

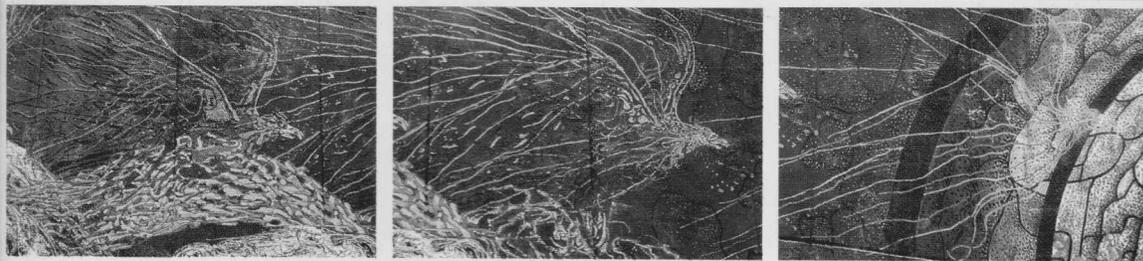
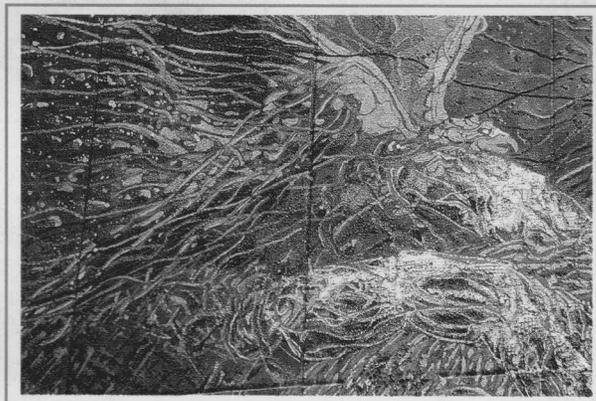


UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO  
Campus Iztacala  
División de Investigación



**Cuadernos de Investigación  
Interdisciplinaria en Ciencias de la Salud,  
la Educación y el Ambiente**

**VOLUMEN I**



ISBN 968-36-6403-2



9 789683 664037

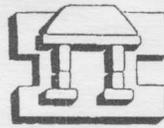
Foto: Peter Mueller

Fragmento del mural de OrtízGris. Biblioteca de la ENEPI  
UNAM Campus Iztacala



UNAM

Dr. Francisco Barnés de Castro  
Rector



IZTACALA

Dr. Felipe Tirado Segura  
Director

Lic. Guadalupe Hernández Cortés  
Secretaria de la Dirección

M.C. Ramiro Jesús Sandoval  
Secretario Académico

Dra. Gloria Vilaclara Fatjó  
División de Investigación

M. en C. Juan Manuel Mancilla Díaz  
Departamento de Investigación

M.C. José Jaime Ávila Valdivieso  
Sección Editorial

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
*campus Iztacala*  
DIVISIÓN DE INVESTIGACIÓN

**Cuadernos de investigación interdisciplinaria en  
ciencias de la salud, la educación y el ambiente**

VOLUMEN I

Compiladores  
Juan Manuel Mancilla Díaz  
Gloria Vilaclara Fatjó

José Jaime Ávila Valdivieso  
Corrección de Estilo

MÉXICO 1998

## La ejecución de estudiantes de educación básica en dos pruebas psicométricas: Comparación entre las versiones lápiz-papel y computarizada

*Alfonso Bustos Sánchez, Xóchilt Galicia Moyeda; Alejandro Miranda Díaz y Felipe Tirado Segura*  
UNAM campus Iztacala

**Resumen:** Se investigaron los efectos de interacción (sujeto-máquina) al comparar la ejecución en dos pruebas, versión tradicional y computarizada, con diferentes niveles de exigencia de interacción. Se seleccionaron el Test de Matrices Progresivas de RAVEN (Menor Interacción) y la subprueba de Ordenación de Dibujos de la Escala de Inteligencia de Wechsler (Mayor Interacción), y se crearon las versiones computarizadas para cada una de estas pruebas. Se analizaron las ejecuciones de niños de primaria (104) de escuelas pública y privada (57 y 47 participantes respectivamente), de 3° (43) y 6° (61) grado. Los resultados indican que no hay diferencias significativas entre las versiones tradicional y computarizada en el RAVEN; sin embargo, en la subprueba de Ordenación de dibujos sí se observaron diferencias significativas, tanto en las escuelas públicas como en las privadas; dichas diferencias pueden ser atribuidas a la calidad de la interacción que establecen los niños con la computadora. **Palabras Clave:** Psicometría, RAVEN, WISC, pruebas computarizadas.

