Intervenciones educativas en era digital: desarrollo de herramientas para la inferencia de algoritmos del cómputo humano

López y Rosenfeld, Matías

2017-09-22

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Intervenciones educativas en era digital: desarrollo de herramientas para la inferencia de algoritmos del cómputo humano

Tesis presentada para optar al título de
Doctor de la Universidad de Buenos Aires
en el área de Ciencias de la Computación

Lic. Matías López y Rosenfeld

Directores de tesis: Dr. Diego Fernández Slezak
Dr. Mariano Sigman
Consejero de estudios: Dr. Agustín Gravano

Buenos Aires, Julio 2017
Fecha de defensa: 22 de Septiembre de 2017

INTERVENCIONES EDUCATIVAS EN ERA DIGITAL: DESARROLLO DE HERRAMIENTAS PARA LA INFERENCIA DE ALGORITMOS DEL CÓMPUTO HUMANO

En los últimos años el avance de la tecnología permitió el acceso masivo a nuevas plataformas digitales. Esta transformación tiene un impacto en la investigación en neurociencia educacional, donde ahora es posible que los estudios realizados en estos contextos sean pensados a partir del uso de los nuevos medios disponibles. El objetivo de esta tesis es construir una herramienta que permita recolectar datos masivos y analizarlos buscando inferir algoritmos del computo humano. Para ello se presenta el diseño y el desarrollo de una plataforma para la administración y evaluación automática de intervenciones educativas basada en videojuegos, llamada Mate Marote. Como casos de estudio, se realizaron trabajos con foco en el uso de tecnologías y su impacto. El primero explora el modo en que el uso de dispositivos móviles afecta a la percepción y la calidad de la comunicación cara a cara. Se encontró que la comunicación con interrupciones mediadas por dispositivos digitales afecta la percepción tanto del hablante como del receptor de la comunicación. Esto indica que hay que tener particular atención al incorporar medios digitales en tareas cotidianas, ya que tienen un impacto en las relaciones entre personas. El segundo analiza estrategias para la transmisión de conceptos utilizando una aplicación de celular y muestra la efectividad de las mismas utilizadas por los participantes. Considerando que uno de los principales medios de comunicación es el texto escrito, encontramos que este medio es notablemente ambiguo a la hora de transmitir figuras geométricas. El estilo utilizado al componer las descripciones afecta la efectividad de la transmisión del mensaje, siendo la descripción algorítmica el estilo más efectivo. Por último, abordamos la flexibilidad cognitiva de niños resolviendo puzzles, con foco en cómo pueden reconocer los propios errores en la resolución del juego y corregirlos. Vemos que la capacidad de modificar un plan al encontrarse con una situación equivocada aumenta con la edad; también es modificada por la manera en que cada niño aprendió a jugar: de un adulto –menor capacidad de cambio– o de un par –mayor capacidad de cambio–. Finalmente se reflexiona sobre la estructura interna, lógica y complejidad de juegos y pruebas cognitivas que deben ser tenidas en cuenta en el uso de estos recursos con fines educativos.

Palabras claves: Intervenciones Educativas, Tecnologías de la Información y la Comunicación, Videojuegos, Educación, Neurociencia educacional.

EDUCATIONAL INTERVENTIONS IN THE DIGITAL ERA: DEVELOPMENT OF TOOLS TO INFER HUMAN COMPUTATION ALGORITHMS

In recent years the advance of technology has allowed massive access to new digital platforms. This transformation has had an impact on Educational Neuroscience research: studies conducted in these contexts are now able to use the newly available technology. The objective of this thesis is to build a tool to collect massive online data and analyze it to infer human computational algorithms. The design and development of a platform of educational interventions based on video games, called Mate Marote, is presented. This thesis includes two studies focusing on the impact produced by the use of new technologies. The first experiment explores how the usage of mobile devices affects the perception and quality of face-to-face communication; digital device interruptions affect the perception of both the speaker and the receiver of a communication. This indicates that incorporating digital media into everyday life is not without its costs; it may impact relationships between people. The second study analyzes the transmission of concepts using a cell-phone app, showing the effectiveness of the different strategies used by the participants. Considering that written text is one of the most common ways to communicate, we find that it is remarkably ambiguous when transmitting geometric figures. Textual style affects the effectiveness of the transmission; algorithmic description is the most effective method compared with others, such as metaphors. Finally, this work addresses the cognitive flexibility of children solving puzzles, focusing on how they can recognize their own mistakes while solving a puzzle and correct them. The ability to change strategy when confronted with an error state increases with age; it is also affected by the way each child has learned to play: learn from an adult - lower ability to change - or learn from a peer - greater capacity to change strategy. Finally, the thesis examines lessons learned from the internal structure, logic and complexity of games and cognitive tests that must be taken into account when using these resources for educational purposes.

