Institute of Education

UCL

Evaluación de Impacto de Math-Whizz

El caso de Aguascalientes

Este documento fue preparado por el Dr Manolis Mavrikis, UCL Knowledge Lab, University College London con el apoyo de la Dr Alice Hansen, David Sebastian Schlepps, Carolyn Ryan y Ani Mavrioz como 'UCL Consultants'.

Mas ayuda con el manejo de la base de datos fue proporcionada por José Manuel de la Fuente y Chun Pong Lau. Quisiéramos agradecer a Martha Alicia Ramos Rodríguez, Mónica Ofelia Ruiz Coronado, Christian José Correa Villalón y Juan Duch de parte del Instituto de Educación de Aguascalientes los cuales aportaron apoyo invaluable en el diseño de la investigación, la recolección de datos, así como la interpretación y el análisis. Esta evaluación no hubiera sido posible sin su valioso apoyo.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son exclusivamente del autor y no representan los puntos de vista de UCL, de la empresa Whizz Educación ni del Instituto de Educación de Aguascalientes.

Tabla de contenidos

1. Objetivos

Financiado por el Instituto de Educación en Aguascalientes, se realizó un estudio cuasi-experimental para evaluar el impacto de uno de los componentes del programa STEM en Aguascalientes, el tutor virtual Math-Whizz en clases de primaria.

Aunque hemos considerado la implementación a lo largo de los diferentes años desde su comienzo en febrero de 2015 hasta la fecha de la evaluación en Sep-Oct 2016, nos hemos centrado especialmente en el 4° grado (edades de 8 a 9 años) con base a la prueba nacional PLANEA que se llevó a cabo en septiembre del presente año. La premisa fundamental era conocer si Math-Whizz incrementa el aprendizaje de las matemáticas y la motivación de los estudiantes.

Además, el objetivo fue evaluar el proceso de aplicación, brindar recomendaciones sobre futuro desarrollo y evaluación del impacto a mayor escala.

1. Objectives

Funded by the Ministry of Education in Aguascalientes, we undertook a quasi-experimental study to evaluate the impact of one of the components of the Aguascalientes STEM program, namely the introduction of the virtual tutor Math-Whizz in primary classrooms.

Although we considered the full implementation throughout the different years since its beginning in February 2015 until the date of the evaluation in Sep-Oct 2016, we focused particularly on Year 4 (age 8 to 9) to rely on the country-wide PLANEA tests that take place in September. The overarching question was whether Math-Whizz increases learning of mathematics and students' motivation.

Additionally, the aim was also to evaluate the process of the implementation and to provide recommendations for future development and evaluation of the wider intervention.

2. Conceptos relevantes

El Instituto de Educación en Aguascalientes implementó Math-Whizz como una de las iniciativas del programa STEM en el Estado. Math-Whizz es un tutor virtual personalizado en línea para estudiantes de 5 a 13 años que genera una experiencia única de aprendizaje de acuerdo al ritmo y necesidades individuales de cada estudiante, ayudándolos en el momento en que más lo necesitan y buscando que todos tengan la oportunidad de alcanzar su máximo potencial. La suite de Math-Whizz le permite a las escuelas:

- Instrucción diferenciada a través de una experiencia personalizada para cada estudiante
- Sencilla dirección y análisis sobre el progreso y logro en matemáticas con reportes en tiempo real a nivel distrito escolar, escuela, clase y estudiante.
- Herramientas para la planeación de clase
- Visibilidad para padres de familia sobre el progreso de sus hijos

2.1 Acerca de Math-Whizz

2.1.1 Math-Whizz tutor

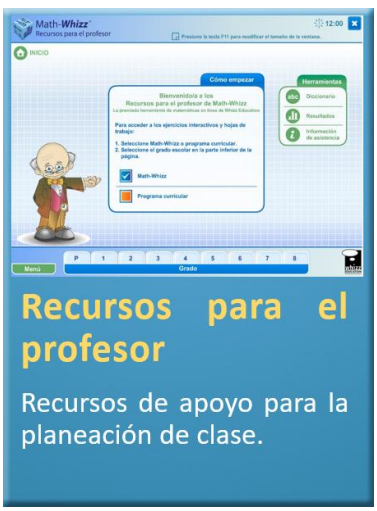
- ¿Qué es Math-Whizz?

Math-Whizz es un tutor virtual personalizado en línea para niños entre 5 y 13 años de edad que reconoce las fortalezas y debilidades de cada estudiante y le presenta lecciones y ejercicios de acuerdo a su edad matemática. El tutor virtual siempre se ajusta a las necesidades de cada alumno incrementando o disminuyendo el nivel de dificultad con objeto de crear una experiencia de aprendizaje única.

- Evaluación inicial

La evaluación inicial toma una "fotografía" de la habilidad matemática actual de cada estudiante tema por tema. Esto es presentado en un formato tipo "examen". La evaluación inicial determina el nivel inicial de dificultad para cada alumno, por lo tanto, es muy importante que los niños no reciban ninguna ayuda mientras completan la evaluación inicial. La evaluación debe ser completada con el fin de visualizar los reportes de desempeño del estudiante.

- Componentes de Math-Whizz



- ¿Qué es la edad matemática?

Edad Matemática es un término único de Math-Whizz que define el nivel de habilidad matemática de cada estudiante. Se utilizan objetivos estandarizados de equivalencia de grados para calcular la edad matemática con relación al cuarto del año más cercano. Por ejemplo, una edad matemática de 8.25 es equivalente a la habilidad matemática de 8 años y un cuarto.

Las lecciones dentro de Math-Whizz se encuentran organizadas por temas de acuerdo a la edad matemática, por lo tanto, los temas disponibles para cada alumno dependerán de la edad matemática de cada quien.

A continuación se presenta el listado de temas que se integran dentro de Math-Whizz:

- Valor posicional
- Propiedades de los números
- Fracciones
- Decimales
- Comprobación rápida: suma y resta
- Estrategias de cálculo mental: suma y resta
- Escritura de sumas
- Escritura de restas
- Comprobación rápida: multiplicación y división
- Estrategias de cálculo mental: multiplicación y división
- Escritura de multiplicaciones
- Escritura de divisiones
- Problemas contextualizados
- Manejo de la información
- Medidas
- Formas y espacio
- Proporcionalidad y funciones
- Uso de la calculadora
- Enteros, potencias y raíces
- Ecuaciones, fórmulas e identidades
- Secuencias, funciones y gráficas
- Probabilidad

- Organización de las lecciones

Cada lección cubre un objetivo específico de aprendizaje y el alumno podrá progresar al siguiente objetivo de aprendizaje cuando haya completado los tres elementos que componen a una lección:

1. Tutorial

   Cada lección comienza con el componente de enseñanza en donde se explica de forma interactiva y animada el concepto a aprender en la lección.

2. Serie de ejercicios

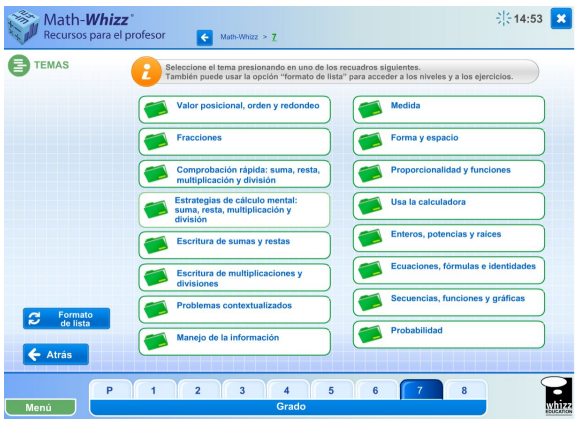
   Una vez que el estudiante ha finalizado con la sección de enseñanza de la lección, procederá a contestar una serie de ejercicios animados relacionados al objetivo de aprendizaje. Si el estudiante contesta de forma incorrecta a la pregunta se le dará una explicación en ese momento para reforzar el concepto; en caso de contestar correctamente a la pregunta recibirá una recompensa animada como motivación a seguir trabajando. El alumno deberá contestar el 70% o más de las preguntas acertadamente con objeto de pasar a la prueba de la lección; en caso de no alcanzar el porcentaje requerido, el tutor cambiará de tema al estudiante y al regresar a este tema el nivel de dificultad será ligeramente más bajo para reforzar el concepto antes de proceder, este modelo define continuamente la experiencia de aprendizaje de los alumnos.

3. Prueba

   Una vez aprobada la serie de ejercicios, el alumno deberá completar una pequeña prueba no animada relacionada con la lección. Esta prueba no cuenta con elementos interactivos ya que se busca el aseguramiento del dominio de la lección. Se aplica el mismo esquema de avance o retroceso que en la serie de ejercicios.

2.1.2 Recursos para el profesor de Math-Whizz

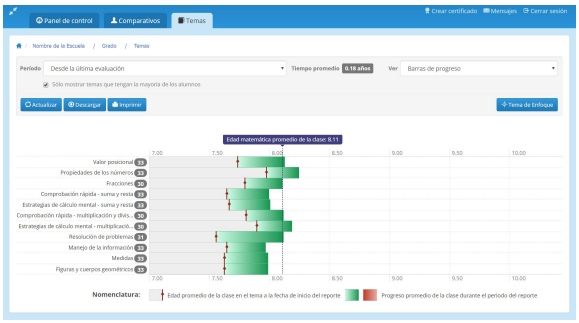
La sección de Recursos para el profesor contiene un aproximado de 1,200 lecciones interactivas para apoyar con la planeación de clase al profesor. Entre las herramientas adicionales de este apartado están una pizarra interactiva, un diccionario y hojas de trabajo. Estas herramientas tienen el objetivo de liberar tiempo de planeación de clase y que el profesor pueda pasar más tiempo atendiendo y revisando las necesidades de aprendizaje individuales de sus alumnos con apoyo de los reportes de desempeño en tiempo real de Math-Whizz.



2.1.3 Herramientas de monitoreo y reportes en tiempo real

Ya que cada estudiante cuenta con su propia cuenta dentro de la escuela, los servicios de reportes en tiempo real se vuelven una poderosa herramienta. Los directores puede ver el desempeño, progreso y niveles de uso de toda la escuela, una clase en particular o estudiantes de forma individual. Estos reportes

también están disponibles a nivel distrito escolar en donde se puede realizar una revisión y comparación entre las escuelas de una forma rápida y sencilla.



2.1.4 Desarrollo profesional

Con objeto de realizar una exitosa implementación en el Estado, se cuenta con Asesores Pedagógicos capacitados en la herramienta de Math-Whizz que realizan el acompañamiento de trabajo con las escuelas y las apoyan con la resolución de dudas y consultas que tienen durante la operación diaria. Se realizan sesiones de capacitación con directores y profesores continuamente tanto en las instalaciones del Centro de Tecnología Umbral como en las instalaciones de las escuelas atendiendo a las necesidades de cada institución.

2.2 What the research says

2.2.1 Learning with digital tools in mathematics education

There exists a wide variety of digital tools that can be used successfully in mathematics education and claim to improve student learning. Recently a number of meta analyses have been undertaken to show the impact that digital tools - and in particular games - are having in education. We define games as computer-based activities with goals, that are interactive, and are rewarding (i.e. give feedback) (Clark, Tanner-Smith & Killinsworth (2014), including virtual learning environments (Bhuiyan & Mahmud, 2015) such as Math-Whizz.

Clark, Tanner-Smith & Killingsworth (2014) undertook a large-scale, systematic review and meta-analysis of 69 study samples providing information on a total of 6,868 unique participants to consider the effect of computer games on learning outcomes. Their analysis shows that digital games conditions were on average more effective than the non-game instructional conditions included in the comparisons (.33 standard deviation improvement). Some of the Whizz lessons and tasks contain playful elements and others involve very structured practice-based questions. Previous EU-funded research in the iTalk2learn project that included Math-Whizz as one components has demonstrated that open-ended (playful or exploratory) activities have the potential to support conceptual understanding learning when combined with structured activities more than just structured activities that focus on procedural knowledge (Rummel et al., 2016).

Clark, Tanner-Smith and Killingsworth (2014) are critical of research that focuses just on the digital resource, and call for acknowledgement that the design of the learning experience within and around the resource is also critically important for students' learning. One core element within Math-Whizz's design is that it provides immediate feedback to students and offers several reporting tools to teachers as presented in Section 2.1. The purpose of feedback is to reduce the gap between students' current understanding or performance and a desired goal (Hattie and Timperley, 2007). In his synthesis of over 500 meta analyses involving more than 20 million students, Hattie (2007) concluded that feedback, where students receive information about a task and how to do it more effectively, results in higher effect sizes than simply praising, rewarding and punishing.

Teachers, students and parents are all instrumental in providing and acting upon feedback effectively. Students can increase effort by undertaking more challenging tasks, developing their ability to detect errors and developing their strategies to complete tasks. Teachers and parents can help students to reach set goals by providing effective learning strategies and feedback. This is why, from the outset, one could envisage a potential impact of the blended learning approach as implemented in Aguascalientes with a substantial element of encouragement for schools, teachers that trickles down to parents and students to use Math-Whizz as homework. Several studies and metastudies have shown the potential of homework (e.g. Maltese, Robery & Fan, 2012) including a recent study with the use of homework with the online homework tool ASSISTMENT (Rochelle et al. 2016)

In their model of feedback to enhance learning, Hattie and Timperley (2007) suggest four levels:

1. Task level: how well tasks are understood / performed
2. Process level: the main process required to understand/perform tasks
3. Self-regulation level: self monitoring, directing, and regulating of actions
4. Self level: personal evaluations and affect (usually positive) about the learner

They argue that self level (level 4) is least effective because it rarely contains task-related information (e.g. "good girl", "That was an intelligent response"), but include it because it is very common in classrooms. Hattie and Timperley (2007) state that the process and self-regulation levels (levels 2 and 3) are powerful for deep processing and mastery of tasks. Task level (level 1) is powerful when the task information is used to improve strategy processing or to enhance self-regulation, which, Hattie and Timperley argue, rarely happens.