Los usos de la computadora en la educación han sido de gran diversidad, entre ellos se puede destacar la evaluación computarizada y el desarrollo de programas educativos, los cuales han sido empleados para impulsar el aprendizaje asistido por computadora y como un medio de aprendizaje experimental; también han ayudado a fortalecer el logro académico al utilizarlos como tutoriales personalizados, e incluso para investigar procesos cognitivos, o motivacionales (Muth, 1984; Johnson y cols, 1985; Carey, 1986; Rambally y Rambally, 1987; Lepper y Gurtner, 1989; Béjar, 1991; Tirado, 1994).

Probablemente la psicometría computarizada será en corto plazo una de las áreas de mayor desarrollo en el campo de las aplicaciones de la computación. Esto por varias razones, entre ellas las históricas, ya que la psicología ha tenido un gran desarrollo en el área de evaluación y ha llegado a generar grandes aportaciones en el campo de la estadística, desde 1905, en que se empezó a utilizar la primera prueba de inteligencia desarrollada por Alfred Binet y Théodore Simon.

Otra característica relevante de los test psicológicos ha sido su amplia diversidad. Se utilizan hoy día para evaluar inteligencia o capacidad intelectual de un sujeto, al estimar sus habilidades personales para interactuar de manera razonada y efectiva ante diversas circunstancias, referidas a una norma. Están también las pruebas de logro escolar que se

usan con fines de diagnóstico y pronóstico; los test de aptitudes que permiten predecir -la ejecución futura- en determinadas tareas de personas que no han sido específicamente entrenadas; los inventarios de intereses, en los que el sujeto indica sus preferencias personales ante diferentes situaciones o actividades; las pruebas de personalidad en las que se identifican actitudes, sentimientos y comportamientos que permiten englobar estilos de personalidad; las pruebas proyectivas, en la que los sujetos elaboran un dibujo o hacen una lectura de un estímulo indefinido, atribuyéndole atributos, o bien forman asociaciones de palabras o complementan oraciones. En todos estos procedimientos se asume que se podrán ver reflejados los estados anímicos y conflictivos de las personas.

Pero además de la diversidad y la ya larga experiencia histórica de la psicometría, la computadora parece ahora agregar nuevas ventajas que prometen ser muy significativas, como un instrumento no sólo de medición sino también de aprendizaje. Entre estas ventajas se pueden citar: a) permite una interacción dinámica; b) las consecuencias de las respuestas pueden variar en función del tipo de contestación; c) se puede obtener un aprendizaje individualizado, diseñado a la medida del sujeto o estudiante, ramificando las rutas de acuerdo a su ejecución (Lepper, y Gurtner, 1989); d) es posible relacionar la ejecución con consecuencias que favorecen la motivación del participante (Lepper, y Gurtner, 1989); dadas sus propiedades gráficas se pueden mejorar los niveles de atención, con lo que se disminuye el nivel de errores (Backhoff y cols., 1996; Lepper, 1985); f) provee al practicante con una retroalimentación inmediata acorde con sus respuestas (Lepper, 1985; McCullough y cols., 1986.); g) evalúa de manera individualizada, automática e inmediata, el desempeño del estudiante (Lepper y Gurtner, 1989), h) emula de manera similar a un jurado calificador, el juicio que se emite respecto a los sujetos evaluados (Béjar, 1991); i) posibilita llevar el registro de todas las respuestas emitidas en una base de datos; j) se puede registrar y evaluar el tiempo de reacción a niveles de fracciones de segundo; k) permite procesar de manera inmediata los análisis estadísticos para calificar con referencia a la norma; l) presenta de manera gráfica los resultados, facilitando de manera significativa su lectura e interpretación.