Keywords: Educational interventions, Information and Communication Technologies, Video games, Education, Educational neuroscience.

AGRADECIMIENTOS

Un doctorado es un viaje, y el mío fue viaje inimaginable. Allá por los 2000 yo terminaba la secundaria. Arrancaba la facultad y empezaba toda una nueva etapa. Una etapa en la que inicialmente tenía como único destino imaginable: terminar la facultad e irse del país a trabajar en el exterior, no quedarse en un país quebrado y en llamas. Inimaginable fue el rumbo que tomó mi vida desde ese momento. Gran parte de esto fue por culpa de las políticas tomadas en materia científica (entre muchas otras) por los gobiernos de Néstor Kirchner y Cristina Fernández de Kirchner, que transformaron esa realidad de *hay que irse del país* en la posibilidad de quedarse acá, y que dedicarse a la ciencia fuera algo posible. Otra gran parte de la culpa se la atribuyo a Mariano Sigman y Sebastián Lipina, que dieron forma a un proyecto de investigación (y al montón de personas que se sumaron y lo hicieron posible) con aplicación con el que pude empezar a soñar que dedicándome a la ciencia podría llegar a *cambiar el mundo* (aunque sea un poco).

Más allá de esto, la página de los agradecimientos siempre es un tema. Ya sucedió en la tesis de Licenciatura. Es un momento difícil, que se escribe estresado, cansado, agotado... y en la cuál uno nunca termina escribiendo exactamente lo que quería o cómo lo quería. No obstante lo cual, acá vamos con el último esfuerzo.

Creo que en mi formación de doctorado me crucé con mucha gente y muchas experiencias. Todas ellas me dejan como enseñanza cosas que quiero ser y cosas que quiero no ser. Me crucé con gente generosa, gente leal, gente de una ética profesional correcta, gente buena, gente que te trata bien, gente atenta y gente que no. Será mi esfuerzo de aquí en adelante rescatar lo bueno, no repetir los errores y evitar las malas experiencias, de las cuales uno no es consiente hasta que las vive.

Quisiera especialmente agradecer Álvaro y Diána por su amor y apoyo incondicional, y a Andre y Juli, participes de tantas aventuras.

A todos mis abuelos, que aunque ya no están sé que todos estarían orgullosos, con el pecho inflado comentándoselo a sus amigotes. A Gabi, que tampoco va poder estar y también tendría el pecho inflado. A todos los tíos y tías, de sangre y adoptados que me apoyaron a lo largo del viaje.

A las grandes amistades que siempre han estado: Nacho, Ali, Mica, Tommy, Román, Pachi, Fran, Facu, Fran, Dani, Gutes, Herman, Dani, Juli, Leti, Joa, Nori, Nico, Vero, Pato, Charlie, Sergio, David, Mariano, Tebi, Esteban, Facu. A Maite que acaba de llegar. A todos los que están perdidos por el mundo Marian, Dani, Ale, Jota, Eleze, Jesse, Belén, Dora, Pau, Germán, Mike, Greg. A los incontables aportes de Yli desde la lejanía. A todos aquellos que me sostuvieron con sus charlas en momentos de crisis y en buenos momentos también: Sebas, Nico, JPG, Pablito, Eli, Sole, Caro, Fer, Juanjo, Guido, Ce, Derek.

A los compañeros del agua que me acompañan, se preocupan, me cuidan y me

han dado energía para seguir: Diego, Kary, Carlitos, Mabel, Elvio y todo el equipo. Al mundo de los veleros y sus tripulaciones.