The duration and frequency of using digital tools is also important to learning outcomes, and can be adjusted as necessary in the learning design. Sitzmann (2011) found that students having unlimited access to a game, demonstrated significantly higher learning gains (relative to the non-game control conditions) than comparisons in which the students had limited access to the game. Similarly, Wouters et al. (2013) found that comparisons in which participants engaged with a game for only one session did not result in significant learning gains relative to the non-game control condition, whereas comparisons in which participants interacted with the game for more than one session resulted in significant learning gains relative to participants in the non-game condition. This is echoed by Higgins, Xiao & Katsipataki(2012) who found in their meta-analysis of learning technologies that technology can be particularly impactful as a short but focused intervention to improve learning, particularly when there is regular and frequent use (about three times a week) over the course of 5 - 10 weeks.

2.2.2 Teacher training for learning with digital tools in mathematics education

International studies of teacher professional development show that the implementation of learning technologies is one of the most challenging aspects of modern teaching (see, for example, Villegas-Reimers, 2003), and that despite often significant training being undertaken, many teachers often fail to have integrate technology effectively in their teaching (Daly, Pachler & Pellitier, 2009; Higgins et al, 2012; NCETM, 2010). One reason for this is that the focus on technology-based skills rather than pedagogy (Daly et al, 2009; Higgins et al, 2012; Clark-Wilson, Oldknow & Sutherland, 2011). Teachers and students need support in getting the best from technology in terms of their learning and that *ongoing* professional development and support to evaluate the impact on learning is likely to be required (Higgins et al, 2012).

The UK's National Centre for Excellence in the Teaching of Mathematics (NCETM, 2010) highlights the need for improved access to teacher professional development in learning technologies. In addition to considering teachers' underlying attitudes, perceptions and technological skill development, the report concludes that collaborative and continued involvement with groups of peers, supported by appropriate expertise has proved successful in sustaining uses of digital technologies. Clark-Wilson, Oldknow & Sutherland (2011) concur. They explain how the process of adopting digital technologies requires teachers to take risks in their teaching, supported by appropriate technical, didactical and mathematical expertise. They recommend that change has to be supported by both school leaders and accompanied by *sustained* professional development opportunities for teachers.

Higgins et al's (2012) meta-analysis of the impact of digital technology on learning found that teacher training was a core element. They conclude that the most successful learning gains were achieved when there is at least a full day's training or ongoing professional inquiry-based approaches to support the introduction of a new technology.

2.2.1 Evaluation expectations

Our own previous research and that of others including previous evaluations of the Math-Whizz environment predict that, under the right conditions, the use of Math-Whizz has the potential to support students' learning. For example, recently Roschelle et al. conducted a large randomized field trial that combining the online homework tool ASSISTments with teacher training significantly increased student scores on an end-of-the-year standardized mathematics assessment as compared with a control group that continued with existing homework practices. Similarly, our previous research has indicated that the use of educational technology is having an effect on student attitude and affective dispositions towards mathematics (Mavrikis et al. 2016). However, we are realistic in that we know that it is hard to demonstrate learning gains particularly in short time and to isolate the effects of a complex socio technical environment to just the technology alone. For example, a much larger study in the USA of the Cognitive Tutors (Pane et al. 2014) found only limited evidence of the effectiveness of these tutors in a relative uncontrolled implementation over conventional teaching.

Moreover, previous investigations in the context of large-scale educational assessments and other studies come to ambivalent findings with regard to computer use and performance of students (OECD 2015, pp. 154). In fact, evidence from countries participating in PISA 2012 suggests that after accounting for their socio-economic background, students who did not use computers in their mathematics lessons, outperformed those that did. In addition, the study suggests that very high engagement in certain activities on computers might supersede learning opportunities of higher effectivity.

Whether positive relationships between usage and outcome that emerge in the literature, depends on what types of computer usage and outcomes were investigated (OECD 2015, pp. 154 - 163). While the mere provision of technological resources does not appear to have a positive relationship in the study, the findings suggest some amounts of usage to be more optimal, i.e. lower achievement to be related to average usage amounts far below or above the OECD average (OECD 2015, pp. 146).

3. Metodología

Este informe se basa en un enfoque de métodos mixtos, incluye información cualitativa y cuantitativa a partir de distintas entidades involucradas. La parte cualitativa se basa en la observación del trabajo en las escuelas, entrevistas con profesores, padres y los mismos alumnos, así como la observación de sesiones de capacitación a maestros. El análisis cuantitativo se basa en un **estudio cuasi-experimental**, particularmente en un diseño de grupo de control no equivalente (Fife-Schaw, 2006). Las dos condiciones se especifican de la siguiente forma

(Condición Math-Whizz) Las clases en las escuelas donde se implementó el sistema de Math-Whizz y forman parte del programa de capacitación a maestros.

(Condición no-usuarios) Selección de escuelas donde Math-Whizz y el programa de capacitación a maestros no había iniciado al momento de la evaluación.

Utilizamos el software R para el análisis de datos (R Core Team, 2016; R Studio Team, 2015) y paquetes específicos como se menciona en el Apéndice 4

Aunque hemos considerado los datos de toda la implementación en Aguascalientes, los **participantes** para el análisis del presente informe se consideran sólo de 4° grado y estudiantes de educación primaria comprendidos entre los 8 y 9 años de edad provenientes de las escuelas indicadas en el Apéndice (elegidas de zonas rurales, suburbanas y áreas dentro de la ciudad). 439 maestros participaron en el cuestionario en línea.

Las limitaciones del estudio son presentadas en la Sección 6.

3. Methodology

This report relies on a mixed methods approach that includes both qualitative and quantitative data from a range of stakeholders. The qualitative part relies on school observations in a range of schools, interviews with teachers, parents and students themselves as well as observation of teacher training sessions. The quantitative analysis relies on a **quasi- experimental design**, particularly a non-equivalent control group study (Fife-Schaw, 2006). The two conditions were specified as follows

(Math-Whizz condition) The classes within schools that implementing the Math-Whizz software and taking part in the teacher training program.

(non-users condition) Selected range of schools that where the Math-Whizz implementation and associated teacher training had not started at the time of the evaluation.

We use the R software for data analysis (R Core Team, 2016; R Studio Team, 2015) and specific packages as referred to in Appendix 4

Although we considered the data of the whole implementation throughout Aguascalientes, the **participants** for the analysis reported here are considered to be only the Year 4 and primary school students aged between 8 and 9 years old from the schools in the Appendix (picked from rural, suburban, and inner-city areas). 439 teachers took part in the online questionnaire.

The limitations of the study are discussed in Section 6.

4. Medidas

Para medir los niveles de aprovechamiento de los estudiantes obtuvimos datos del estado de Aguascalientes de la prueba nacional PLANEA en lectura y matemática que se realizó en septiembre.

De acuerdo al cronograma de actividades, no se contaba con tiempo suficiente para utilizar una prueba estatal para medir el nivel de aprovechamiento al final de la intervención. Debido a otros factores, incluyendo las evaluaciones bimestrales, contamos con una ventana muy limitada para aplicar una prueba distinta. Asi que con el debido permiso, se seleccionaron 10 preguntas de PLANEA de los temas "Números y sistemas de numeración" y "Problemas aditivos" ya que son temas de trabajo dentro de Math-Whizz. Los estudiantes recibieron un punto por cada pregunta contestada correctamente y consecuentemente una calificación agregada. La consistencia interna para el estudio de Oct fue Cronbach α=.76 y para el de Sep α=0.85. Whizz Education no realizó ningún cambio en el algoritmo aunque conocían con antelación los temas a evaluar.

Monitoreamos el uso de Math-Whizz a través del sistema que registra cada vez que el estudiante inicia sesión dentro y fuera del aula. Aparte de la información de uso, el sistema también registra el progreso de la "Edad Matemática", un término de Math-Whizz que define el nivel de habilidad matemática de cada estudiante.

Además, se observaron cinco aulas en distintas escuelas y perfiles utilizando Math-Whizz, se entrevistó a 10 docentes y cinco padres de familia. También se interactuó y pidió la opinión de más de 20 niños (cuatro de cada escuela) durante su interacción y retrospectivamente en grupos de enfoque.

4. Measures

To measure students' levels of achievement we obtained data from the state of Aguascalientes from the national PLANEA students' achievement test for reading and mathematics that took place in Sep.

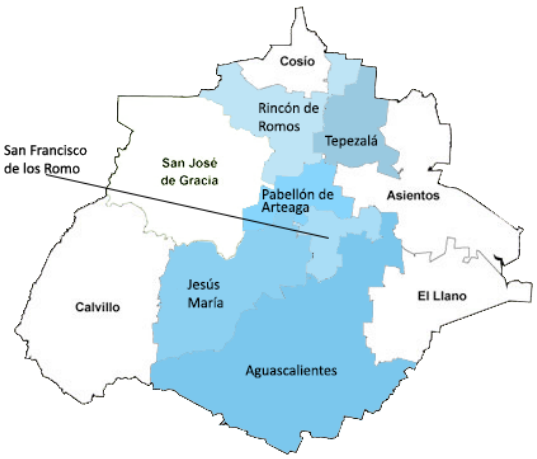
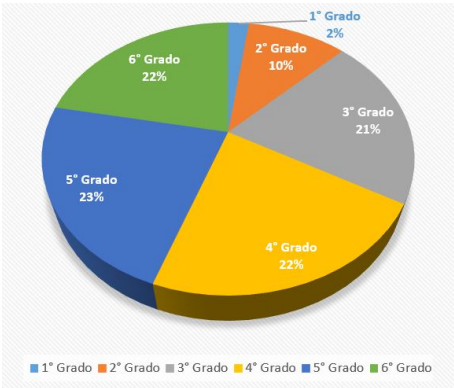
The timeframe of our evaluation meant that we could not rely on a state-administered test to measure students' levels of achievement at the end of the intervention. Due to other factors, including the bi-monthly tests that students are taking, we also had limited time window to apply a different test. As such, we modified PLANEA with permission and selected 10 questions from the 'numbers and counting" and "addition" problems as this is also the topics that were covered by Math-Whizz. The students received one point for each correctly answered question and consequently an aggregated score. Internal consistency for the Oct test was Cronbach α=.76 and for Sep α=0.85. Note that Whizz Education did not make any changes to the delivery algorithm neither were they aware of the choice of PLANEA in advance.

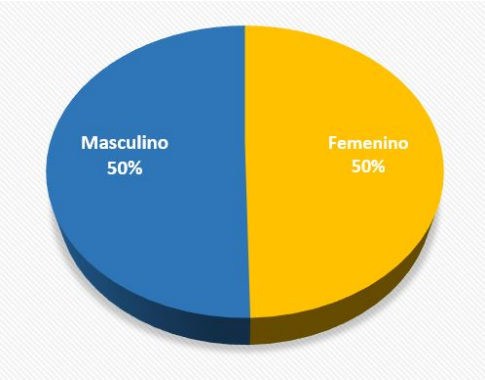
We tracked usage of Math-Whizz through the system that logs data for each student whenever whether in or outside the classroom. Apart from usage data the system also logs the progress the so-called "Math-Age", a Math-Whizz term that defines the level of math skills of each student.

In addition, we observed five schools/classrooms from different backgrounds and interviewed 10 teachers and five parents. We also interacted and asked the opinion 20 children (four from each school) during their interaction with Math-Whizz and retrospectively in focus groups.