Todas estas ventajas han entusiasmado y motivado a muchos investigadores a utilizar la computadora en diferentes áreas de la evaluación. Hay quienes la utilizan como una herramienta para analizar y concentrar los datos obtenidos en la aplicación de tests psicológicos, y así facilitar el establecimiento de normas estadísticas en poblaciones determinadas (Backhoff, 1993). Pero también se han aprovechado las cualidades de la computadora para realizar, mediante ellas, evaluaciones psicológicas, de tal manera que se investigue la confiabilidad y validez de las pruebas computarizadas (McCullough y cols., 1986; Fekken y Holden, 1989; Holden y Fekken, 1990). Por ejemplo, en los usos de la computadora dentro de la evaluación psicométrica, se han privilegiado las características

siguientes: La creación de bases de datos, entrevistas y pruebas computarizadas, rigurosamente normalizadas —ello es, hoy día, de gran ayuda para el especialista, ya que con ellas se pueden explorar una amplia variedad de áreas psicológicas—, pueden ser empleadas por el personal no especializado, se establece una lectura y evaluación estandarizada de la prueba simultánea a la aplicación, permite un mayor confort para responder preguntas confidenciales, posibilitan incluir medidas individualizadas y generalizadas, logra elaborar buenos niveles de diagnóstico, pueden emplearse para realizar investigaciones sistemáticas, reducen el tiempo que los evaluados y evaluadores requieren en comparación con los métodos convencionales, además de ser relativamente más baratas y fáciles de usar comparado con los medios tradicionales. Dentro de estas ventajas se ha incluido el supuesto de que no se requiere haber tenido alguna experiencia previa de interacción con computadoras; pero esto se ha presentado en aquellos casos en que, por el tipo de prueba y sus reactivos, los evaluados simplemente reciben instrucciones para oprimir una tecla, pues la prueba tiene la característica de que cada reactivo se contesta con un valor numérico, y esa misma respuesta es la que produce la presentación de un nuevo reactivo (Backhoff y cols. 1996; McCullough y cols., 1986).

Sin embargo, no se han realizado análisis precisos de las exigencias respecto a la interacción sujeto-máquina y el nivel de conocimiento de cómputo requerido para ejecutar adecuadamente una prueba, así como tampoco de las características de esta última para facilitar la relación interfaz-usuario. Consideramos que muchos de los instrumentos utilizados en el área de evaluación que sean llevadas a la computadora, seguramente requerirán niveles de interacción usuario-computadora mucho más complejos y que podrían variar la valoración que se haga del sujeto.

Estas consideraciones no han sido tomadas en cuenta en los estudios en los cuales ha surgido el interés por comparar las ejecuciones que se realizan en pruebas administradas por instrumentos en su versión tradicional (lápiz-papel), con la ejecución de la misma prueba en una versión computarizada.

Al respecto existen algunas evidencias que deben ser tomadas en consideración. En el estudio de Fekken y Holden (1989), bajo un paradigma de reaplicación del *Personality Research Form*, se evaluaron la validez y confiabilidad entre las versiones computarizadas y lápiz-papel en sujetos universitarios. Se encontró que las versiones son paralelas pues, las medias escalares y las desviaciones estándar en la ejecución computarizada fueron similares a las normas estandarizadas.

Una característica de la versión computarizada del PRF era que sólo se requería usar las teclas "T" (*true*) y "F" (*false*) para responder. Al oprimir cualquiera de éstas, la pantalla automáticamente borraba el reactivo en turno y presentaba el siguiente. Para que el sujeto realizara adecuadamente esta tarea, se le dieron instrucciones acerca de las

respuestas que tenía que realizar en la máquina (apretar las teclas "F" o "T"). Cualquier otra respuesta quedaba sin efecto, por lo que la diversidad de respuestas resulta ser muy reducida. El tener que apretar sólo una u otra tecla es un indicador de la baja exigencia de interacción con la computadora.

En otro trabajo, Holden y Fekken (1990) nuevamente encuentran ejecuciones semejantes entre la aplicación lápiz-papel del *Basic Personality Inventory* (BPI) y la aplicación computarizada. Se trabajó con pacientes adultos psiquiátricos, con un nivel de lectura de 6° grado. Cabe mencionar que este estudio tuvo el mismo arreglo metodológico, reaplicación del inventario y el mismo tipo de respuesta ante los reactivos: usar exclusivamente las teclas "T" (*true*), "F" (*false*) y "R" (*return*), donde "R" permitía a los participantes regresar a la pregunta anterior y hacer las correcciones deseadas. De manera similar, se puede advertir que las respuestas que el sujeto debe dar ante la computadora requieren de una interacción muy limitada, lo que suele ser una constante en este tipo de estudios.

