



**CONGRESO
IBEROAMERICANO**
DE CIENCIA, TECNOLOGÍA,
INNOVACIÓN Y EDUCACIÓN

BUENOS AIRES, ARGENTINA
12, 13 Y 14 DE NOVIEMBRE 2014

**CONGRESSO
IBERO-AMERICANO**
DE CIÊNCIA, TECNOLOGIA,
INOVAÇÃO E EDUCAÇÃO

BUENOS AIRES, ARGENTINA
12, 13 Y 14 DE NOVIEMBRE 2014

Escalas de medición de habilidades cognitivas, sociales y tecnológicas (EMHCoST) para niños de 9 a 12 años

Salinas, M.; Santiago, J.; Pesqueira, N.; Barrientos, S.

Escalas de medición de habilidades cognitivas, sociales y tecnológicas (EMHCoST) para niños de 9 a 12 años

*Dr. Moisés Salinas Fleithman

msalinasphd@gmail.com

**Dra. Jesús Mónica Santiago Ramírez.

jmonyk@gmail.com

**Dra. Norma Guadalupe Pesqueira Bustamante.

nopesque@hotmail.com

**Mtra. Soledad Barrientos López.

sbarrientoslopez@gmail.com

*Universidad ORT México

**Centro Regional de Formación Docente e Investigación educativa

Resumen

El presente es un reporte del diseño, estandarización y aplicación de un instrumento de medición de habilidades cognitivas, sociales y tecnológicas en niños de 4°, 5° y 6° de primaria. La evaluación de estas habilidades es de suma importancia ya que en general no están reflejadas en la evaluación de desempeño académico tradicional, pero son fundamentales para el éxito personal y profesional en el siglo XXI. El estudio tiene un corte cuasi experimental con grupo de control, el tratamiento aplicado es el uso libre de la computadora XO. Los sujetos de investigación están contextualizados en una comunidad rural y en el ambiente urbano marginado de una ciudad del Norte de México. La técnica de recolección de datos es el auto reporte con base en escalas de medición de las habilidades de: investigación científica, tecnológica, razonamiento y pensamiento crítico, pensamiento creativo, aprendizaje independiente y desarrollo humano y crecimiento personal. El presente es un reporte de la elaboración y validación del instrumento

Palabras clave: Escalas de medición, Técnicas de medición y Autopercepción

Introducción

La integración y uso de tecnología en el contexto escolar es un proceso fundamental si queremos generar las competencias necesarias para que los niños sean competitivos en el siglo XXI, a nivel mundial, pero particularmente en América Latina donde uno de los programas más ambicioso con el uso de computadoras, es el Plan Ceibal, en el que se introdujeron computadoras XO, diseñada por el proyecto *One Laptop Per Child*. Ellos han documentado sistemáticamente el proceso de entrega de computadoras y han generado indicadores a través del Departamento de monitoreo y evaluación del plan Ceibal, en este estudio reportan beneficios significativos con la incorporación de la XO en el uso de algunas herramientas de programación de la XO. Así mismo, la evaluación del Plan Ceibal reporta un comparativo de como el desarrollo de habilidades digitales como: chatear, bajar música, buscar información en internet, escribir, dibujar y pintar con la XO es superior en alrededor de 15 a 20 puntos porcentuales en comparación a una computadora PC.

Otra experiencia Latinoamericana es la que se desarrolla en Perú. En ella se realizó una Evaluación Experimental Del Programa "Una Laptop Por Niño", con el propósito de medir los resultados del impacto del programa en las áreas: de

habilidades cognitivas con medidas de aprendizaje curricular, habilidades TIC'S, habilidades no cognitivas: capacidad de resolución de problemas, de colaboración, de trabajo en equipo, de aprendizaje auto guiado y conductas: deserción, asistencia, motivación, esfuerzo, utilización del tiempo dentro y fuera de la escuela