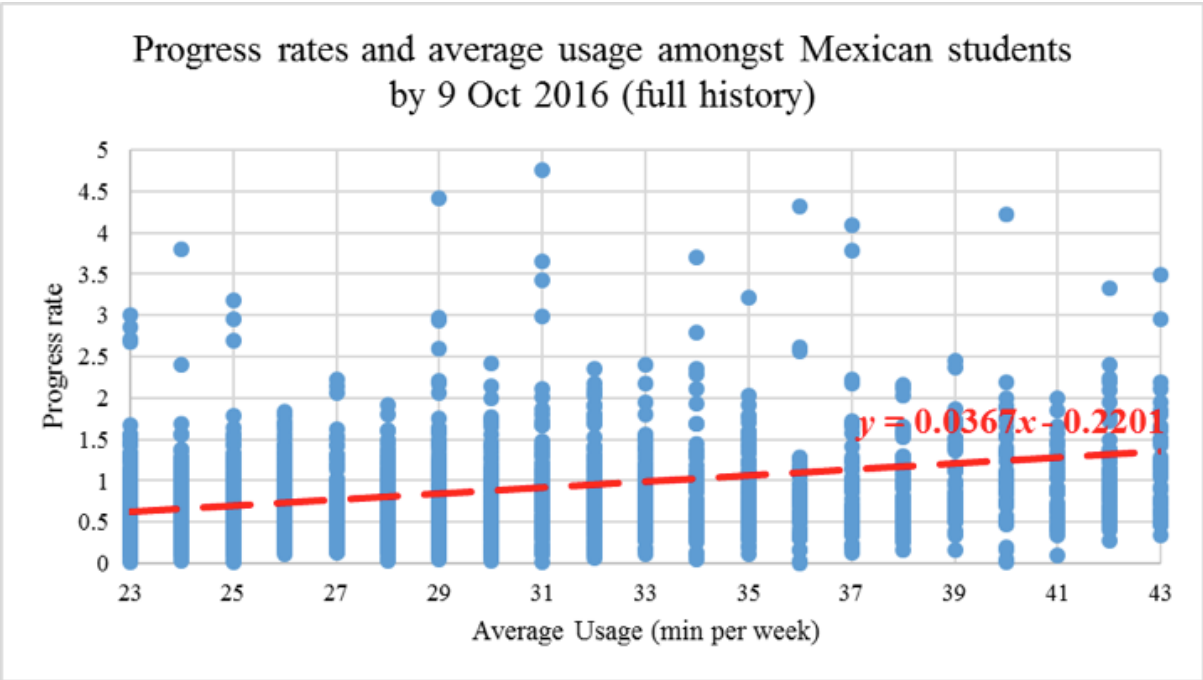
5. Resultados

5.1 Sobre la implementación de Math-Whizz en Aguascalientes

Número total de escuelas en Math-Whizz	132
Municipios en donde se utiliza Math-Whizz 	Aguascalientes Jesús María Pabellón de Arteaga Rincón de Romos San Francisco de los Romo Tepezalá
Localidades en donde se utiliza Math-Whizz	Aguascalientes, Jesús Gómez Portugal, Jesús María, Pabellón de Arteaga, Pocitos Puertecito de la Virgen, Rincon de Romos San Francisco de los Romo, Tepezala
Estudiantes dentro de Math-Whizz	45,673
Estudiantes con evaluación inicial	28,018
Tiempo total de tutoría recibida a través de Math-Whizz	155,264 horas
Distribución aproximada de estudiantes por grado académico: 1° Grado - 2% 2° Grado - 10% 3° Grado - 21% 4° Grado - 22% 5° Grado - 23% 6° Grado - 22%	

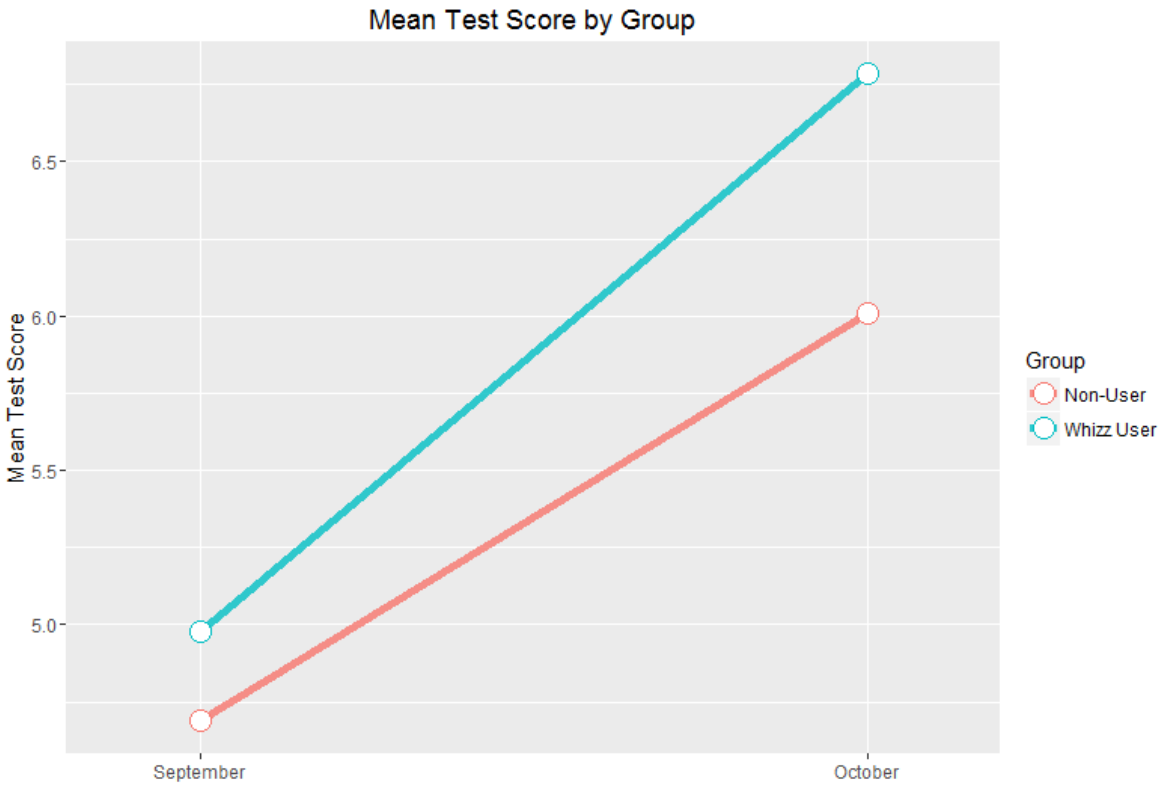
Distribución de estudiantes por género: Masculino - 50% Femenino - 50%	

Analizando todo el cohorte, y seleccionando el subconjunto de alumnos cuyo uso promedio está entre 23 y 43 minutos por semana desde el inicio de la implementación y hasta el 9 de octubre de 2016, en promedio, una regresión lineal simple muestra que un estudiante en el cohorte de Aguascalientes necesita utilizar Math-Whizz aproximadamente 33 minutos por semana para lograr una tasa de progreso de uno. Una regresión lineal en aquellos entre 35 y 55 minutos por semana, demuestra que un estudiante debe utilizar Math-Whizz cerca de 44 minutos por semana para lograr una tasa de progreso de 1.5.



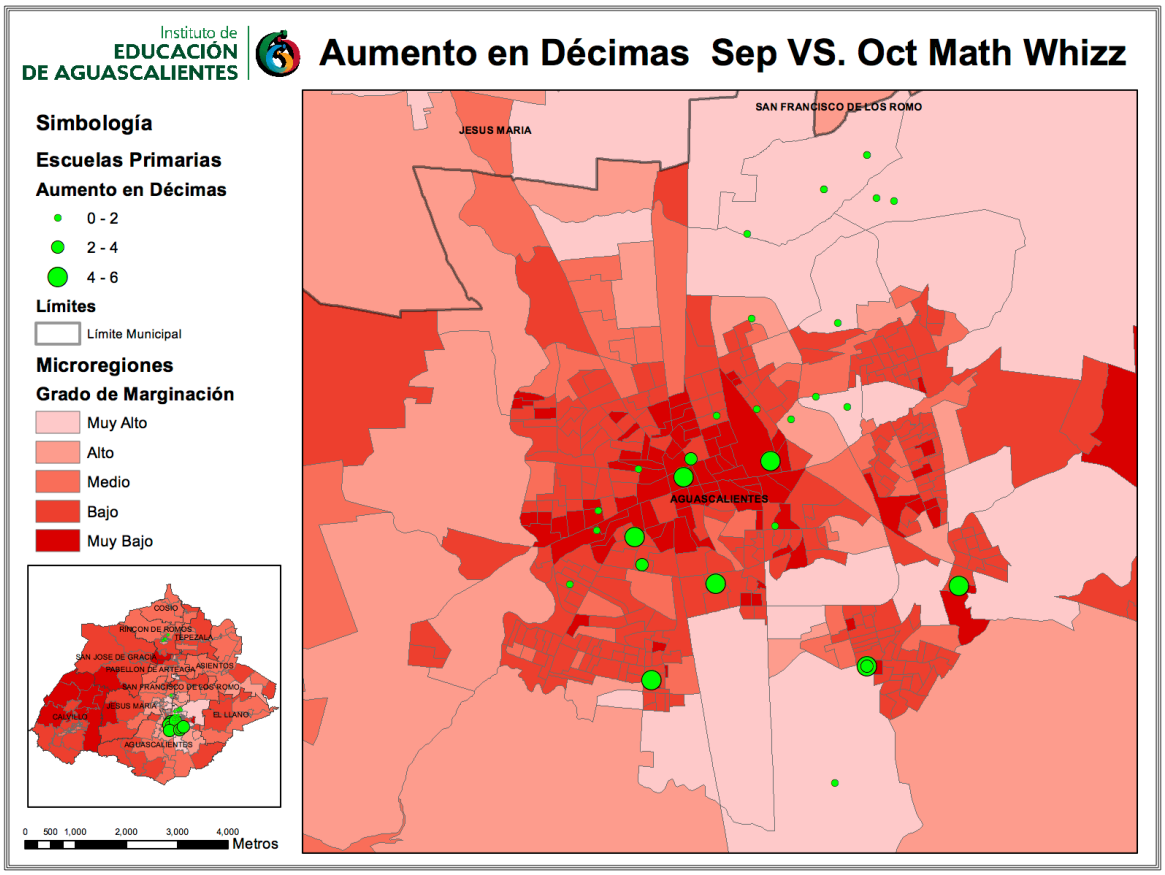
El análisis estadístico detallado presentado a continuación (sección 5.3 en Inglés) del cohorte de 4° Grado, lleva a la conclusión de que el grupo que utilizó Math-Whizz tuvo un mayor progreso, entre septiembre y octubre, en comparación con el grupo de control. Tomando en cuenta las diferencias en los resultados de la prueba de septiembre, el logro previsto para los alumnos usuarios de Math-Whizz en la muestra es 0.659 puntos superior en una escala de 10 que la de los no-usuarios de Math-Whizz, mientras que si dividimos a los usuarios de Math-Whizz en grupos más pequeños, dividiendo a quienes utilizan el sistema más de 45 minutos en promedio, entre 34-44 minutos de uso promedio, 5-33 minutos de uso promedio y a menos de cinco minutos de uso promedio, las relaciones se diferencian aún más. La diferencia entre los no usuarios y el grupo de menos de cinco minutos de uso promedio sigue siendo insignificante (indicando la validez del análisis) y el grupo de uso alta está asociado a un progreso adicional de 1.26 puntos. **Estos resultados deben ser leídos con cautela en el contexto de todo el análisis presentado en la sección 5.3 y las limitaciones que se presentan en la sección 6.**