Notemos que el nivel de interacción con la máquina, requerido en este tipo de aplicaciones computarizadas, tiende a ser el mínimo, lo que permite la aplicación de pruebas a sujetos cuyo nivel de conocimientos del uso de la computadora sea mínimo. Así lo especifican Backhoff y cols. (1996) al reportar que sujetos preuniversitarios sin conocimiento de computación, pudieron fácilmente ejecutar una prueba computarizada en la que requerían usar únicamente cinco teclas de respuestas ("a", "b", "c", "d", y "e") para indicar la opción que consideraban la mejor respuesta a cada reactivo que se les presentaba.

De la revisión bibliográfica realizada, se puede concluir que se tiene escaso conocimiento de investigaciones sobre versiones computarizadas en los que se requiere de una interacción más compleja, que la de simplemente marcar una tecla. De ahí se propone que en las investigaciones en las que se pretenda llevar a las computadora una prueba psicológica deben de estar presentes por lo menos dos preguntas fundamentales: a) si existen cambios significativos entre la ejecución tradicional (lápiz-papel) y la computarizada y b) si los grados de interacción exigidos en las interfaces para las pruebas computarizadas pueden ser una variable que influya en la ejecución de los sujetos y, por ende, en la calificación obtenida por ellos.

Esto hace pensar que la segunda generación de pruebas computarizadas no serán ya meras traducciones de las versiones de lápiz-papel, sino que se orientarán más a poder aprovechar la extraordinaria versatilidad de la computadora, alejándose de las interacciones simples.

Es por ello que el objetivo de este trabajo ha sido comparar la ejecución en pruebas psicométricas estandarizadas en su versión tradicional (lápiz-papel) versus computarizada, así como analizar la ejecución cuando se requiere una interacción simple en lugar de otra

más compleja. Para atender este propósito, se seleccionaron dos pruebas que han sido ampliamente utilizadas desde hace ya mucho tiempo, por lo cual sus índices de validación y confiabilidad han llegado a ser muy altos, dado el proceso de adecuación continuo que han tenido.

Las pruebas que se seleccionaron fueron el Test de Matrices Progresivas, Escala para la Medida de la Capacidad Intelectual, Escala Especial, para edades de 4 a 11 años (RAVEN), considerada como una prueba que, al llevarla a la computadora, mantiene una estructura simple (solamente se emplea el clic del mouse), y la Subprueba de Ordenación de Dibujos de la escala de Inteligencia Wechsler para Nivel Escolar, Revisada en México (WISC-RM), considerada para efectos de esta investigación como una prueba que requería de una interfaz de mayor nivel de interacción (además del clic de mouse, se requería clic sostenido y arrastrar y soltar objetos).

## MÉTODO

### Materiales

Se emplearon la prueba (RAVEN) y la Subprueba de Ordenación (WISC-RM), las cuales fueron manejadas en sus versiones tradicional (lápiz-papel) y se construyó la traducción a versiones computarizadas.

Se utilizaron dos computadoras con procesador 486, monitor a color con la interfaz gráfica Windows. Los programas computarizados se desarrollaron en Visual Basic, con una interacción basada en el aditamento indicador de mesa (ratón-*mouse*).

### Sujetos

Participaron 104 escolares, 43 cursaban el tercer grado y 61 el sexto grado de primaria. 57 lo hacían en escuela pública y 47 en privada. A los niños de tercer grado se les aplicó el Test de Matrices progresivas (RAVEN) y a los de sexto grado la prueba de Ordenación de Dibujos (WISC-RM). 53 de ellos contestaron en versión lápiz-papel y 51 en versión computarizada. La asignación a las condiciones según con la versión de la prueba, lápiz-papel o computarizada, se realizó considerando la lista alfabética de cada grupo; la primera mitad del grupo pasaría por la condición lápiz-papel y la segunda por la versión computarizada.

En la aplicación del Test de Matrices Progresivas, las edades de los sujetos estuvieron dentro de un intervalo de 7 a 10 años en la primaria pública; en la primaria particular, el intervalo de edades fluctuó de 8 a 10 años. En la prueba de Ordenación de Dibujos, las edades de los participantes abarcaban un intervalo de 11 a 13 años en la escuela privada, mientras que el intervalo de edades fue de 10 a 13 años en la escuela pública.