Sin embargo, a pesar de las evaluaciones que se han llevado a cabo en Latinoamérica, y que han mostrado beneficios del uso de las computadoras XO, aun falta diseñar y realizar evaluaciones mas puntuales sobre el impacto del uso de tecnología en general, y de las computadoras XO en particular, en aquellas habilidades y competencias que han sido identificadas como más significativas para la competitividad individual y social en el mundo digital y globalizado del siglo XXI, entre ellas: pensamiento analítico, Inteligencia social, innovación, habilidades tecnológicas, y colaboración, (Davies, Fidler, y Gorbis, 2011). En general el impacto generado en las habilidades de los alumnos y desempeño académico en los citados ejemplos anteriores, fueron evaluados utilizando indicadores tradicionales de desempeño académico y centrados en los contenidos del currículo formal de los estudiantes; sin embargo más allá del éxito curricular, los niños del siglo XXI, se enfrentan al reto de integrar en su formación habilidades que les permita tener éxito en la vida social y laboral a la cual se enfrentarán en un futuro próximo. En ese sentido existe un vacío que permita medir habilidades como las mencionadas por Davis y sus colegas en los niños. Más aún, para este propósito era necesario realizar un instrumento normado y validado en español que nos permita realizar las mediciones a gran escala, dado que la tendencia en la política educativa en Latinoamérica es implementar programas que involucran el uso de las computadoras portátiles con estudiantes de nivel básico.

En ese sentido la presente investigación tuvo como propósito:
Diseñar un instrumento para medir en niños de 4°, 5° y 6° de primaria las habilidades de:

1. Investigación científica.
2. Pensamiento creativo.
3. Razonamiento y pensamiento crítico.
4. Humanas y de crecimiento personal.
5. Aprendizaje independiente.
6. Tecnológicas.

La metodología se llevó a cabo en tres fases: 1) identificación de conceptos, habilidades y base de datos de reactivos, 2) desarrollo del instrumento, y 3) validación del instrumento.

Identificación de conceptos y de reactivos. En una primera fase, se hizo una revisión colegiada de instrumentos que miden habilidades y competencias no tradicionales, pero que han sido identificadas como fundamentales para el aprendizaje significativo (Salinas y Garr, 2009). Los instrumentos examinados fueron: Assessment of Learner Centered Practices (Mc Combs & Miller 2007), Khatena-Morse Multi-Talent Perception Inventory (Khatena & Morse, 1994), Classroom Environment Scale (Moos & Trickett, 2002) y Process Skills Rating Scales (Karnes, F. & Bean, S. 2004), como se indica en la tabla 1.

Factores	Instrumento	Escalas consideradas	Numero de Reactivos
Motivación y auto regulación	<i>Assessment of Learner Centered Practices</i>	Todas	75
Creatividad	<i>Khatena-Morse Multi-</i>	Creatividad	50

	<i>Talent Perception Inventory</i>	Iniciativa Liderazgo Versatilidad	
Trabajo en grupo	<i>Classroom Environment Scale</i>	Afiliación Innovación	20
Pensamiento científico y crítico, habilidades de proceso	<i>Process Skills Rating Scales</i>	Investigación independiente Crecimiento personal Pensamiento creativo Razonamiento y pensamiento crítico Investigación científica Habilidades tecnológicas	360

Tabla 1, instrumentos analizados para la construcción de las escalas

El análisis se llevó a cabo a través del método de juicio de expertos. Nueve miembros del equipo académico, conformado por personal profesional del Centro Regional de Formación Docente e Investigación Educativa, así como expertos académicos en Massachusetts, EEUU, y México, revisaron y evaluaron los instrumentos base, y luego se tomó una decisión colegiada sobre los factores y reactivos que formarían la base de datos para el desarrollo. Se eligió el *Process Skills Rating Scales Revised* (PSRS-R) de Frances A. Karnes y Suzanne M. Beane (2004), que ya incluía la mayoría de los factores que los expertos consideraron eran centrales para el presente programa.

El PSRS-R es un instrumento de auto-reporte, diseñado como apoyo en el proceso de enseñanza aprendizaje, y enfocado en habilidades de pensamiento, razonamiento y aprendizaje auto-regulado. Su objetivo es determinar el nivel que los estudiantes tienen en 13 áreas no académicas, las cuáles predicen el éxito escolar. Esta prueba no es estandarizada, con reactivos tipo Likert de frecuencia que van de "Siempre" a "Nunca", y puede utilizarse a través de técnica de autor-reporte o bien a través de la observación de padres o maestros. Las escalas pueden ser utilizadas para diferentes situaciones, entre otras: monitorear los avances de los alumnos en las diferentes áreas, retroalimentación para estudiantes, padres o maestros sobre los procesos de desarrollo que siguen los estudiantes, ayudar a los profesores a determinar los avances de los estudiantes, guía para proveer de información a los estudiantes sobre sus propios avances y a los padres de familia sobre las áreas que pueden apoyar a sus hijos en casa. Así mismo, el instrumento era adecuado como base para el presente estudio ya que fue diseñado para ser utilizado con estudiantes de nivel primaria y secundaria.