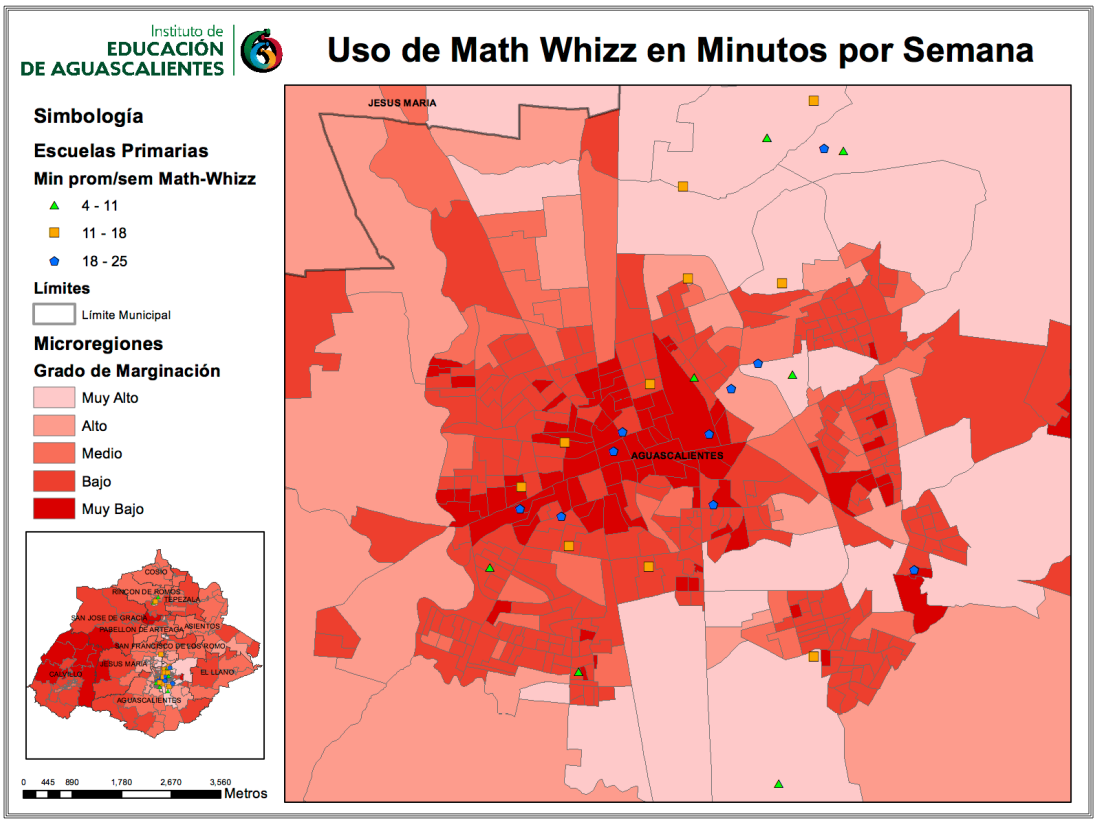
Gráfica 5.1

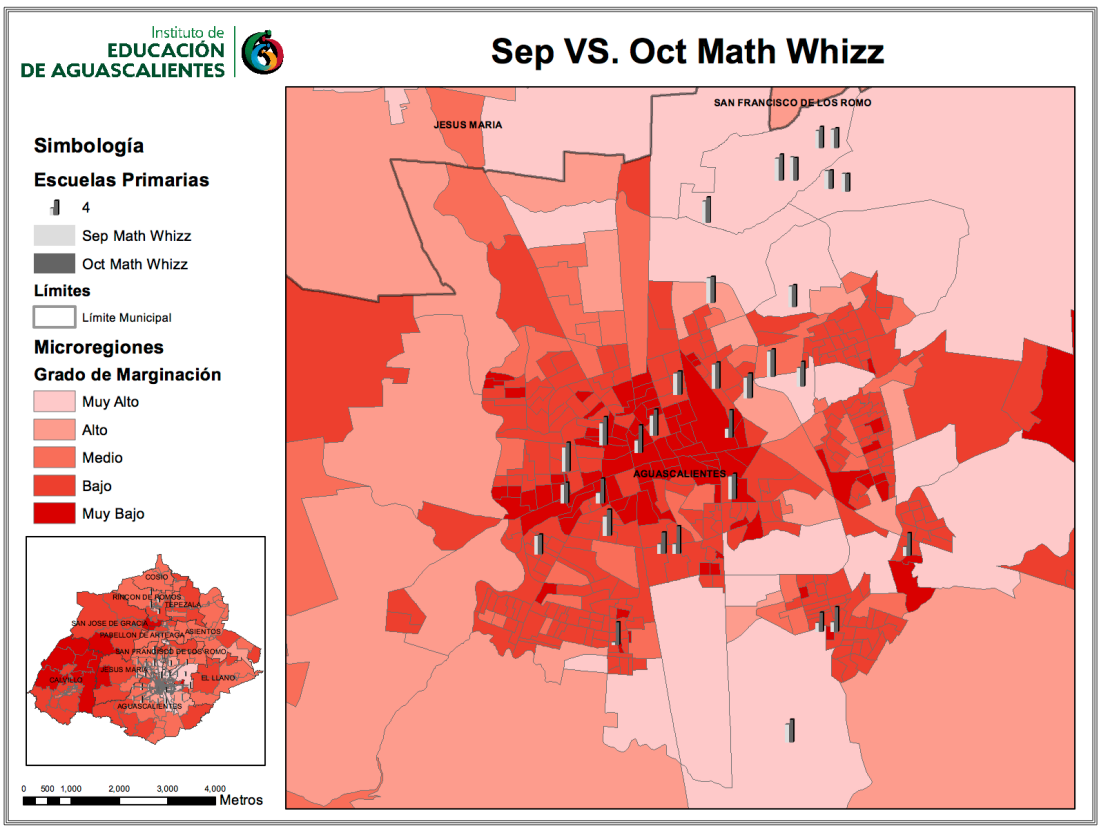


5.2 Representación visual

Centrándose sólo en el grupo que utilizó Math-Whizz en 4° Grado, los siguientes mapas proporcionan una representación visual para aquellas escuelas que teníamos información demográfica y resultados estadísticos validados. Podemos observar una interesante distribución del uso y las diferencias de aprendizaje que se extiende a todas las microrregiones (de muy alto a muy bajo). El análisis estadístico adicional quedó fuera del alcance de este trabajo, pero los resultados indican una implementación robusta y que valdrá la pena recopilar sistemáticamente estos datos a partir de ahora para tener en cuenta en una implementación más amplia.




Instituto de
EDUCACIÓN
DE AGUASCALIENTES
Uso de Math Whizz en Minutos por Semana
Simbología
Escuelas Primarias
Min prom/sem Math-Whizz
4 - 11
11 - 18
18 - 25
Límites
Límite Municipal
Microregiones
Grado de Marginación
Muy Alto
Alto
Medio
Bajo
Muy Bajo
COSIO
RINCON DE ROMOS
TEPEZALA
SAN JOSE DE GRACIA
ASIENTOS
PABELLON DE ARTEAGA
SAN FRANCISCO DE LOS ROMO
JESUS MARIA
CALVILLO
EL LLANO
AGUASCALIENTES
0 445 890 1,780 2,670 3,560
Metros
JESUS MARIA
AGUASCALIENTES


Instituto de
EDUCACIÓN
DE AGUASCALIENTES
Sep VS. Oct Math Whizz
Simbología
Escuelas Primarias
4
Sep Math Whizz
Oct Math Whizz
Límites
Límite Municipal
Microregiones
Grado de Marginación
Muy Alto
Alto
Medio
Bajo
Muy Bajo
COSIO
RINCON DE ROMOS
TEPEZALA
SAN JOSE DE GRACIA
ASIENTOS
PABELLON DE ARTEAGA
SAN FRANCISCO DE LOS ROMO
JESUS MARIA
CALVILLO
EL LLANO
AGUASCALIENTES
0 500 1,000 2,000 3,000 4,000
Metros
JESUS MARIA
SAN FRANCISCO DE LOS ROMO
AGUASCALIENTES

5.3 Detailed Statistical analysis

5.3.1 Learning

As suggested by the graph 5.1, the mean test score of the Whizz group was a bit higher than that of the non-Whizz group. A supplementary regression analysis did not show significant differences in the mean September test scores for the Whizz group and the non-Whizz group ($\beta = 0.289$, $p > 0.05$). However, in terms of mean scores, graph 5.1 shows that the Whizz group appears to have made more progress between September and October. The results in table bleow were estimated by linear regression model that resembles the ANCOVA method for measuring change in time (Wright & London 2009, pp. 49) and using cluster robust and heteroscedasticity-consistent standard errors (Graham, Arai & Hagströmer 2016, e.g. Snijders & Bosker 2012, pp. 197). Accounting for the differences in students' test achievement in September, the predicted achievement score for Math-Whizz users in the sample is 0.659 points higher than that of non-Whizz (Model 1, $\beta = 0.659$, $p < 0.05$) in a scale of 0 to 10. According to model 1 and if the change is not due to unobserved variables, this significant difference in progress might be associated with the use of the software. The entire model 1 explains around 18.6 % of the differences in posttest scores. The corresponding multilevel model suggests very similar values ($\beta = 0.6$, $p < 0.05$). From the standarized coefficient of the multilevel model 1 ($\beta = 0.1$, $p < 0.05$), we calculate an effect size of about $d=0.22$[1] a value that is commensurate with other studies in the area (e.g. Roschelle et al. 2016).

From model 2 (Table 5.1), we also conclude that the positive relationship between Math-Whizz usage and students' progress in the sample was not dependent on students' previous achievement, i.e. the interaction effect between the September score and the Math-Whizz group variable was not significant (Model 2, $\beta = -0.086$, $p > 0.05$).

In a similar analysis, the Whizz group was split into those users that used the software more and those who have used the software less than five minutes per week. For the latter group, one might argue that it is rather questionable whether these users could be expected to benefit from the intervention. Controlling for prior achievement in September, there is no significant difference (Model 3, $\beta = 0.081$, $p > 0.05$) in the progress of those participants who used the system less than five minutes on average, compared to the non-user group. For the same analysis, the usage group of above five minutes is associated to an additional progress of 0.87 points (Model 3, $\beta = 0.867$, $p < 0.01$) in the test compared to the non-user group. The similarity between the non-user and the less-than-five-minutes user group, becomes visual in figure 5.2 below.

The similarity between the group that did not participate in the intervention and the one that was meant to participate, but whose degree of participation is doubtful, is certainly a very interesting finding. However, while this could be interpreted in favour of the intervention, it should also be beared in mind that the users self-select into their group of usage intensity. This self-selection is likely to be dependent on student characteristics itself, such as less motivated and proficient learners using the system less (EEF, 2016). Future studies of higher complexity should attempt to develop approaches to control for such confounding effects.

[1] http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD22.php

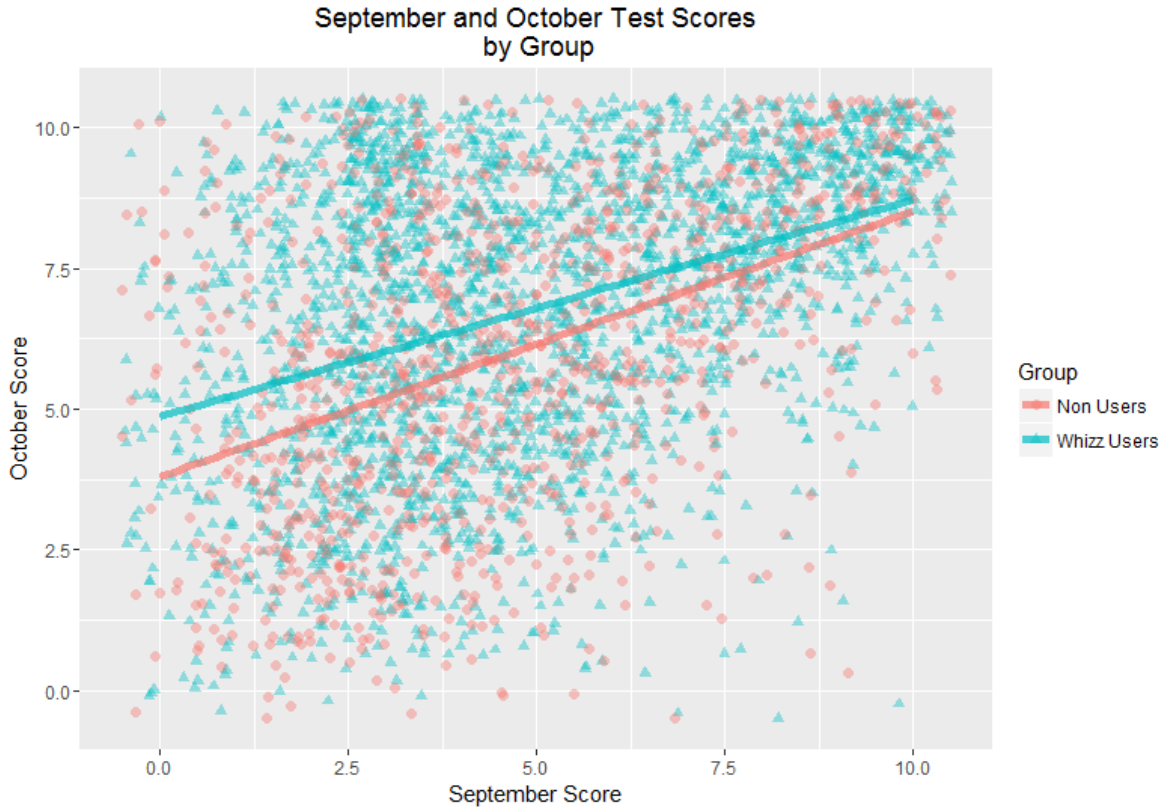
The above findings are notwithstanding the earlier described finding that the share of participants who have met the recommendation of more than 33 minutes of usage per week is very small. The positive and significant coefficient of more than five minutes average usage remains indicative of a positive effect (considering limitations of the current study described in the next section). Moreover, when splitting up the group of users further into the smaller groups of those who used the system more than 45 minutes on average, 34-44 minutes on average, 5-33 minutes on average and less than five minutes on average, the relationships differentiate further: The difference between the non-users and the group of less than five minutes average usage remains insignificant (Model 4, $\beta = 0.081$, $p > 0.05$).

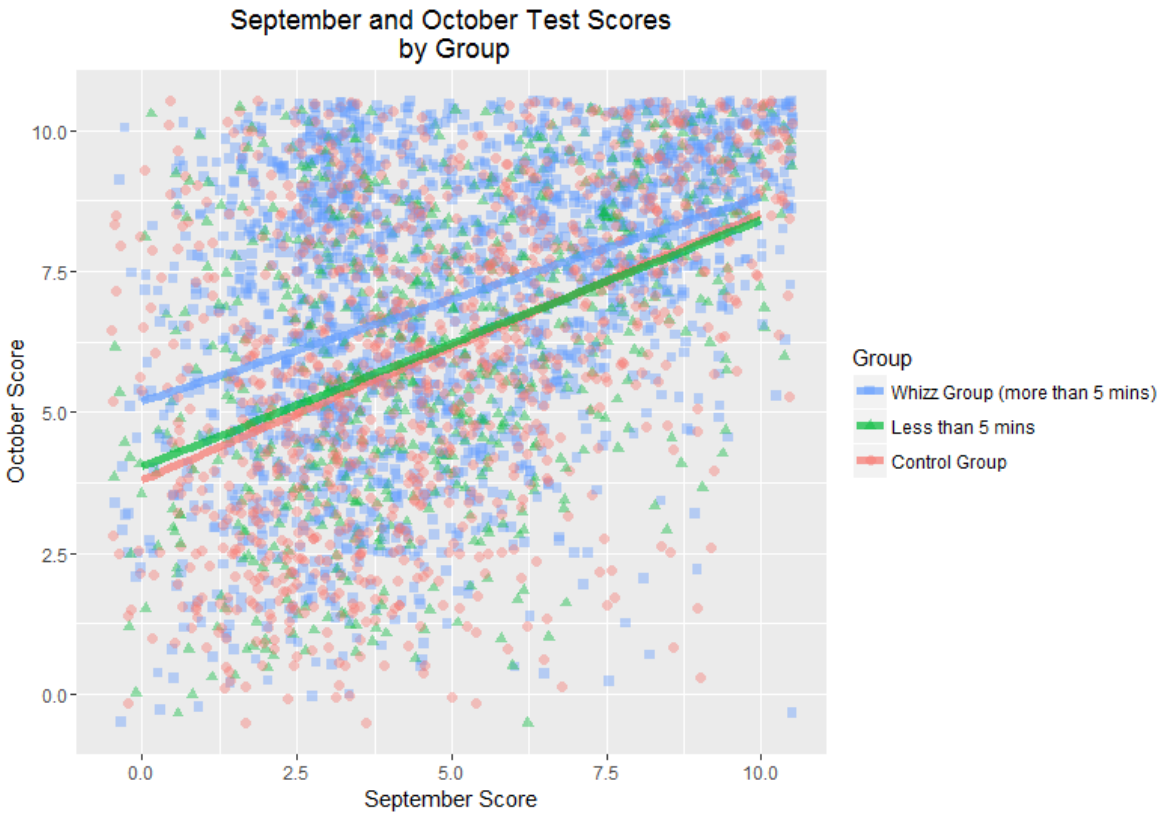
Table 5.1 Linear regression models with cluster-robust standard errors