### **Procedimiento**

El procedimiento de aplicación y calificación para la versión lápiz-papel de ambas pruebas fue exactamente el especificado en el protocolo de las pruebas. En las aplicaciones de las versiones computarizadas, la indicación fue apegarse a las instrucciones que la máquina iba presentando, dado que esta misma se encargó de todo el registro de la ejecución. Como parte complementaria se hicieron registros anecdóticos de las sesiones. Cabe señalar que, para el objeto de esta investigación, la prueba de Ordenación de Dibujos se aplicó en su totalidad, es decir, aun cuando los sujetos cumplieran con el criterio de terminación de la prueba (realizar tres errores consecutivos), tanto el aplicador de la prueba como la computadora continuaban presentando los reactivos hasta llegar al final.

### **Aplicaciones lápiz-papel**

Se decidió comenzar la aplicación con las versiones lápiz-papel, tanto del WISC como del RAVEN. Por ello se determinaron las características principales de cada una de ellas dentro de su aplicación, y las diferentes clases de respuesta que los niños generaron en su ejecución. Así, al momento de realizar la programación, se contaría con una guía de los elementos más relevantes de la aplicación con el formato lápiz-papel.

La aplicación se realizó en las escuelas durante la primera mitad del día, en un aula de clases que fue asignada para este trabajo.

### **Aplicaciones computarizadas**

#### **1. Desarrollo de versiones computarizadas**

La programación de las versiones se efectuó paralelamente a la aplicación lápiz-papel, lo que permitió —como ya se dijo— programarlas tomando en cuenta el mayor número posible de factores que intervenían en la aplicación.

En la versión computarizada, la estructura se basó en cinco componentes:

- Presentación
- Entrenamiento (Tutorial en el uso del mouse y del teclado)
- Hoja de datos (Registro de datos generales del usuario)
- Desarrollo de la prueba
- Fin del programa

Cada una de las pruebas, en tanto que presentan características de ejecución muy particulares, exigió un arreglo distinto que diferencia el nivel de interacción solicitado al usuario. Así, para la prueba de Matrices Progresivas, el trabajo se centró en que la interfaz permitiera a los sujetos señalar o indicar la figura que correspondiera a cada reactivo, por lo

que sólo se manejó el uso del "clic" del mouse como estrategia de interacción. Por otro lado, la prueba de ordenación de dibujos requirió, además del uso del "clic" del ratón, la función Arrastrar y soltar (drag and drop), para que los sujetos colocaran cada figura del reactivo en su lugar correspondiente. Esto permite suponer un nivel de interacción "más complejo" en la prueba de ordenación de dibujos que en la de matrices progresivas.

Las dos pruebas fueron llevadas a la versión computarizada apegándose a todas sus características. En la prueba de matrices progresivas, sólo en algunos casos los colores no fueron igualados con exactitud, mientras que en la prueba de ordenación de dibujos, las dimensiones de las tarjetas de algunos reactivos (específicamente los de cuatro y cinco dibujos) fueron reducidas en sus dimensiones originales debido a las limitaciones de espacio del monitor de 14"; sin embargo, estimamos que estas variantes no son suficientemente relevantes como para modificar los resultados de una manera significativa.

## 2. Registros

El programa computarizado tiene integrado una base de datos donde se llevan todos los registros de la ejecución de cada participante. Como un complemento, se levantaron registros anecdóticos con el objeto de ubicar estilos de interacción, dificultades o deficiencias de interacción, de instrucciones, de tutoriales y desarrollo de la prueba.

## RESULTADOS

La estimación de resultados se basó principalmente en análisis de varianza, en los que se compararon las diferencias obtenidas en las puntuaciones para cada versión en cada una de las pruebas. De esta manera, se realizaron dos análisis distintos (uno correspondiente a cada prueba) que, a su vez, se dividieron de dos formas:

- 1) Una comparación entre las ejecuciones de lápiz-papel y las computarizadas (intraescuela).
- 2) Una comparación entre las escuelas (interesuela) de distinto tipo (pública y privada), donde se equipara la ejecución entre pruebas equivalentes (lápiz-papel y computarizada).

## RAVEN

Bajo estos esquemas de análisis, los resultados en la Prueba de Matrices Progresivas no muestran diferencia estadísticamente significativa ( $p=.375$  y  $p=.917$ ) al hacer las comparaciones entre la versión computarizada y la de lápiz-papel (tabla 1).