A la larga lista de compañeros de laburo que tuve: del LNI; del LIAA; de la UNA; del Bunge Lab; a a todos con los que compartí la experiencia Orga1 que son un montón; a la LASchool con todos sus participantes y organizadores. A Susan que me dio la oportunidad de ver uno de los mejores centros de investigación del mundo. A Silvia que me recibió en su laboratorio y me permitió no solo conocer a grandes amigos, sino pensar la ciencia distinta a lo que venía acostumbrado.

A Sebastián Uchitel, por darme un lugar para que yo pudiera acceder a una beca. A Aída y todos sus secuaces de secretaría. A mis directores, Diego y Mariano, por la oportunidad que me dieron.

A todo el Departamento de Computación, a Exactas, a la UBA, a CONICET y a todas las instituciones que hicieron posible este trabajo.

A todos los que en este momento no vienen sus nombres a mis dedos pero que igual están.

A Olmo.

Y sobre todo a quien me abrió su corazón y su vida: a Male.

A Male

CONTENTS

1. INTRODUCTION

El futuro llegó hace rato.
–Patricio Rey y sus Redonditos de Ricota

The world is going through one of the most important technological advances of history: Information and Communication Technologies (ICT) are becoming part of everyday life: in education, work and, leisure. Although there is still a technological gap existing due to the cultural and socioeconomic development among regions [52, 99, 125], access to ICT has become massive in a world wide scale [59, 58].

Since ICT are deployed in a wide variety social and cultural contexts, the implementation of cognitive training based on ICT has many challenges: generalization of results, evaluation of negative effects, replicability of the findings, among others [6]. These challenges require an interdisciplinary approach in order to explore the multiple dimensions involved in this phenomenon. Including technology in different stages of life generates changes in cognitive and learning competencies. Hence, the relevance of multidimensional and interdisciplinary approaches to study the complexity of these processes. However, interdisciplinary integration for research in ICT is incipient [59, 87].

Some aspects in which incorporation of ICT generated changes include the building of knowledge (i.e., learning) and communication patterns [8, 133, 52, 87, 114, 138, 141]. However, understanding these changes and impacts is neither clear nor obvious. On the contrary, there is an extensive and diverse discussion through books, articles in newspapers and magazines, television and even the opinion of professionals from different disciplines. This background contributes to generating a scenario in which divergent opinions for and against ICT coexist. However, this scenario lacks support on adequately constructed evidence for a clear and orderly discussion.

Argentina, in particular, is living a unique opportunity where an enormous amount of primary and secondary school students have their own laptop, provided by institutional programs such as Conectar Igualdad and Joaquín V. González (La Rioja), Plan Sarmiento (CABA). 70% of the population of Argentina had access to Internet in 2015 [7]. But this was not the case a few years before. Indeed, in 2010 the penetration was less than 50%. The fact that learners have their own personal computers allows novel educational applications and large-scale interventions.

As a consequence, new educational methodologies based on ICT were born. Massively Open On-line Courses (MOOCs) are virtual free on-line courses with an huge number of students in each course. These courses came across a series of challenges (e.g. preventing drop-off, effectiveness of learning) [77, 54, 140, 50]. Some of the most popular MOOCs are Coursera[1], Udacity[2], and Khan Academy[3].

Students' engagement is key for the new methodologies. Therefore, when attempting to get students more involved some educational projects bring video games into the classroom as part of their pedagogical tools [47, 76, 109, 12, 33]. These games pursue making the

[1] `https://www.coursera.org`
[2] `https://www.udacity.com/`
[3] `https://www.khanacademy.org/`

process of learning more engaging and enjoyable. Chocolate Fix (designed by the company Think Fun in 2008), Refraction (Center for Game Science in 2010), and DragonBox (We Want to Know in 2012) are a few examples of efforts in this area. In spite of the existing controversy of the effects of video games [48, 47], there is evidence which shows that video games usage can train cognitive abilities and even transfer to school learning domains [43, 69, 49, 80].