Linear Regression with Cluster-Robust Standard Errors

	Dependent Variable			
	October Score			
	Model 1	Model 2	Model 3	Model 4
Intercept	4.079*** (0.312)	3.797*** (0.496)	4.092*** (0.309)	4.100*** (0.309)
September Score	0.412*** (0.034)	0.472*** (0.067)	0.409*** (0.033)	0.407*** (0.033)
Whizz User Group	0.659* (0.260)	1.070 (0.582)		
Whizz User*September Score		-0.086 (0.077)		
Less than five minutes P/W			0.081 (0.254)	
More than five minutes P/W			0.867** (0.270)	
High usage				1.255*** (0.319)
Min. recommended usage				0.969** (0.338)
Low usage				0.831** (0.276)
Very low usage				0.081 (0.254)
Observations	3702	3702	3702	3702
Multiple R-squared	0.186	0.187	0.198	0.199

Note: *p<0.05, **p<0.01, ***p<0.001


September and October Test Scores
by Group
October Score
September Score
0.0
2.5
5.0
7.5
10.0
Group
Non Users
Whizz Users


September and October Test Scores
by Group
October Score
September Score
0.0
2.5
5.0
7.5
10.0
Group
Whizz Group (more than 5 mins)
Less than 5 mins
Control Group

5.3.2 Students' attitude and sentiment towards mathematics

While it is difficult to derive conclusions from sort questionnaires about emotional constructs, the results appear in the Appendix and suggest that further research is needed in the area. A logistic regression in the case of effort suggest that the odds for a Whizz user to report "strongly agree/agree" over the odds of non users to report that is exp(0.632)= 1.88137. The odds for Whizz users are 88,1% higher than the odds for non-user.

Logistic Regression with Cluster-Robust Standard Errors

	Dependent Variable				
	Happy	Solving New Problems	Effort	Afraid	Dedicating Time
Intercept	0.590***	1.273***	1.074***	0.434***	1.438***
	(0.126)	(0.131)	(0.126)	(0.123)	(0.156)
Whizz User Group	0.079	-0.029	0.285*	0.212*	0.632***
	(0.118)	(0.124)	(0.118)	(0.106)	(0.150)
September Score	0.026	0.076***	0.100***	0.090***	0.165***
	(0.017)	(0.022)	(0.017)	(0.018)	(0.027)
Observations	3678	3641	3641	3647	3604
Note:					*p<0.05; **p<0.01; ***p<0.001

5.4 Classroom observations, teacher and parent interviews

En general, podemos decir que hay un apoyo abrumador a Math-Whizz y apreciación por su implementación. Las entrevistas con los padres destacaron el papel particular que potencialmente pueden desempeñar con tal implementación. Todos los padres entrevistados se refirieron en su observación sobre la transformación de sus hijos tanto en cuanto al comportamiento como a los resultados del aprendizaje. Los padres refirieron la paciencia, la diligencia, la reducción del miedo a las matemáticas e incluso una mayor motivación hacia otros deberes. Algunas citas características se proporcionan en el Apéndice 2.

Dependiendo de sus circunstancias, los padres entrevistados siguen diferentes estrategias para apoyar a sus hijos a acceder a Math-Whizz, incluyendo comprar nuevas computadoras, prestarles sus propios dispositivos (por ejemplo, tabletas) o llevarlos a la biblioteca. Hemos observado que los antecedentes matemáticos de los padres pueden estar determinando su papel en el hogar. La mayoría están ayudando a sus hijos con dudas que tienen y parecen entender la importancia de no intervenir ya que el sistema utiliza los datos para adaptarse (esto se resalta en el entrenamiento de los profesores y los padres). Uno de los padres reflexionó sobre la oportunidad que esto le proporciona para estar más comprometida en las actividades para que pueda aprender o simplemente recordar algunas de las matemáticas involucradas en las tareas. Esto le da la capacidad de ayudar a su hija más. Dos de los padres informaron sobre los beneficios positivos de esta oportunidad de conectarse mejor con su hijo y sentirse parte de lo que está pasando en la escuela.

La observación de las aulas y mientras conversamos con los maestros, también estuvieron abrumadoramente a favor de la implementación y agradecidos del apoyo que

In general, all the interviews and observations we conducted suggest that teachers have an overwhelming support for Math-Whizz and appreciation to the overall implementation. The parent interviews highlighted the particular role that they can potentially play in such an implementation. All parents interviewed referred to their observation on their children transformation both regarding behaviour and learning outcomes. The parents referred to patience, diligence, reduction of fear of mathematics and even increased motivation towards other homework. Some characteristic quotes are provided in Appendix 2.

Depending on their circumstances the parents interviewed follow different strategies to support their children access Math-Whizz including buying new computers, lending them their own devices (e.g. tablets) or taking them to the library. We observed that the parent's mathematics background may be determining their role at home. Most are helping their children with particular doubts they have and seem to understand the importance of not intervening as the system uses the data to adapt (this is stressed at the training for teachers and parents). One parent reflected on the opportunity this provides for her to be more engaged in the activities so as she can learn or simply remember some of the mathematics involved in the tasks. This gives her the ability to help her daughter more. Two of the parents reported on the positive benefits of this opportunity to connect better with their child and feel part of the what is going on at school.

Observing the classrooms and discussing with the teachers, they were also overwhelmingly in favour of the implementation and grateful to the support that is provided to them through the state. They reflected on the role of the trainers who frequently visit them and analyze

les brindan a través del estado. Reflexionaron sobre el papel de los entrenadores que frecuentemente los visitan y analizan con ellos los informes que el programa Math-Whizz proporciona, para ver qué necesidades están siendo presentadas por cada escuela. Más allá del apoyo técnico, dos profesores reflexionaron en particular sobre las oportunidades de capacitación pedagógica que esto ofrece a medida que los capacitadores los apoyan más allá de los informes con consejos sobre cómo integrar a Math-Whizz en su aula. Investigamos más este hallazgo en la siguiente sección.

En general, podemos resumir los hallazgos refiriéndonos al hecho de que tanto los maestros aprendiz como los maestros experimentados identifican potenciales positivos similares. Esto señala a la calidad de la formación y el diseño de las características Math-Whizz que se aprecian fácilmente. Además, hay un aumento cuantitativo y cualitativo en el tipo de respuestas después de que los maestros han experimentado la plataforma en el aula en que proporcionan respuestas que se relacionan con el pensamiento matemático y el potencial de motivación del estudiante. Esto apunta a una mayor comprensión por parte de los maestros de la implementación posterior a Math-Whizz. En cuanto a las limitaciones, los profesores se refieren principalmente a los desafíos de infraestructura como el acceso a Internet y a la computadora, pero algunas respuestas apuntan también a la falta de integración profunda de la herramienta de monitoreo en la práctica cotidiana, sugiriendo que a medida que la implementación gane impulso, herramientas para identificar las fortalezas y debilidades de los estudiantes pueden ser el siguiente paso positivo.

with them the reports that the Math-Whizz program provides, to see what needs are being presented by each school. Beyond the technical support two teachers reflected particularly on the pedagogical training opportunities that this provides as the trainers support them further to the reports with advice on how to integrate Math-Whizz in their classroom. We elaborate upon this finding in the next section, using the questionnaire data.

In general, we can summarise the findings by referring to the fact that both trainee and experienced Whizz teachers identify similar potential strengths. This points to the quality of the training and the design of Math-Whizz features that are easily appreciated. In addition, there is both a quantitative and qualitative increase in the type of responses after teachers have experienced the platform in the classroom in that they provide answers that relate to mathematical thinking and the potential for student motivation. This point to teachers' increased understanding of Math-Whizz post implementation. In terms of limitations, the teachers refer mostly to infrastructure challenges such as internet and computer access but some responses point also to the lack of deep integration of the monitoring tool in everyday practice, suggesting that as the implementation gains momentum, a shift towards using the tools to identifying students' strengths and weaknesses might be a positive next step.

[Confidencial – solo para IEA – No publicar en línea]

5.5 Teacher Questionnaire findings

419 teachers responded to the online questionnaire. In the questionnaire, teachers were asked open-ended questions about their first impressions of Math-Whizz, how they used the monitoring tool, and what they found easy and difficult to use in the reporting system. We have used the data provided by the teachers to evaluate their use of Math-Whizz. The detailed results of this questionnaire appear in Appendix 3.

In general, teachers were very positive indeed about the introduction of Math-Whizz. The comments below are indicative of the types of positive responses received:

- In the time we have used the program, as a group, we have found that students have improved their enthusiasm in the contents of mathematics and Math-Whizz motivates them to continue working.
- I hope to continue to innovate as this program is a great support to increase intelligence and learning.
- We are very pleased with the program, and will continue working with the same and experience with the school year
- I congratulate those who proposed this program because very up-to-date and interesting for all students.
- Congratulations. A new alternative way to work and improve practice of students to develop skills and knowledge.
- Personally the program makes an excellent support tool for learning the mathematics.
- Congratulations, an excellent project.
- I'm interested in learning that the program will to optimize the mathematics work with my students.
- Math-Whizz I regard as an excellent resource for teachers, students and parents.
- Good program supports both to students and teachers.
- It is excellent, funny, innovative and current.
- The program is excellent, students motivates them a lot and serves to put in practice their math skills and develop skills like mental calculations and of course enables students to access the use of technological devices that most of them have at school.
- It has been a great experience working with this tool.
- I consider it a novel resource and is motivating for students, as students continually ask for work activities
- Math-Whizz has helped both me and my students to better understand the mathematics. Thank you.

However, there were also a number of concerns from teachers about infrastructure issues, where there was a poor internet connection in schools or homes.

- It is an excellent program that motivates students to get a taste for mathematics. I hope better use can be made of it with improved internet connections.
- It is a problem not having a special cabinet to store and recharge tablets.
- We do not have enough computers for students to work simultaneously and serve two groups with different activities.
- Students do not have access to the program at home, using only time in school.
- It requires a good computer room
- Students that have internet at home to keep practicing at home

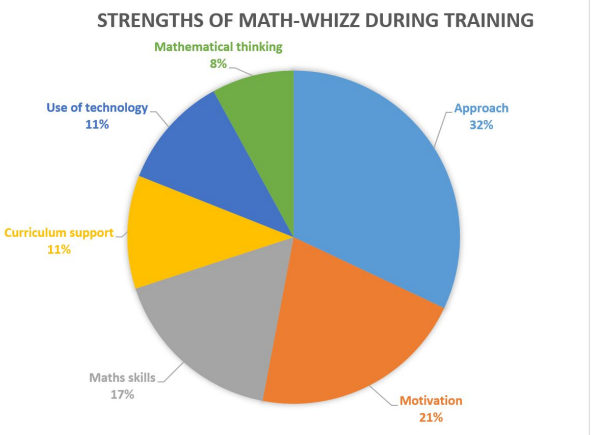
- Where are the computers for fifth grade that were promised? Computers that the school has the parents acquired with much effort.
- We were unable to start the program because of the low speed of the internet.

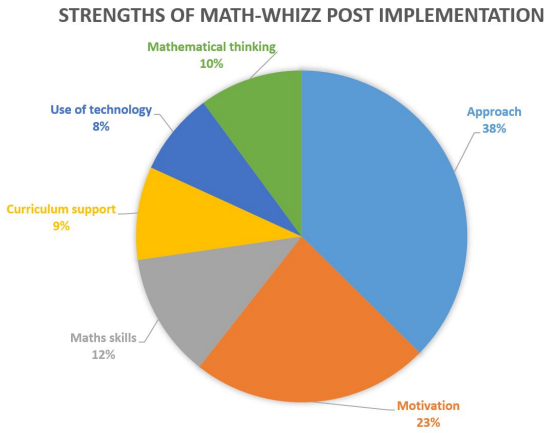
Below these broad headlines are explored in more detail.

Strengths of Math-Whizz

The teachers were asked what they perceived as the strengths of Math-Whizz at two phases of implementation: during their training and now they have used it in their teaching. 117 teachers identified aspects they considered strengths prior to use (during the training) and 170 teachers identified aspects after they implemented Math-Whizz in their classrooms. This increase in the number of teachers who were able to respond about the strengths of Math-Whizz after they began to implement it may reflect the teacher's' ability to engage more fully with the program after they had become more familiar with it.