Desarrollo del instrumento. Una vez seleccionadas las escalas y reactivos base para el desarrollo del instrumento, el equipo experto realizó un análisis ítem por ítem para determinar su relevancia para el presente estudio. Los reactivos se seleccionaron o eliminaron basados en los siguientes criterios: 1) relevancia conceptual para el estudio, 2) relevancia para la población del estudio, ya sea por edad o por contexto cultural, y 3) lenguaje que fuera apropiado para el segmento de edad de los participantes. Así mismo, se propusieron reactivos adicionales para la base de datos, con el propósito de evaluar la habilidad de uso de las computadoras XO de *One Laptop Per Child*, de manera que pudiera ser adaptado a las necesidades específicas del programa para el que el presente instrumento fue diseñado.

En total, Se hizo una reducción total acumulada de 18% de los ítems. Finalmente, se realizó un piloto técnico en un grupo de 10 niños de diferentes edades

fluctuantes entre los 8 y 12 años de edad para revisar su redacción y pertinencia en el campo.

Una vez realizado el primer pilotaje se decidió a través de un juicio de expertos, que era importante añadir algunos reactivos relacionados con el uso de la XO y el aprendizaje independiente, se diseñaron reactivos y se colocaron en el instrumento para realizar un segundo pilotaje que nos diera como resultado un instrumento definitivo.

Se aplicó la prueba una vez más en una población de 60 niños de 4to., 5to. y 6to. de primaria para probar y refinar el protocolo de aplicación.

Posteriormente, se llevo a cabo un análisis factorial confirmatorio utilizando un método de extracción de análisis de componentes principales, y una rotación varimax. El análisis en general confirmo la estructura de seis escalas. Reactivos con una carga menos a .50 en alguno de los factores fueron evaluados por el comité de jueceo para determinar su relevancia, y si se consideró que no eran fundamentales, fueron eliminados. El número de reactivos se modificó de acuerdo a los procesos mencionados como lo muestra la tabla 2

Escala	Número de ítems originales	Número de ítems del instrumento
Habilidades tecnológicas	45	40
Investigación científica	86	41
Razonamiento y pensamiento crítico	77	46
Investigación independiente	29	26
Pensamiento creativo	43	29
Habilidades humanas y de crecimiento personal	80	57

Tabla 2. Número de ítems finales.

Validación del Instrumento. La aplicación del instrumento se llevó a cabo en dos escuelas, la escuela experimental y la escuela de control. Para la validación del instrumento se procedió a realizar varias etapas, la primera de ellas consistió en la integración y depuración de la base de datos de alumnos, considerando solo a los alumnos que aplicaron en ambos momentos (pre y post test). El número total de alumnos por grado y escuela se muestra en la tabla 3. En total se contó con 160 registros, 87 de la escuela experimental y 73 de la escuela de control.

Se verificó el ajuste psicométrico de los ítems que conforman cada escala, de acuerdo con la Teoría de Respuesta al Ítem. Para ello se utilizó el modelo de crédito parcial, que forma parte de los modelos Rasch para ítems politómicos. Este permite analizar ítems que corresponden a respuestas graduadas en varios niveles de cuantificación, como es el caso de la escala tipo Likert (Martínez, Hernández & Hernández, 2006; Fisher & Molenaar, 1995). El procesamiento se realizó con el programa Winsteps (Linacre, 2012). Además se presenta el alpha de Cronbach como medida de la consistencia interna.

Las ventajas de la utilización del modelo Rasch son varias, entre ellas destacan las siguientes:

- a) Medición conjunta. Los parámetros de las personas y de los ítems se expresan en las mismas unidades y se localizan en el mismo continuo.
- b) Objetividad específica: La medición de personas en un atributo no depende de los ítems específicos con los fueron medidos. Igualmente, las características del test es independiente de los sujetos de la muestra (Rasch, 1977).