En las comparaciones entre escuelas públicas y privadas (tabla 1), tampoco fueron significativas en ninguna de las dos versiones de la prueba ( $p=.154$  para la versión lápiz-papel y  $p=.394$  para la versión computarizada).

**Tabla 1. Medias aritméticas y análisis de varianza obtenidos en la ejecución del test de matrices progresivas raven.**

	PRIVADA		PÚBLICA		ANOVA	G.L.
	MEDIA	DESV.EST.	MEDIA	DESV.EST.		
L-P	78,636	12,468	53,182	31,723	0.154	1
COMP	58,125	29,633	67,000	32,762	0.394	4
ANOVA	0.375		0.917			
G. L.	1		4			

### ORDENACIÓN DE DIBUJOS

Para el caso de la prueba de Ordenación de Dibujos (tabla 2) no se encontraron diferencias estadísticamente significativas al comparar las ejecuciones en la versión lápiz-papel *versus* computarizada ( $p=.905$  y  $p=.898$ , respectivamente). Sin embargo, al hacer las comparaciones entre las escuelas (pública *versus* privada), se aprecia una diferencia que sí es estadísticamente significativa, a favor de la escuela privada en el caso de la versión lápiz-papel ( $p=.05$ ), pero en el caso de la versión computarizada no hubo una diferencia significativa ( $p=.144$ ).

**Tabla 2. Medias aritméticas y análisis de varianza de la ordenación de dibujos WISC-RM calificaciones con descalificación**

	PRIVADA		PÚBLICA		ANOVA	G.L.
	MEDIA	DESV.EST.	MEDIA	DESV.EST.		
L-P	11	2,253	10,235	2,513	0.058	6
COMP	8,857	2,983	3,250	2,697	0.144	6
ANOVA	0.898		0.905			
G. L.	6		6			

Para esta prueba se llevó a cabo otro procedimiento de aplicación y una manera alterna de analizar los resultados, la cual consistió en ignorar la descalificación como lo específica el protocolo, es decir, la discontinuación de la prueba se efectúa después de tres fracasos seguidos. Esta decisión se tomó al considerar algunas observaciones consignadas en los registros anecdóticos, en los que se mostraba claramente que los niños, al interactuar

con los primeros estímulos, se equivocaban constantemente, lo que provocaba su descalificación desde los primeros reactivos. Al permitirles proseguir, mejoraban notablemente la ejecución en los estímulos posteriores, por lo que se les permitió continuar y tener claramente establecida la diferencia, esto es, se registró una calificación de corte al cometer tres errores consecutivos, tal como lo indica el protocolo de la prueba y otra de ejecución ininterrumpida, lo que permite analizar las diferencias en los resultados encontrados.

De esta manera se realizaron los análisis intraescuela e interescuela, con la particularidad de que las calificaciones tomadas para este análisis fueron sin descalificación y un análisis de varianza entre los dos tipos de calificación asignada (p. e.: se efectuó la comparación entre la calificación lápiz-papel de escuela pública con discontinuación, con la calificación lápiz-papel de escuela pública sin discontinuación).

Los análisis de varianza de los resultados sin descalificación (tabla 3) no arrojaron diferencias estadísticamente significativas en la comparación intraescuela para los dos tipos ( $p = .966$  para la pública y  $p = .992$  para la privada), ni interescuela ( $p = .108$  para lápiz-papel y  $p = .352$  para computadora).

**Tabla 3. Medias aritméticas y análisis de varianza de la ordenación de dibujos WISC-RM calificaciones sin descalificación.**

	PRIVADA		PÚBLICA		ANOVA	G.L.
	MEDIA	DESV.EST.	MEDIA	DESV.EST.		
L-P	11	2,253	10,353	2,290	0.108	6
COMP	9,375	2,098	4,188	2,373	0.352	5
ANOVA	0.992		0.966			
G. L.	5		6			

No obstante, en el análisis de los dos tipos de calificación asignada sí encontramos diferencias en la calificación de la ejecución en las versiones computarizadas de ambos tipos de escuelas. La diferencia de calificación discontinuada y sin discontinuar para la versión computarizada en la escuela pública y en la privada, fue estadísticamente significativa en ambas ( $p = .000$  en la privada y  $p = .001$  en la pública), observando una mejoría en las calificaciones asignadas sin discontinuar la prueba para ambas escuelas; en las versiones lápiz-papel no se encontraron diferencias (tabla 4).