Strengths	***During training***	***After implementation***
Approach Appreciation of the individualised, varied and interactive approach of the platform, taking into account the needs of students regardless of their level of attainment, and supports their progress.	32%	37%
Motivation Motivating impact of Math-Whizz because it is attractive, playful, eye-catching, dynamic and so on.	21%	23%
Maths skills Students learn basic maths skills.	17%	12%
Curriculum support Math-Whizz is a good support for teaching the mathematics curriculum.	11%	9%
Use of technology Using technology for technology's sake; acknowledging that technological skills are learnt within mathematics lessons.	11%	8%
Mathematical thinking Developing students' mathematical thinking, more specifically logical reasoning or logical thinking.	8%	10%



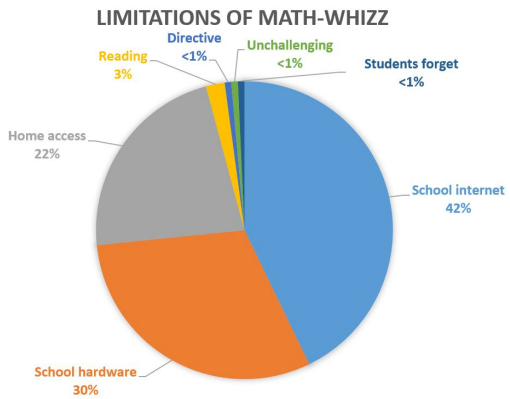


The strengths that were identified by the respondents remained constant, demonstrating that the training may have at least introduced teachers to all the necessary elements of Math-Whizz. Math-Whizz's approach received a high number of responses. The teachers appreciated the progress the students make using the individualised, varied and interactive activities. Furthermore, the teachers felt that students found Math-Whizz motivating. Taken together, these two aspects (Math-Whizz's approach and motivation), appeared in 53% of the comments during training and 60% after implementation. The slight increase in the number of responses related to mathematical thinking may also reflect the teachers' increased understanding of Math-Whizz post implementation.

Perceived limitations of Math-Whizz

The teachers were also asked to identify any limitations of Math-Whizz. 136 responses were made. The overwhelming aspects identified relate to resources: Nearly half (42%) of the responses stated a lack of internet in school, ranging from either no internet accessibility at all, the internet cutting out (and therefore freezing/cutting off the students' sessions) or not being able to cope with so many tablets connected at the same time. Nearly one third (30%) referred to the lack of computers/tablets in school, which makes student access difficult. A further fifth (22%) of the responses were concerning the students' lack of access to resources at home internet, hardware or both.

In addition to these infrastructure concerns, there were a very small number of pedagogically-related comments made. Three teachers (2%) identified that for younger students reading the activities can be an issue. One teacher identified the lack of training for teachers, another about the lack of space/freedom for students within Math-Whizz's design, and another about the lack of challenge in some cases. Finally, one teacher noted that when students stop practicing, they forget easily.

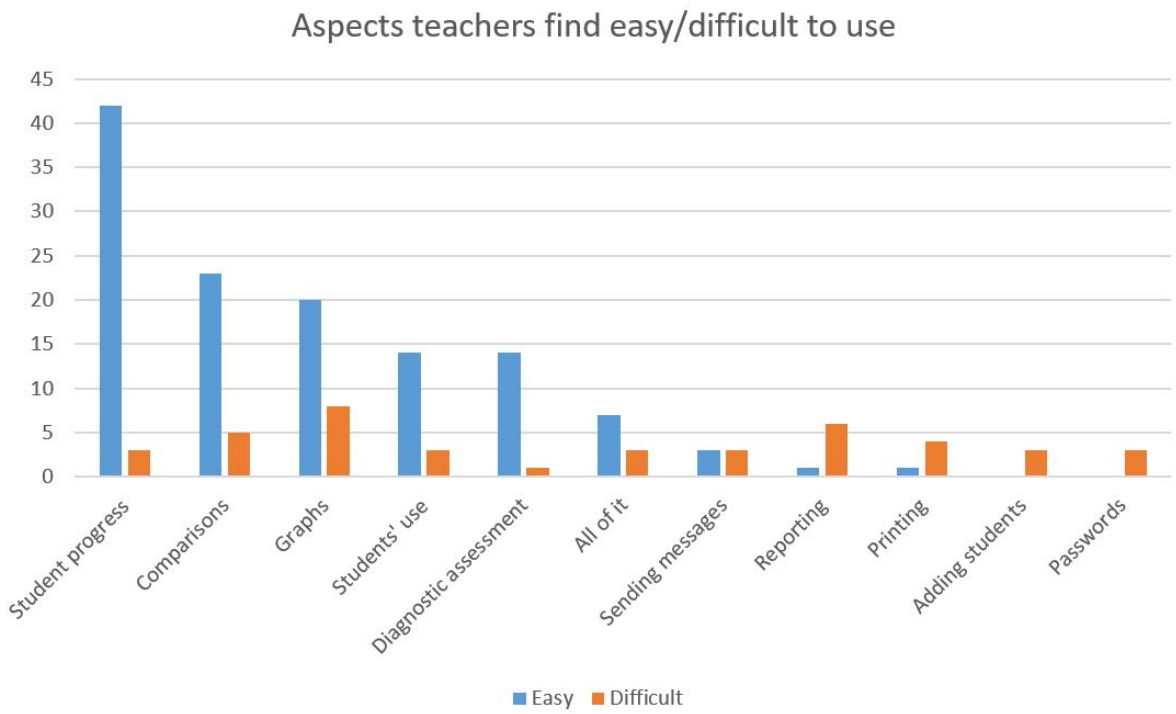


Clearly school and home access to Math-Whizz is an issue for a significant proportion of the responders (94% of responses are related to these issues).

The monitoring and reporting tool: aspects teachers find easy or difficult to use

Given the central importance of the monitoring and reporting tool, we were interested to identify the aspects teachers found easy and difficult to use in the monitoring and reporting tool. A total of 138 responses were given for ease of use and 56 responses were provided that identified difficulties related to the monitoring and reporting tool.

Easy to use (number of responses)	**Aspect provided by teachers**	**Hard to use (number of responses)**
42	Student progress	3
23	Comparisons (student/teacher/school)	5
20	Graphs	8
14	Student use of Math-Whizz (time on program)	3
14	Formative assessment (identifying students' strengths and weaknesses)	1
7	All of it	3
3	Sending messages/feedback	3
1	Reports (downloading, interpreting, format)	6
1	Printing	4
0	Adding students / preparing a class	3
0	Passwords (changing / using)	3



We received more responses from teachers about what they found easy to use, indicating that Math-Whizz is straightforward for the majority of teachers. Some teachers found some aspects difficult to use, including interpreting graphs.

Teachers' use of feedback

84 responses were received in the open question regarding how teachers used the monitoring tool. The responses can be broadly grouped into three categories:

1) using the monitoring tool as a **method of establishing/verifying students' use of Math-Whizz**, such as checking which tasks have been completed, giving rewards to students as they complete tasks and checking homework. E.g, "To inform the parent the importance of the student to use outside school", "To compare the time that students have used".

2) using the monitoring tool as a method of **checking and encouraging progress**, such as using the feedback as motivation to learn, providing "healthy competition" between students, and informing parents of their child's progress. E.g., "[I] print the report and annex it to the ballot / evaluation report which I hand over to parents every two months", "To publicize progress to parents."

3) using the monitoring tool as a **way to enhance teaching/learning practises** by identifying students' strengths and weaknesses, using it to identify follow-up tasks required in class, informing parents of areas they can help their children with at home, and adapting their teaching to address the needs of their students. E.g., "To plan activities that reinforce the contents with the teaching of mathematics."

Use	Percentage	Total for category
Category 1: Establishing use		
Verifying students' use of Math-Whizz	13%	17%
Providing homework	2%	
Instruct	1%	
Send message	1%	
Category 2: Encouraging participation		
Check students progress	26%	61%
Provide encouragement	16%	
Inform parents of children's progress	15%	
Inform students of their progress	4%	
Category 3: Enhancing learning/teaching		
Identify students' strengths and weaknesses	11%	22%
Improve teaching by identifying areas of need	5%	
Identify follow-up activities	4%	
Encourage parental support	2%	

We see relationships between Hattie and Timperley's (2007) levels of feedback to enhance learning with the categories identified within the teachers' responses. They highlight how important it is that both teachers and students respond to feedback to enhance learning. Beyond the observed students' use of feedback at a more micro-level (i.e. within the lessons and exercises they solve), there is a close association with Hattie and Timperley's first three levels and the teachers' responses in the categories.

1. *Task level: how well tasks are understood / performed.*
   At this level, teachers perceive Math-Whizz as a task where students are required to complete certain requirements, such as time spent on the platform. Their feedback is related to fulfilling the requirements, such as: "to reward those who use it", "to verify who works at home with the program" or "grant a prize to those who complete one hour per week".

2. *Process level: the main process required to understand/perform tasks.* Here teachers are focusing on the process of students using Math-Whizz. They are monitoring students' progress and are informing parents within that set process, e.g. "To publicize progress to parents", "to inform parents on the progress of their children".
3. *Self-regulation level: self monitoring, directing, and regulating of action*. Those teachers operating at this level are using feedback to self-direct and regulate their teaching and the students' learning, "to see the content I need to strengthen", "adapting my maths classes for the learning styles of my students".
4. *Self level: personal evaluations and affect (usually positive) about the learner.* While the questionnaire answers to not provide specific evidence of the teachers using self-level (level 4) feedback, our observations and teacher interviews suggest that there are some personal affect-related feedback practices taking place.

The feedback teachers receive through the monitoring tool in Math-Whizz appears to be helping the majority of teachers and students (according to 60% of responses) to be encouraging participation to understand their performance and desired goals. Additionally, one fifth of teachers are using the feedback to enhance their students' learning and their own teaching. The number of teachers who use feedback in this way at this early stage of implementation is heartening as it demonstrates these teachers are embedding Math-Whizz as an additional support for their teaching and the students' learning. A shift towards more use of identifying students' strengths and weaknesses to identify new tasks for the students and as a basis of discussion with parents would be a positive next step for the 60% of users who are currently simply encouraging student participation in Math-Whizz.

6. Limitantes

La separación de una relación causal entre el uso de cualquier tecnología y aprendizaje o resultados de rendimiento y otros factores que pueden distorsionar este tipo de relaciones, es un problema conocido en el campo (p. ej. OCDE, 2015, p. 146, Hanley y otros, 2015). El presente estudio sólo puede considerar las diferencias observadas entre los estudiantes o escuelas que participaron y las que no. Especialmente bajo condiciones donde una aleatorización de acceso o de uso no es posible o factible, investigaciones como ésta tienen que sujetarse a considerables limitaciones. Mientras que es posible que las relaciones observadas entre el uso del software y una prueba reflejen efectos causales, estas relaciones podrían ser revertidas o explicada por factores que no son (o no podían ser) medibles.

Por ejemplo, los estudiantes tienden a elegir distintos niveles de uso, dependiendo de sus propias habilidades, capacidades u otros aspectos (OCDE, 2015, pág. 147, Eickelmann y otros 2012). Además, no ha sido probado que los logros de estudiantes en escuelas con un cierto acceso a las TIC son comparables con otros. Algunos ejemplos son que el acceso a los recursos de TIC y su uso (eficiente) podría depender de los maestros, escuelas o directores, como su motivación en relación a la implementación, disponibilidad de recursos, personal y financiación (OCDE, 2015, pág. 147). Estudios en este ámbito carecen de la habilidad para determinar la dirección de la verdadera relación causal, aun cuando exista.

Mientras que en el presente estudio, parte de la información estaba disponible a través de los logros de cada usuario en Math-Whizz y un grupo control de estudiantes, una exhaustiva recopilación de datos sobre el contexto socioeconómico, la motivación para participar en la prueba o el programa y otros factores están más allá del alcance del presente estudio. Debido a las restricciones de tiempo y disponibilidad de esta información, los resultados de la prueba de septiembre servirán como indicador para factores como el nivel socioeconómico o auto-selección en

6. Limitations

The disentanglement of causal relationships between the use of any technology and learning or performance outcomes and other factors that might distort the view on such relationships, is a well-known problem in the field (e.g. OECD 2015, p. 146, Hanley et al., 2015). The current study can only consider the observed differences between students or schools who participated in the programme and those that did not. Especially under conditions where a randomization of access or usage is not possible or plausible, investigations like this one have to bear with considerable limitations. While it is possible that observed relationships between software use and a test score reflect causal effects, these relationships might actually be reversed or explained by factors that were not (or could not be) measured.

For example, students are likely to self-select into different usage intensities, depending on their own computer skills, abilities or are other aspects (OECD 2015, p. 147, Eickelmann et al. 2012). Moreover, it is not certain that students in schools with a certain access to ICTs and their achievements are comparable to others. Examples are that the access to ICT resources and their (efficient) use might depend on characteristics of teachers, schools or principals, such as motivation concerning implementation, availability of resources, able staff and funding (OECD 2015, p. 147). Studies in this field lack the ability to determine the direction of the true causal relationship, even when one might exist.

While in the current study, some information was available on the previous achievement of the Whizz user and control group students, a thorough collection of data on the socio-economic background, students motivation to participate in the test or the programme and other confounding factors had to lie beyond the scope of the current study. Due to the resource restrictions and lack of such data, the September score of the students might serve as a proxy for some confounders such as socio-economic status or self-selection rather than act as an accurate measure of student

lugar de actuar como una medida precisa del conocimiento del estudiante. Además, es probable que pueda haber diferencias sistemáticas presentes en casa o en distintas escuelas. Por lo tanto, los resultados deben ser interpretados cautelosamente.

Asimismo, la muestra no fue tomada de forma aleatoria de la población de escuelas y estudiantes en el distrito ya que su disponibilidad y la viabilidad del estudio en el corto período de tiempo fue importante. Una gran cantidad de usuarios ha tenido acceso o utilizado el software desde antes de la medición de septiembre. Junto con la falta de información sobre las variables demográficas que están asociados con el aprovechamiento de los alumnos, esto hace más difícil evaluar la relación entre el nivel de aprovechamiento y el uso de Math-Whizz así como su intensidad de uso. Dicho esto, la figura 5.1 muestra un efecto que corresponde a lo esperado con el mayor uso de Math-Whizz. Así que a pesar del hecho de que los dos grupos no eran significativamente diferentes en términos de los resultados de la prueba de Septiembre, esto no asegura que los dos grupos son realmente iguales (cf. Five-Schaw 2006, p.81).