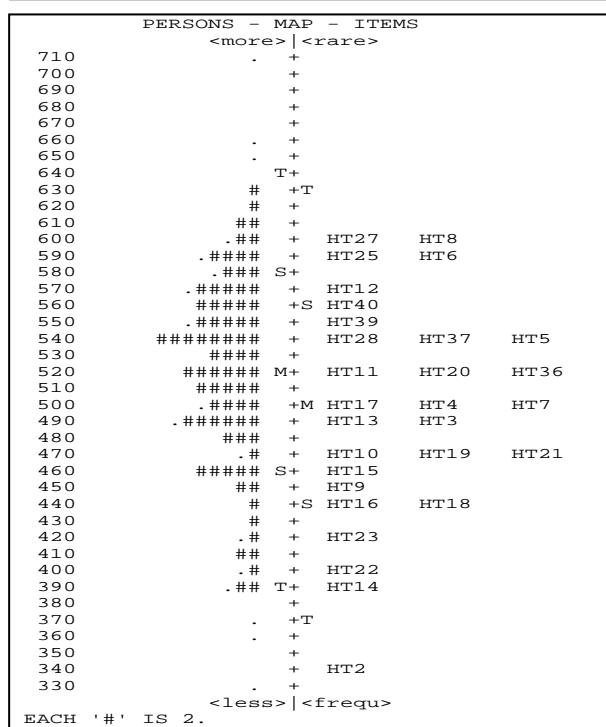
- c) Especificidad del error típico de medida. Permite cuantificar la cantidad de información con la que se mide en cada punto de la dimensión y seleccionar los ítems que incrementan la información en regiones del atributo previamente especificadas.

Los criterios de bondad de ajuste que se consideraron para cada uno de los reactivos son los siguientes:

- 1) El ajuste interno (INFIT) y el ajuste externo (OUTFIT). El primer estadígrafo mide si el reactivo que se analiza está siendo respondido por personas cuyo nivel de rasgo es cercano a la exigencia del reactivo. En reactivos de opinión o percepción los reactivos que componen una escala deben ajustar al modelo de contraste con valores de INFIT y OUTFIT (Linacre, 2011) en el intervalo de 0.50 a 1.50.
- 2) Correlación punto biserial (PTBIS), se interpreta como una correlación que mide el grado de asociación entre el reactivo individual y el puntaje total observado por la escala. Los valores altos de la correlación biserial son un indicativo de que el ítem pertenece a la escala unidimensional respectiva. Se estableció el criterio de PTBIS igual o mayor a 0.20.
- 3) Índice de Discriminación (DISCR) es la capacidad del ítem para distinguir entre los respondentes de bajo y alto nivel en el rasgo que se mide. Bajo el modelo de Rasch, el valor ideal para este índice es de 1.00. El criterio aquí establecido es de mínimo 0.90.

Una de las ventajas del modelo Rasch, es la de presentar en una misma escala la habilidad de los sujetos y la dificultad de los ítems. En la figura 1 se muestra el escalamiento conjunto de personas (izquierda) y reactivos (derecha), donde se aprecia que la media de las puntuaciones de los alumnos es ligeramente mayor a la dificultad media de los ítems.

Figura 1. Escalamiento conjunto de reactivos y personas. Escala de Habilidades Tecnológicas



En las figura de la 2 a la 6 se muestran las mediciones conjuntas de las escalas, en donde se aprecia, por ejemplo, que faltan ítems de alta dificultad en la mayoría de los constructos. Particularmente, en las escalas Habilidades de pensamiento y razonamiento crítico y Habilidades humanas y crecimiento personal los datos muestran que la mayoría de los alumnos tienen una alta probabilidad de contestar en las opciones positivas, es decir, es muy “fácil” que los alumnos estén de acuerdo con las afirmaciones presentadas.

Figura 2. Escalamiento conjunto de la escala habilidades de investigación científica

