem en particular en base a las variables utilizadas para su clasificación. En este sentido, tal y como puede observarse en las Tablas 3 y 4, los resultados mostraron valores de media altos para las tres variables relevancia, coherencia y claridad, y bajos para la variable tendenciosidad. Es decir, existiría entre los jueces un acuerdo sustancial a la hora de valorar el contenido de las diferentes situaciones del programa MS como relevantes, coherentes y claros con respecto a los dominios a los que se refieren, esto es, consumo de drogas o acoso escolar. A su vez, los valores bajos en las medias de las valoraciones referentes a la variable tendenciosidad nos indicaría la existencia de acuerdo o concordancia entre los jueces por cuanto a la baja tendenciosidad que mostraría el contenido de las situaciones del programa MS.

Mención aparte merece el análisis de la concordancia realizado a cada uno de los ítems de manera individual. En este sentido, el análisis individualizado reveló valores w de Kendall significativos para todas las situaciones del programa MS, con la excepción de la situación 13 (véase Tabla 5). De manera específica, el hecho de que los índices w de Kendall se encuentren en un intervalo comprendido entre 0.651 y 0.932 implicaría un acuerdo sustancial o casi perfecto (Landis & Koch, 1977) entre los jueces, respecto a la validez del contenido de cada una de las situaciones para la detección de comportamientos de acoso escolar y consumo de drogas.

Por su parte, con respecto a los resultados del estudio II, relativos a la validez de

criterio predictiva del programa MS para detectar alumnos consumidores de sustancias o implicados en acoso escolar, los datos obtenidos mediante el estadístico McNemar muestran que no existen diferencias significativas entre las respuestas obtenidas mediante el uso del programa de simulación MS y las obtenidas mediante las entrevistas individualizadas. Existe, por tanto, una correspondencia entre lo que los alumnos dicen cuando son entrevistados con respecto a su consumo de drogas y su implicación en situaciones de acoso escolar, y lo que respondieron seis meses antes en el programa informático MS cuando se vieron inmersos en ese mismo tipo de situaciones simuladas virtualmente.

Con respecto a la elección de la muestra específica seleccionada para las entrevistas, cabe puntualizar que dado que el número de acosadores y acosados, así como de consumidores de determinadas sustancias, como cocaína o MDMA, fue minoritario, lo que coincide con las prevalencias observadas para estas edades (Observatorio Español Sobre Drogas, 2008; Serrano & Iborra, 2005), la muestra final utilizada fue seleccionada para incluir a aquel número reducido de alumnos que mostraban estos tipos de comportamientos, por lo que esta no es proporcional al total de la población original de participantes que respondieron seis meses antes al programa MS, sino representativa de aquellos implicados en situaciones de acoso escolar y consumo de drogas, así como de no consumidores y no acosados/acosadores, en donde la muestra final utilizada corresponde al 10 % del total de la muestra inicial de 570 alumnos. De esta manera, fue posible analizar y validar dichas situaciones del programa informático MS en base a la muestra utilizada.

A modo de conclusión, con respecto al estudio I, podemos afirmar que la importancia de los resultados obtenidos con base al acuerdo sustancial observado entre los jueces #con respecto a la validez del contenido de las situaciones que componen el programa MS# se presentaría como un aval empírico que fundamentaría el uso de dicho programa para la detección de estos tipos de comportamientos

en jóvenes. En segundo lugar, con respecto al estudio II, el uso de entrevistas individualizadas habría permitido corroborar la validez predictiva de la herramienta MS para la detección de comportamientos de acoso escolar y consumo de drogas en jóvenes estudiantes. En definitiva, los resultados obtenidos se presentarían como un aval empírico de la validez de contenido y de criterio, predictiva del programa de simulación 3D MS para detectar alumnos en riesgo de consumo de sustancias y acoso escolar.

Referencias

- Bates, S. C., & Cox, J. M. (2008). The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviours. *Computers in Human Behavior*, 24(3), 903-916. <https://doi.org/10.1016/j.chb.2007.02.021>
- Buchanan, T., Johnson, J.A., & Goldberg, L.R. (2005). Implementing a five-factor personality inventory for use on the Internet. *European Journal of Psychological Assessment*, 21(2), 115-127. <https://doi.org/10.1027/1015-5759.21.2.115>
- Carmona, J. A., Cangas, A. J., & Langer, A. I. (2012). Applications of 3D simulation in Mental Health: Utilities and new developments. En L. L'Abate (Ed.), *Mental Illnesses. Evaluation, Treatments and Implications* (pp. 37-56). Rijeka: Intech.
- Carmona, J. A., Cangas, A. J., García, G. R., Langer, A. I., & Zárate, R. (2012). Early Detection of Drug Use and Bullying in Secondary School Children Using a 3-D Simulation Program *CyberPsychology, Behavior, and Social Networking*, 15(1), 43-49. <https://doi.org/10.1089/cyber.2010.0589>
- Carmona, J. A., Espínola, M., Cangas, A. J., & Iribarne, L. (2010a). MII School: New 3D Technologies Applied in Education to Detect Drug Abuses and Bullying in Adolescents. En M. Lytras et al. (Eds.), *Technology Enhanced Learning: Quality of*