Otra importante limitante del estudio corresponde a las pruebas que fueron utilizadas. Restricciones de tiempo y recursos hicieron necesario el uso de un instrumento de evaluación breve. Aunque las pruebas utilizadas en Septiembre y Octubre tienen un grado moderado-alto de fiabilidad, la brevedad del cuestionario significa que estas pruebas sufren del efecto 'suelo' y 'techo', esto es, hay un gran número de alumnos que no pudieron respondieron las 10 preguntas en la primera evaluación y un gran número de alumnos que contestaron a todas las preguntas de forma correcta en la segunda evaluación. Esto podría conducir a un sesgo en las puntuaciones, efectos y violaciones de los supuestos detrás de los análisis de regresión realizados. Aunque realizamos análisis para la validación de los resultados, necesitamos una evaluación validada y prolongada para capturar la habilidad de los alumnos.

knowledge. In addition, it was not possible to compare software usage at home or in schools. It is conceivable that be systematic differences are introduced at home or at schools. As such the results have to be interpreted very cautiously.

Similarly, the sample was not a random sample from the population of schools and students in the district since their availability and feasibility of the overall study in the brief time period was important. A large share of users has had access or been using the software already before the September measurement. Together with the lack of information on demographic variables that are associated with student achievement, this is making it more difficult to assess a relationship between outcome achievement and the usage of Math-Whizz as well as its intensity. Having said that, figure 5.1 shows a "fanning out" effect that corresponds to what we would expect when higher usage of Math-Whizz corresponds to higher progress in test achievement. So despite the fact that the two groups were not significantly different in terms of the September score test, this does not ensure that the two groups are truly equal (cf. Five-Schaw 2006, p.81).

Another main limitation of the study concerns the tests that were used. Constraints in terms of time and resources necessitated the use of a rather brief assessment instrument. Although the tests used used in September and October have a moderate to high reliability the brevity of the questionnaire means that these tests suffer from strong 'ceiling' and 'floor' effects, i.e. there is a high number of students who were not able to answer any of the ten questions in the first test and a high number of students who answered every question correct in the second measurement. This might lead to a downward bias in scores and effects and violations of the assumptions behind the conducted regression analyses. While more complex analyses were conducted to validate the findings, these cannot compensate for a validated assessment to capture students' ability, a longer evaluation period and scientific scaling methods.

7. Recomendaciones

En esta sección ofrecemos recomendaciones en tres partes que el Instituto de Educación quizás desee considerar en el futuro. Las recomendaciones sirven para celebrar el inicio positivo de la implementación de Math-Whizz y para aprovechar estos éxitos para mejorar aún más la experiencia de los estudiantes en matemáticas. El primero está relacionado con la implementación inicial con un enfoque en los maestros: la capacitación que recibieron y cómo están usando Math-Whizz. El segundo se relaciona con la infraestructura que se requiere para apoyar totalmente la implementación completa de Math-Whizz. La última está relacionada con la propia metodología de evaluación y recomienda cómo podría llevarse a cabo una revisión más sólida y sistemática.

7.1 Implementación inicial

Recomendación 1.1: *Diseñar e implementar una segunda oportunidad de desarrollo profesional para apoyar a un mayor número de maestros para integrar completamente Math-Whizz en su pedagogía.*

En última instancia, la oportunidad de desarrollo profesional, en cualquier formato, debería centrarse en un mayor uso de la herramienta de monitoreo y presentación de informes como una herramienta de evaluación formativa para más maestros mediante su uso para identificar las fortalezas y debilidades de los estudiantes. Base de la discusión con los padres. La oportunidad de desarrollo profesional podría tomar una variedad de formas tales como:

- un nuevo taller para que los profesores aborden sus preocupaciones específicas y les brinden la oportunidad de reflexionar sobre su pedagogía actual y las "mejores prácticas"

- perfeccionar el papel de los embajadores de Math-Whizz en la escuela a través del desarrollo profesional especializado para que estén mejor colocados para apoyar a sus colegas usando un enfoque en cascada

- el desarrollo de redes de escuelas (tanto en contextos similares como en contextos diferentes) cuyos profesores puedan apoyarse mutuamente.

7.2. Infraestructura

Recomendación 2.1: *Asegurar que todas las escuelas tengan suficiente hardware (computadoras o tabletas) para que cada niño en un salón de clases pueda tener acceso a Math-Whizz en el tiempo dedicado.*

Asegurar que todos los estudiantes tengan acceso a Math-Whizz permitirá una lección dedicada a ser emprendida. También demostrará el compromiso hacia la intervención, mientras que se apoyan a otras áreas, además de las matemáticas.

7.1 Initial implementation

Recommendation 1.1: *Design and implement a second professional development opportunity to support a greater number of teachers to fully embed Math-Whizz into their pedagogy.*

Ultimately the professional development opportunity, in whatever format, should focus upon more use of the monitoring and reporting tool as a formative assessment tool for more teachers by using it to identify students' strengths and weaknesses, to identify new tasks for the students and as a basis of discussion with parents. The professional development opportunity could take a range of forms such as:

- a new workshop for teachers to address their specific concerns and provide an opportunity for them to reflect on their current pedagogy in light of 'best practice'

- honing the role of Math-Whizz ambassadors in school through specialist professional development so they are better-placed to support their colleagues using a cascading approach

- developing networks of schools (either in similar contexts or different contexts) whose teachers are able to support each other in their implementation;

7.2. Infrastructure

Recommendation 2.1: *Ensure all schools have sufficient hardware (computers or tablets) for every child in a classroom to be able to access Math-Whizz at the dedicated time.*

Ensuring all students have access to Math-Whizz will enable a dedicated lesson to be undertaken. It will also demonstrate the commitment towards the intervention, while supporting other areas than mathematics.

Recomendación 2.2: *Asegurar que todas las escuelas tengan suficiente acceso a Internet para que todos los niños puedan acceder simultáneamente a Math-Whizz*.

Una conexión a Internet estable y consistente, que continúa funcionando cuando todos los estudiantes están conectados es necesaria para el uso continuo y el compromiso con la intervención.

Recomendación 2.3: *Asegurar que haya contingencias realistas disponibles para los niños y sus familias que no tengan acceso a Math-Whizz en casa*

Tiempo adicional en Math-Whizz parece tener el potencial de un impacto positivo en los logros del aprendizaje de los estudiantes. Por lo tanto, es importante que aquellos estudiantes que no tengan acceso en casa puedan usar la plataforma de aprendizaje en otros momentos como parte de su tarea.

7.3. Evaluación adicional

Recomendación 3.1: *Diseñar y poner en práctica una evaluación experimental a gran escala relativa que tenga en cuenta las iniciativas gubernamentales sobre la prueba de alumnos*, incluido el procedimiento de muestreo científico, el diseño cuidadoso de los instrumentos de encuesta y una forma más sistemática de incluir opiniones y observaciones de estudiantes y maestros del uso de supervisión

Recommendation 2.2: *Ensure all schools have sufficient internet access for all children in a to be able to access Math-Whizz simultaneously.*

An internet connection that is stable and consistent, that continues to work when all students are connected is necessary for continued use and commitment to the intervention.

Recommendation 2.3: *Ensure realistic contingencies are available for children and their families who do not have access to Math-Whizz at home*

Additional time on Math-Whizz seems to have the potential towards a positive impact on students' learning gains. Therefore it is important for those students who do not have access at home to be able to use the learning platform at other times as part of their homework.

7.3. Further evaluation

Recommendation 3.1: *Design and implement a relative large-scale experimental evaluation that takes into account government initiatives on pupil testing,* included Scientific sampling procedure, careful design of survey instruments and a more systematic way to include student and teacher opinions and observations of the use of monitoring

References

Bhuiyan, T. & Mahmud, I. (2015) Digital game-based education: a meta analysis. International Conference of Inclusive Innovation and Innovative Management (ICIIIM 2015). Available from https://www.researchgate.net/publication/295859083_Digital_Game-based_Education_A_Meta_Analysis

Clark, D., Tanner-Smith, E., Killingsworth, S . (2014). Digital Games, Design and Learning: A Systematic Review and Meta Analysis (Executive Summary). Menlo Park, CA: SRI International. Available from: https://www.sri.com/sites/default/files/publications/digital-games-design-and-learning-executive_summary.pdf

Clark-Wilson, A., Oldknow, A. & Sutherland, R. (2011) Digital technologies and mathematics education. A report from a working group of the Joint Mathematical Council of the United Kingdom. Available from http://www.jmc.org.uk/documents/JMC_Report_Digital_Technologies_2011.pdf

Daly, C., Pachler, N. & Pelletier, C. (2009) Continuing Professional Development in ICT for teachers: A literature review. WLE Centre, Institute of Education, University of London. Available from http://eprints.ioe.ac.uk/3183/1/Daly2009CPDandICTforteachersprojectreport1.pdf

Fife-Schaw, C. (2006). Quasi-experimental designs. Research methods in psychology

Higgins, S., Xiao, Z. & Katsipataki, M. (2012) The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation Full Report Education Endowment Fund / Durham University. Available from: https://v1.educationendowmentfoundation.org.uk/uploads/pdf/The_Impact_of_Digital_Technologies_on_Learning_FULL_REPORT_(2012).pdf

National Centre for Excellence in the Teaching of Mathematics, Mathematics and digital technologies: New Beginnings. 2010. Available from https://www.ncetm.org.uk/public/files/3399763/NCETMDigitTechReport2010.pdf

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Rummel, N; Mavrikis, M; Wiedmann, M; Loibl, K; Mazziotti, C; Holmes, W; Hansen, A; (2016) Combining Exploratory Learning with Structured Practice to Foster Conceptual and Procedural Fractions Knowledge. In: Looi, CK and Polman, J and Cress, U and Reimann, P, (eds.) Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016, Volume 1. (pp. pp. 58-65). International Society of the Learning Sciences (ISLS): Singapore. Green open access

Sitzmann T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology 64, 489–528.

Villegas-Reimers, E. (2003) Teacher professional development: an international review of literature. UNESCO: International Institute for Educational Planning. Available from http://unesdoc.unesco.org/images/0013/001330/133010e.pdf

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013, February 4). A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. Journal of Educational Psychology, 105(2), 249-265.

Appendix 1 - Estudios de caso ilustrativos

Parent of a fourth grader (9 year old)

Interviewer: As a mother, how has your experience of Math-Whizz been?

Parent: It has been gratifying and surprising because D is a child who used to count with his fingers. He was not able to do basic maths without his fingers. It was suprising for me to see how very quickly D changed … now he doesn't use his fingers to count and this is the only thing of many others he has improved since using Math-Whizz. He has improved a lot in problem solving, multiplication tables, adding, as well as many others. Now he can quickly solve all of this in his head in a much more efficient way.

While he was in Grade 1 and 2 he told me that he didn't like math because he didn't understand it, and now with this fantastic tool Math-Whizz, he actually enjoys math. It is much easier for him and now he gets wonderful grades. His maths grades have improved from a 6 or 7 to a 9 in a 10-point scale that in first grade he didn't think he could get.

It has been very gratifying because he has improved in math and now he is not afraid to use technology. Before Math-Whizz he was a little cautious about technology and now the useage of computers and tablets is something natural for him. I have seen this improvement in this last academic year and this is wonderful.

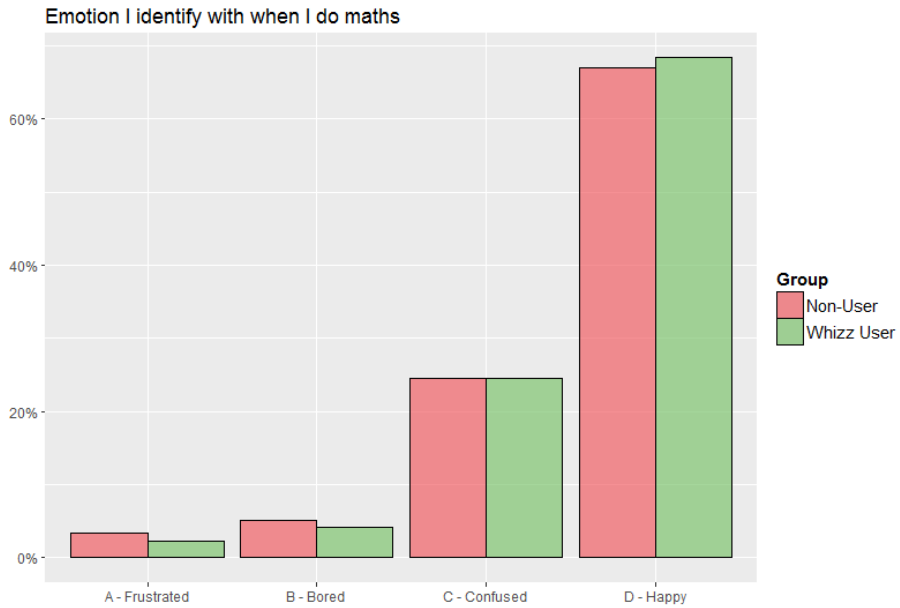
Principal from a pilot project school

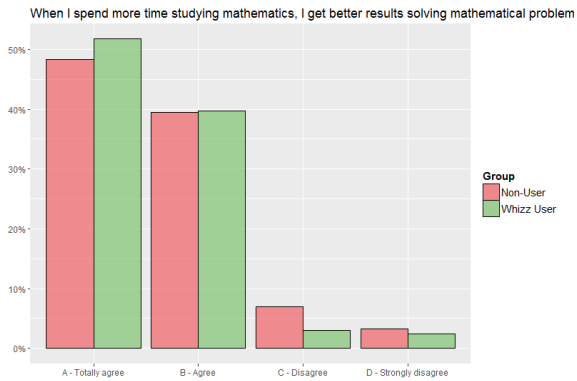
The results gained through the pilot project were very positive, children improved a lot in their math knowledge. Children that were struggling with the traditional classroom approach got excellent results through Math-Whizz. We are greatly satisfied in this regard, children love to visit the computer lab in order to log in to Math-Whizz. Most of them also have the opportunity to access Math-Whizz at home, we only have a few exceptions who don't have internet.