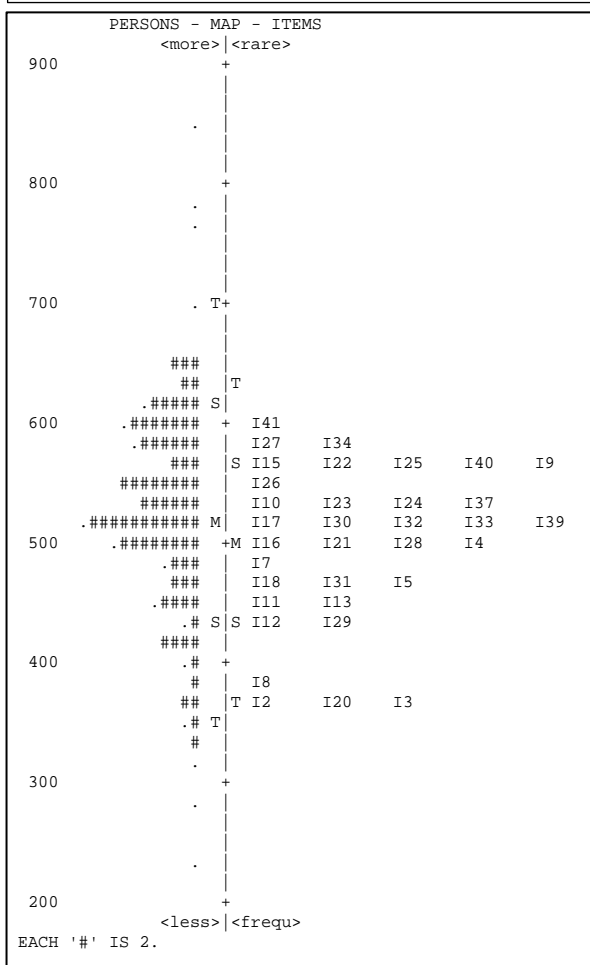


Figura 3. Escalamiento conjunto de la escala habilidades de pensamiento y razonamiento crítico

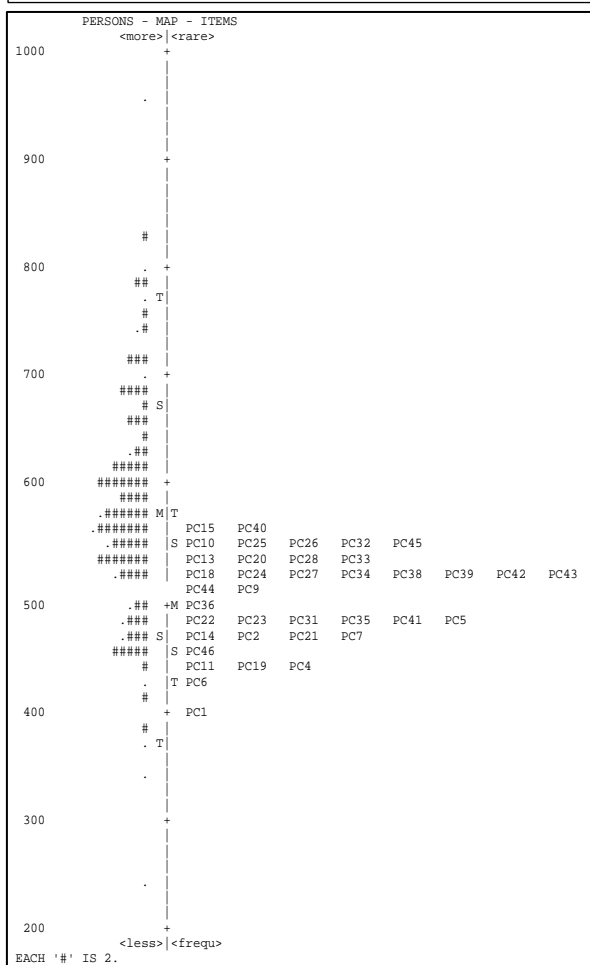


Figura 4. Escalamiento conjunto de la escala habilidades humanas y de crecimiento personal