- Teaching and Educational Reform* (pp. 65-72). Heidelberg: Springer.
- Carmona, J. A., Espínola, M., Cangas, A. J., & Iribarne, L. (2010b). Detecting Drug Use in Adolescents Using a 3D Simulation Program. *Psychology, Society, & Education*, 2(2), 143-153. Recuperado de <http://ojs.ua.es/ojs/index.php/psyc>
- Carmona, J. A., Espínola, M., Cangas, A. J., & Iribarne, L. (2011). MII-School: A 3D Videogame for the Early Detection of Abuse of Substances, Bullying and Mental Disorders in Adolescents. *European Journal of Education and Psychology*, 4(1), 75-85. Recuperado de <http://www.formacionasuni.vep.com/ejep/>
- Escobar-Pérez, J., & Cuervo-Martínez, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6(1), 27-36.
- Freitas, S., & Neumann, T. (2009). The use of "exploratory learning" for supporting immersive learning in virtual environments. *Computers & Education*, 52(2), 343-352. <https://doi.org/10.1016/j.compedu.2008.09.010>
- Gutiérrez-Maldonado, J., Alsina-Jurnet, I., Carvallo-Becú, C., Letosa-Porta, A., & Magallón-Neri, E. (2007). Aplicaciones clínicas de la realidad virtual en el ámbito escolar. *Cuadernos de medicina psicosomática y psiquiatría de enlace*, 82, 32-51. Recuperado de http://www.editorialmedica.com/editorialmedica_publicacion.php
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from Webbased personality inventories. *Journal of Research in Personality*, 39(1), 103-129. <https://doi.org/10.1016/j.jrp.2004.09.009>
- Kendall, M. G., & Babington, B. (1939). The Problem of m Rankings. *The Annals of Mathematical Statistics*, 10(3), 275-287. <https://doi.org/10.1214/aoms/1177732186>
- Kim, B., Park, H., & Baek, Y. (2009). Not just fun, but serious strategies: using meta-cognitive strategies in game-based learning. *Computers & Education*, 52(4), 800-810. <https://doi.org/10.1016/j.compedu.2008.12.004>
- Kyriakakis, C., Van Rooyen, A., & Sisemore, D. (2000). The virtual classroom: a virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology & Behavior*, 3(3), 483-499. Recuperado de <http://www.liebertpub.com/overview/cyberpsychology-behavior-and-social-networking/10/>
- Landis J. R., & Koch G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. Recuperado de <http://www.biometrics.tibs.org/>
- López Feal, R. (1986). *Construcción de instrumentos de medida en ciencias conductuales y sociales*. Barcelona: Alamex.
- McCabe, S. E., Boyd, C. J., Young, A., Crawford, S., & Pope, D. (2005). Mode effects for collecting alcohol and tobacco data among 3rd and 4th grade students: A randomized pilot study of Web-form versus paper-form surveys. *Addictive Behaviors*, 30(4), 663-671. <https://doi.org/10.1016/j.addbeh.2004.08.012>
- McGartland, D., Berg-Weger, M., Tebb, S. S., Suzanne, E., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94-104. Recuperado de <http://www.naswpress.org/publications/journals/swr.html>
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153-157.
- Millman, J., & Greene, J. (1989). The specification and development of test of achievement and ability. En R. L. Linn (Ed.), *Educational Measurement* (pp. 335-366). Londres: Macmillan.
- Observatorio Español sobre Drogas. (2008). *Encuesta Estatal Sobre Uso de Drogas en Enseñanzas Secundarias (ESTUDES)*. Madrid: Delegación del Gobierno para el Plan Nacional sobre Drogas.

- Osterlind, S. J. (1989). *Constructing test items*. Boston: Kluwer Academic Publishers.
- Reips, U. D. (2000). The web experiment method: Advantages, disadvantages, and solutions. En M.H. Birnbuam (Ed.), *Psychological experiments on the internet* (pp. 89-117). San Diego: Academic Press.
- Serrano, A., & Iborra, I. (2005). *Informe Violencia entre compañeros en la escuela*. Valencia, España: Centro Reina Sofía para el Estudio de la Violencia. Recuperado de <http://www.centroreinasofia.es>
- Siegel, S., & Castellan, N. J. (1995). *Estadística no paramétrica aplicada a las ciencias de la conducta*. México D. F.: Trillas.
- Zoll, C., Enz, S., Schaub, H., Aylett, R., & Paiva, A. (2006, abril). Fighting bullying with the help of autonomous agents in a virtual school environment. *The 7th International Conference on Cognitive Modeling*. Trieste, Italia. Recuperado de <http://www.inesc-id.pt/ficheiros/publicacoes/3057.pdf>

Notas

- * Artículo de investigación.