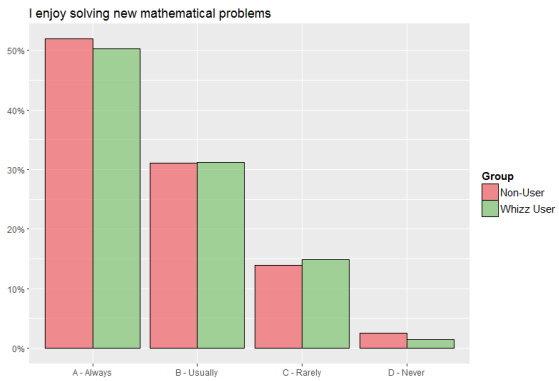
During the pilot we had the support of a trainer, a tech-savvy young man with an academic background. He was there to support the teachers if they had difficulty with the platform but I have to say that we didn't have many of these situations because Maths-Whizz is very easy to use, easy to manage and very user-friendly. He helped us by supervising the class while using Math-Whizz because we have large classes and it was great to have someone help the teachers.

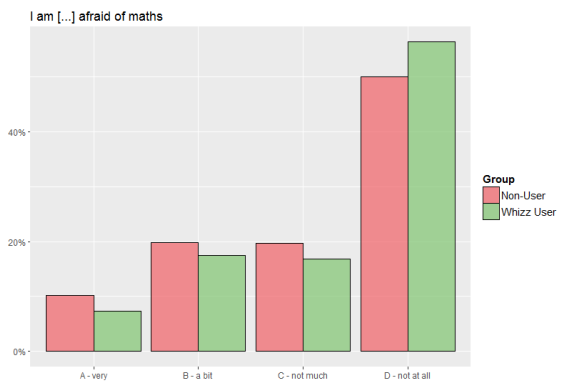
Teachers experienced the improvement of the class, they could see that students now understood the exercises and explanations much faster and easier and children could get better outcomes thanks to Math-Whizz.

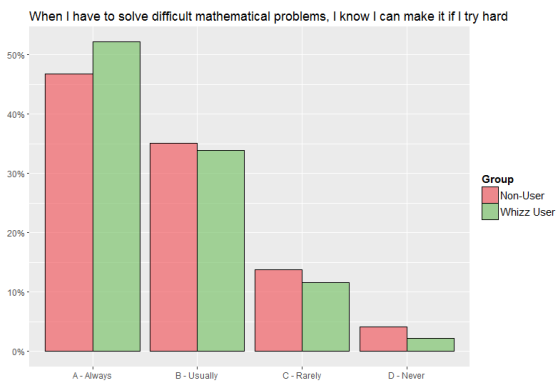
Appendix 2 - Motivación y actitud hacia las matemáticas







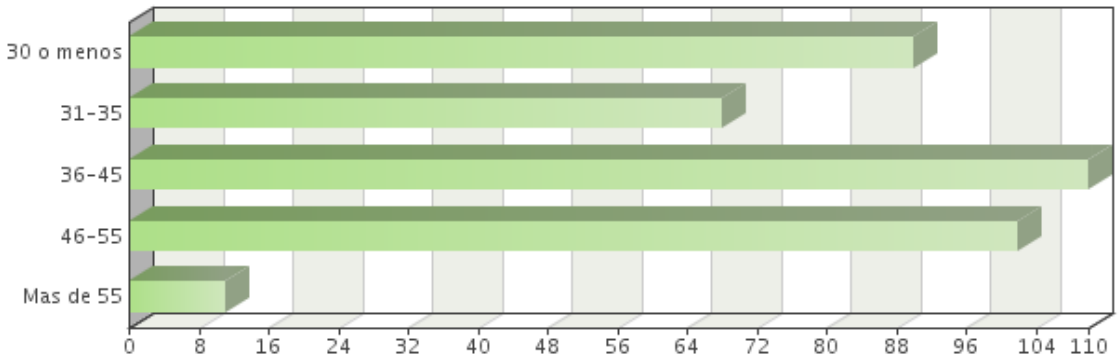




Appendix 3 – Cuestionario de docentes

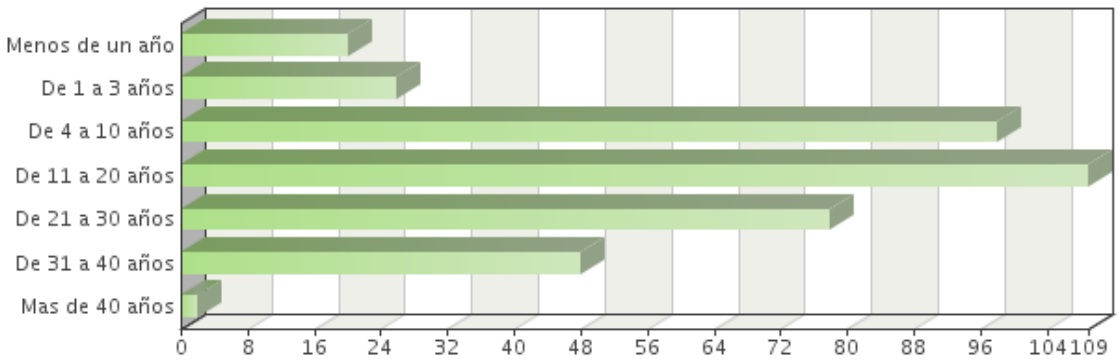
Question 3

¿Cuántos años tiene?



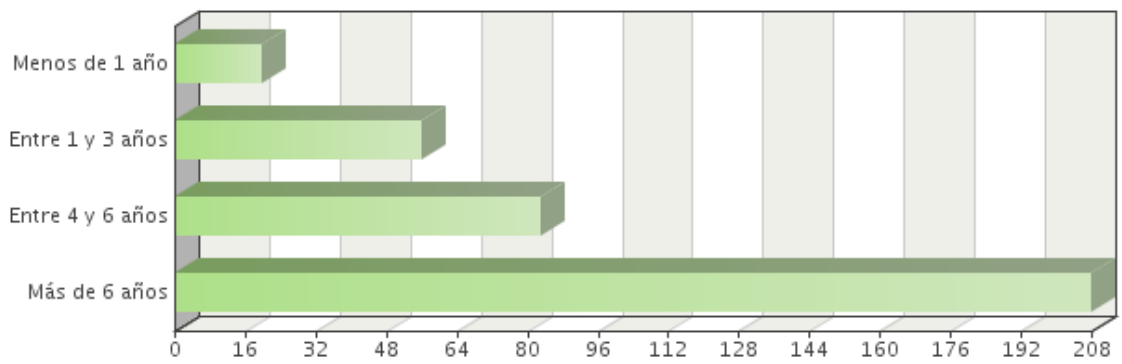
Question 4

Considerando este año, ¿Cuántos años tiene de experiencia en la práctica docente?



Question 7

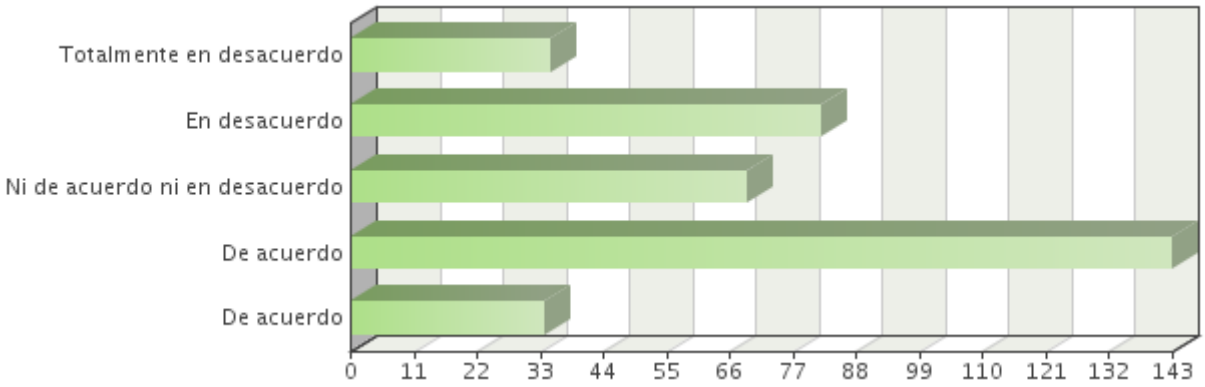
¿Desde hace cuántos años utiliza computadoras o Internet en algún centro



Question 10

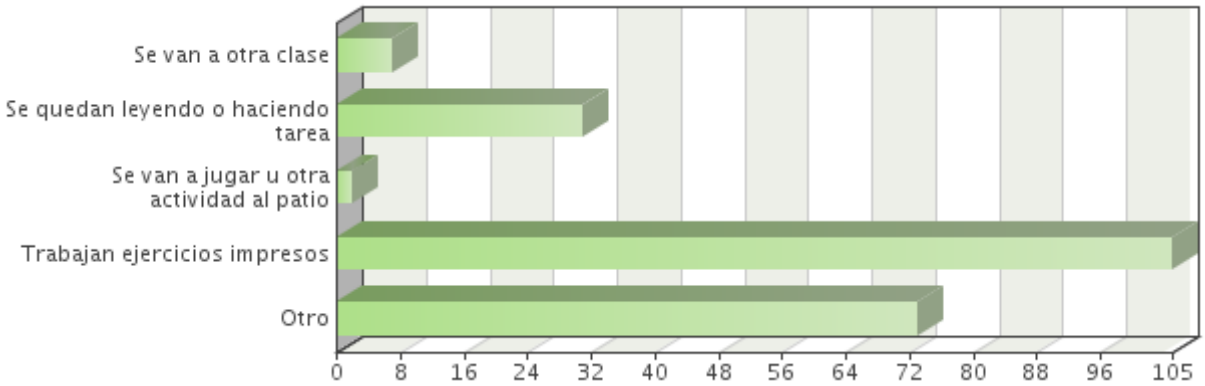
¿Qué tanto está de acuerdo con las siguientes afirmaciones acerca del apoyo para el uso de las TICs?

Recibo apoyo o consejos (en línea, mi escuela u otro lado) acerca del uso de las TICs en mi forma de enseñar.

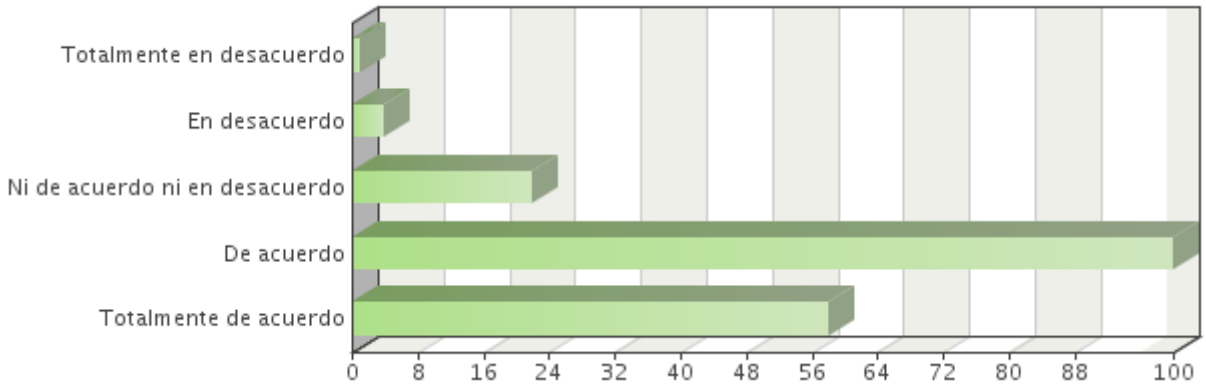


Question 15

¿Qué pasa con el resto de la clase que no usa Math-Whizz en ese momento?



Math-Whizz puede ayudarme a alcanzar mis metas propuestas en el programa curricular.



Appendix 4 - R software packages used for the analysis

Nathaniel Graham, Mahmood Arai and Björn Hagströmer (2016). multiwayvcov: Multi-Way Standard Error Clustering. R

package version 1.2.3. https://CRAN.R-project.org/package=multiwayvcov

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4.

Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.

http://CRAN.R-project.org/package=stargazer

Yihui Xie (2016). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.14.

Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963

Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch

and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN

978-1466561595

Lüdecke D (2016). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.1.0, <URL:

https://CRAN.R-project.org/package=sjPlot>.

Lüdecke D (2016). *sjmisc: Miscellaneous Data Management Tools*. R package version 2.0.0, <URL:

https://CRAN.R-project.org/package=sjmisc>.

Revelle, W. (2016) psych: Procedures for Personality and Psychological Research, Northwestern University,

Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.6.9.