**Tabla 4. Medias aritméticas y análisis de varianza: calificaciones sin comparación entre las distintas asignaciones de calificación de la ordenación de dibujos WISC-RM**

	CON DESCALIF.		SIN DESCALIF.		ANOVA	G.L.
	MEDIA	DESV.EST.	MEDIA	DESV.EST.		
PÚB_L-P	10,235	2,513	10,353	2,290		4
PÚB_COMP	3,250	2,697	4,188	2,373	0.001	6
PRI_L-P	11	2,253	11	2,253		6
PRI_COMP	8,857	2,983	9,375	2,098	0.001	5

### CONCLUSIONES

Las ejecuciones en las versiones computarizada y lápiz-papel, a pesar de presentar ligeras tendencias a favor de una o de otra presentación, no representan la existencia de distancias significativas atribuibles a la presentación de las pruebas, es decir, como ya lo han mencionado Fekken y Holden (1989) y Holden y Fekken (1990) en pruebas clínicas estandarizadas, existe una equivalencia entre las pruebas lápiz-papel y las versiones desarrolladas en computadora. De igual manera, estos resultados son análogos a los encontrados por Backhoff y cols. (1996) en la utilización de un sistema computarizado de exámenes de admisión a nivel licenciatura y su aplicación en lápiz-papel.

Sin embargo, las diferencias en los niveles de exigencia de interacción computadora-infante, tal como se esperaba (considerando que cada prueba requería un menor o mayor grado de interacción), sí nos permite hablar de diferencias cuantitativas y cualitativas, las que evidentemente sería un gran error pasar por alto. Las diferencias a las que nos referimos son las de la interacción máquina-infante, que se observaron como importantes para la puntuación asignada al final de las pruebas, específicamente en la de Ordenación de Dibujos.

Las pruebas seleccionadas para evaluar los diferentes niveles de exigencia en interacción con la computadora mostraron, efectivamente, generar distintos niveles de interacción.

Para la prueba de Matrices Progresivas, la interacción con la máquina fue menor, únicamente era necesario "hacer clic" con el control o aditamento indicador de mesa (*raton-mouse*), es decir, en todo el programa se pedía que el sujeto ejecutara "clicks" en los botones y en las imágenes correspondientes; esto hace a la respuesta muy similar a las pruebas

desarrolladas por Fekken y Holden (1989) y Holden y Fekken (1990), donde sólo se requería usar las teclas "T" y "F" para responder, y -al oprimir cualquiera de éstas- la pantalla borraba el reactivo en turno y aparecía el otro; o a los trabajos de Backoff y cols. (1996), donde también sólo era necesario utilizar las cinco teclas indicadas para responder.

En cambio, la subprueba de Ordenación de dibujos del WISC implicaba una mayor exigencia de habilidades en la interacción con la computadora; ya que además del "clic" en lugares específicos, era necesario la habilidad de "arrastrar" y "soltar" imágenes en determinada secuencia para efectuar la tarea requerida.

Esta diferencia, establecida por el nivel de interacción con la computadora, se comprobó fácilmente en la ejecución de niños de escuela privada, cuyo contacto con las computadoras existe desde los primeros años de formación, con los sistemas más comúnmente usados actualmente (ambiente Windows). Se detectó una diferencia considerable al comparar con los niños de escuela pública, en los que la manipulación de la computadora es escasa, lo que reduce significativamente la probabilidad de familiaridad. Esta falta de familiaridad con la computadora se reflejó en la complejidad que para estos niños representó la interacción, por lo que fue necesario instrumentar una fase de modelamiento en la interacción con la computadora para posibilitar el cumplimiento de la prueba en la escuela pública. La aplicación computarizada incluía una fase de entrenamiento, la cual cubrió los aspectos de conocimiento respecto de las funciones que se usarían durante cada prueba; sin embargo, creemos que no logró generar la suficiente familiaridad con estas tareas.