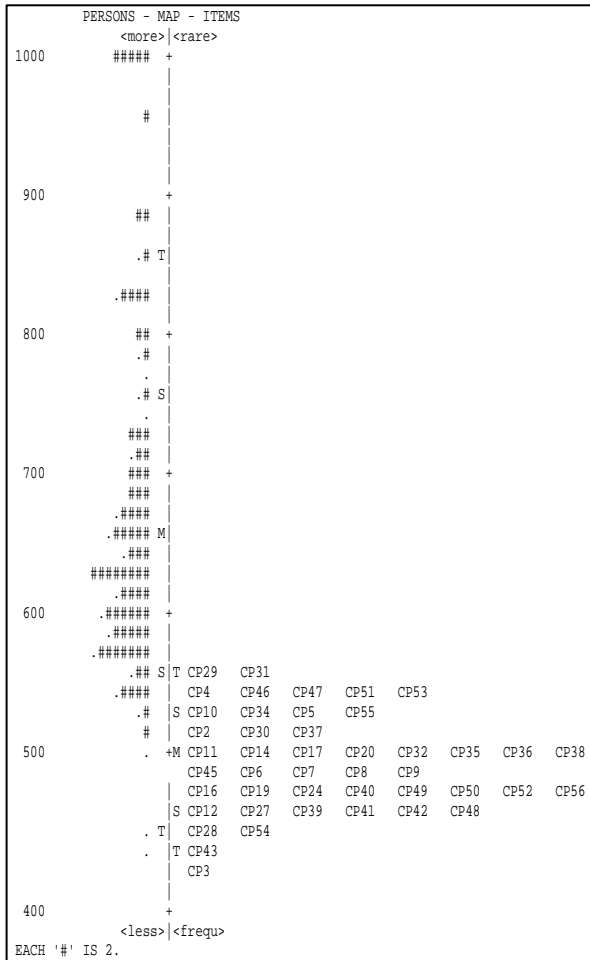
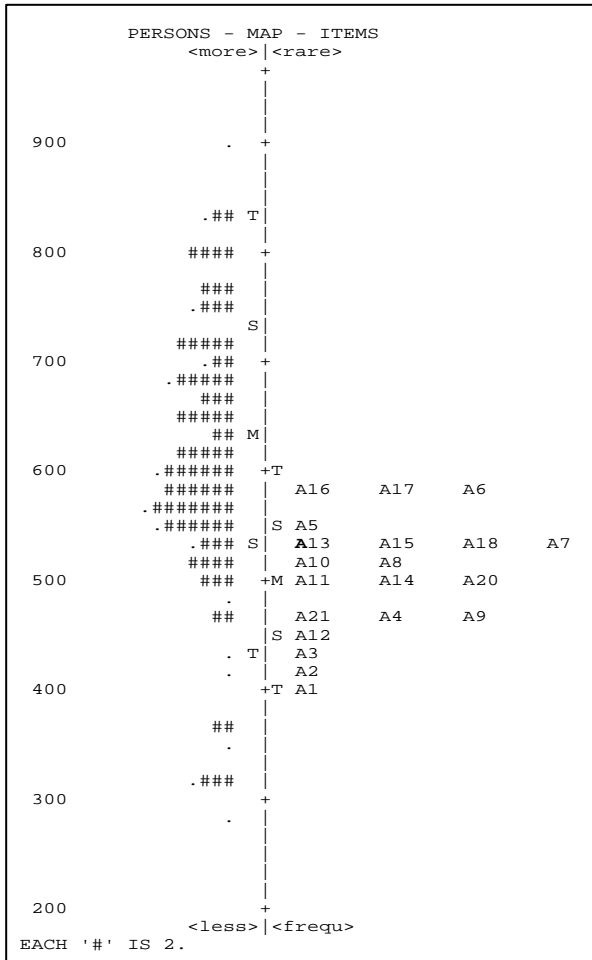
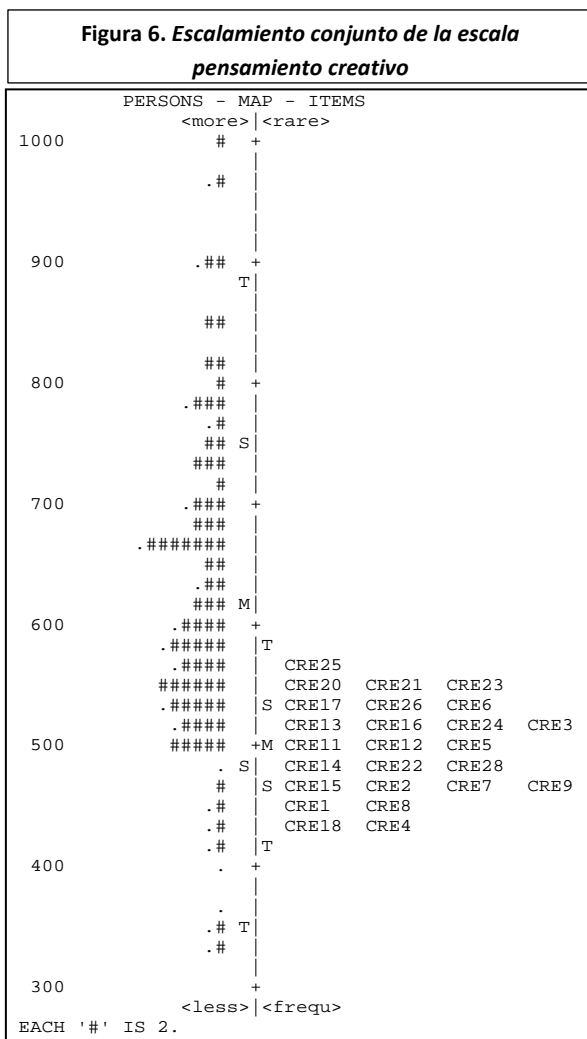