Tal evidencia apunta a considerar como un aspecto importante en la computarización de pruebas ya estandarizadas, el tipo de interacción que el individuo debe tener con la máquina y la familiaridad de los sujetos participantes con el uso de las computadoras, ya que ésta mostró ser un factor que determinó en gran medida el proceso de interacción en las pruebas y su consecuente resultado. Si la respuesta requerida es relativamente sencilla, como el apretar una u otra tecla, así como el sólo hacer clic, entonces se espera que las ejecuciones sean semejantes, tanto en la versión lápiz-papel como en la versión computarizada. Si la respuesta requerida implica una mayor habilidad o destreza en el manejo de diferentes funciones de la computadora, las ejecuciones de los individuos ante la prueba computarizada pueden variar con su correspondiente versión lápiz-papel, de acuerdo con el nivel de familiaridad que se tenga con estos procesos.

Tomando en cuenta tales consideraciones, se debe alertar que, en los casos que se desee llevar a la computadora una prueba ya estandarizada, se considere como una variable importante el nivel de interacción individuo-máquina. De igual forma, nuestros resultados nos sugieren investigar esta variable con mayor amplitud para conocer realmente el tipo de efectos que puede producir.

Finalmente, los resultados de esta investigación fortalecen la idea del uso de la computadora como instrumento de evaluación. En esta investigación, el uso de esta herramienta sí permitió una interacción dinámica, un aprendizaje individualizado, relacionar la ejecución con consecuencias que favorecen la motivación del participante, mejorar los niveles de atención —con lo que se disminuye el nivel de errores,, siempre y cuando se consideren todos los elementos que sobre interacción y familiaridad hemos discutido—, proveer al practicante con una retroalimentación inmediata acorde con sus respuestas y evaluar de manera individualizada, automática e inmediata el desempeño del estudiante, registrando sus respuestas en una base de datos y generando los análisis estadísticos para calificar con referencia a la norma.

La investigación, en su conjunto, sugiere considerar siempre los factores que intervienen en la interacción sujeto-máquina, así como las habilidades requeridas para su ejecución en la generación de las interfaces que para estos fines se estructuran. Además, se aconseja la elaboración de fases de entrenamiento y familiaridad en el uso de la máquina (tutoriales), de tal manera que, hasta que se logre un estándar en la interacción sujeto-máquina, se le permita al sujeto comenzar la prueba de ejecución propiamente dicha, neutralizando de esta manera los efectos de esta variable, en la evaluación general.

#### BIBLIOGRAFÍA

- Backhoff, E. (1993). Normas Estadísticas Hispánicas de Tests Psicológicos: ¿Son Adecuadas para la Población Mexicana?. *Revista Mexicana de Psicología*, 9, 117-125.
- Backhoff, E., Ibarra, M. y Rosas, M. (1996). Desarrollo y Validación del Sistema Computarizado de Exámenes (SICODEX). *Revista de la Educación Superior*, 25(1), 41-54.
- Béjar, I. (1991). A Methodology for Scoring Open-Ended Architectural Design Problems. *Journal of Applied Psychology*, 56, 522-532.
- Carey, S. (1986). Cognitive Science and Science Education. *American Psychologist*, 41, 1123-1130.
- Fekken, C. y Holden, R. (1989). Psychometric Evaluation of the Microcomputerized Personality Research Form. *Educational and Psychological Measurement*, 49, 875-882.
- Holden, R. y Fekken, C. (1990). Clinical Reliabilities and Validities of the Microcomputerized Basic Personality Inventory. *Journal of Clinical Psychology*, 46, 845-849.
- Johnson, R., Johnson, D. y Stanne, M. (1985). Effects of Cooperative, and Individualistic Goal Structures on Computer-Assisted Instruction. *Journal of Educational Psychology*, 77, 668-677.
- Lepper, M. (1985). Microcomputers in Education. *American Psychologist*, 40, 1-18.

- Lepper, M. y Gurtner, J-L. (1989). Children and Computers. **American Psychologist**, 44, 170-178.
- McCullough, L., Farrell, A. y Longabaug, R. (1986). The Development of a Microcomputer-Based Mental Health Information System. **American Psychologist**, 41, 207-214.
- Muth, D. (1984). Solving Arithmetic Word Problems: Role of Reading and Computational Skills. **Journal of Educational Psychology**, 76, 205-210.
- Rambally, G. y Rambally, R.. (1987). Human Factors in CAI Design. **Computers and Education**, 11, 149-153.
- Tirado, F. (1994). La Evaluación Sistemática Computarizada. Hacia Nuevos Procesos Reguladores de la Práctica Educativa. **Asociación Mexicana de Investigadores de la Educación**, 1(2), 18-24.