Figura 5. Escalamiento conjunto de la escala aprendizaje independiente





Conclusiones Generales

El análisis de los resultados permite formular las siguientes conclusiones:

- Las correlaciones de los ítems con la escala son adecuados y presentan una confiabilidad por consistencia elevada.
- Se cumplen con los requerimientos psicométricos de ajuste al modelo, lo que significa que las escalas están midiendo de forma prioritaria los constructos para lo que fueron diseñadas.
- En la mayoría de las escalas, hay regiones del continuo que no se miden suficientemente, particularmente falta incluir reactivos de alta dificultad que evalúen apropiadamente a los alumnos con alta habilidad.

Los resultados de estos análisis permitieron determinar los reactivos susceptibles de ser eliminados para tener una prueba más consistente y de económica aplicación.

Estos resultados indican que la EMHCoST tiene un alto nivel de coherencia interna y una confiabilidad aceptable para su aplicación en la población seleccionada.

En conclusión, los análisis de validez y confiabilidad el EMHCoST es un instrumento razonablemente valido y confiable, que puede ser aplicado en niños de 4° a 6° grado de primaria, y que nos puede dar información muy valiosa sobre el desarrollo de habilidades cognitivas, sociales y tecnológicas que no son evaluadas de manera tradicional en el salón de clases.

Discusión. En una realidad educativa dinámica y cambiante, y un mundo digital y globalizado, es imperativo cambiar el foco de la enseñanza de contenidos y

competencias básicas, a las competencias que permitan al niño contribuir a la innovación y la transformación positiva de nuestro mundo: investigación, pensamiento creativo, razonamiento, habilidades tecnológicas entre otras. Sin embargo, es muy difícil para las escuelas y los sistemas educativos, enfocarse en este tipo de competencias y habilidades, cuando carecemos de buenos modelos e instrumentos de evaluación, y continuamos atrancados en exámenes estandarizados y calificaciones cuantitativas. Por supuesto, la evaluación de estas habilidades que nos permitirán ser competitivos en el siglo XXI va más allá del uso de instrumentos de auto evaluación, e inclusive de instrumentos de evaluación de desempeño. Las habilidades complejas requieren de evaluaciones complejas, auténticas y transdisciplinarias. Sin embargo, es necesario empezar un proceso que nos permita desarrollar modelos complejos de evaluación de estas habilidades. El presente estudio busca aportar a ese proceso, con el desarrollo de instrumentos que nos permitan evaluar más allá de las competencias tradicionales y el conocimiento inerte. El desarrollo de la EMHSCoST es un paso positivo en esa dirección

Lista de referencias

Davies, A. Fidler, D., y Gorbis, M. (2011). *Future Work Skills 2020*. Recuperado en: http://cdn.theatlantic.com/static/front/docs/sponsored/phoenix/future_work_skills_2020.pdf

Fisher, G.H., & Molenaar, I. (1995). *Rasch Models, foundations, recent developments and applications*. Nueva York: Springer Verlag.

Linacre, J.M. (2012). *A user's guide to Winsteps Ministeps*. Rasch model computer programs. Beaverton, Oregon: Winsteps.com.

Martínez, M.R., Hernández M.J., & Hernández, M.V. (2006). *Piscometría*. Madrid: Alianza Editorial.

Karnes, F. y Bean, S. (2004) *Process Skills Rating Scales Revised*. Prufrock Press. Wako Texas.

Khatena, J. y Morse, D. (1994). *Khatena Morse Multitalent Perception Inventory*. Scholastic testing service Inc: Bensenville Illinois.

McCombs, B. y Miller, L. (2007) *Learner Centered Classroom practices and assessments. Maximizing student motivation, learning and achievement*. Corwin press a Sage publications: Thousand Oaks, California.

Moos, R. y Trickett. (2002). *Classroom environment scale*. Mind Garden: Menlo Park, California.

Rasch, G. (1977). On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. *The Danish Yearbook of Philosophy*, 14, 58-94.

Salinas, M. F., & Garr, J. (2009). Effect of learner centered education on the academic outcomes of minority groups. *Journal of Instructional Psychology*, 36, 226-237.