MOOCs deal with an enormous amount of students but have few teachers behind its courses. Therefore, it is a challenge for MOOCs to cater for the interaction between a teacher and a learner. As a consequence, scaffolding, feedback and assessments should be automated. The customizing of the software requires knowledge of the domain, i.e. understanding how people think, how people solve tasks and, how people interact with software specially designed for this purpose.

Beyond the mass distribution of equipment to children and adolescents, the available evidence regarding the impact on their cognitive development does not yet allow for specific recommendations on what, how and when to include ICT in education and parenting contexts to contribute in optimizing cognitive and learning competencies development [52, 106].

This thesis proposes the development of technological tools and the analysis of human behavior that allows generating clear empirical evidence to contribute to the debate about the pertinence, the way and, the moments of the incorporation of technology into the learning processes.

What kind of impacts does the use of different ICT have in formal and informal educational contexts? What are the variables that module such impacts at the individual and contextual level? What are the intrinsic characteristics of ICT? These are some of the questions that guide the work presented in the current document.

This thesis is a combined effort of Computer Science, Cognitive Science and Neuroscience fields which aims to contribute to the design of software –hidden under the skin of a video game– for cognitive training, looking for incorporate adaptive algorithms –to respect the individuality of people– and tools for analysis to study cognitive processes.

A new discipline has emerged not so long ago –called *educational neuroscience*– combining cognitive neuroscience and behavioral methods to investigate the development of mental representations and, possibly, to use neuroscience pre-existing knowledge to improve different teaching approaches [103].

In the learning processes a group of cognitive processes can be identified. To achieve a task there are demands in aspects of control –identify and inhibit irrelevant or conflicting stimulus–; in keeping *on-line* information during the activity; in exchanging solution strategies; in scheduling a complex task solution as a sequence of steps and execute them; or initiating novel sequences of action [20, 40].

During the last two decades, cognitive neuroscience and developmental psychology have designed and implemented a series of interventions aimed at optimizing the cognitive performance of different populations of children. In this context, cognitive training is defined as the process of improvement or optimization of performance through practice provided by exercises or cognitive training, or teaching proposals designed for the same purpose. In general, these studies focused on two objectives. One of them is oriented to answering empirical questions about the cognitive processes to be trained and the mechanisms by which changes occur. The other is oriented to the design of effective interventions, through which changes are verified.

There are different ways to determine the effectiveness of a training. The most common is to measure the performance improvement in the trained domain, in terms of efficiency and reaction times. Other variables that are also studied include the frequency of use of a particular strategy. If performance is measured over the intervention period, it is also possible to determine learning curves or trajectories, which show how learning acquisition changes over time. Likewise, the learning curves vary among the participants, so it is important to anticipate this consideration in the designs by means of which the impacts are analyzed. On the other hand, in addition to the evaluation of the change in performance, it is increasingly common to assess the long-term effects of training and its cumulative effects, since improvements can also be acquired during the exercise as a result of side effects. For example, increased motivation and learning [130].

Finally, another aspect of importance to evaluate is the generalization or transfer of training, both to tasks with similar demands to those trained (i.e., close transfer), as to other untrained ones (i.e., distant transfer). Different studies have verified both types of transfer [43, 66, 112, 113], although the evidence is not enough to enunciate the aspects or factors are associated with them. Future studies may study the mechanisms of mediation by which the changes occur and what is transferred, rather than focusing on the analysis of its effectiveness.

Interdisciplinary collaborations between researchers in cognitive and computer sciences have carried out a reverse engineering program to detect the operations involved in arithmetic calculations. Some of these studies have shown the universal existence of an approximate calculation system [28] which is present early in life [11] and it is ubiquitous in most cultures [30]. As a simple example of this process, any person takes more time to answer $52 > 31$ than $52 > 17$ [29]. Formally, these operations are equivalent; in approximate (analog) systems, the second operation results easier. This fingerprint is even observable in people with vast mathematical training [27]. Similar examples are found in geometry notions that are present even before formal mathematical education [60].

Discovering these kinds of characteristics allows efficient interventions. As knowing the processor architecture helps code optimization, deep knowledge behind the human computations might allow for the optimization of the education activities. Recently, cognitive neuroscience research revealed that development of neurocognitive systems is related to control, numeracy and literacy competencies. These show plasticity during brain organization and reorganization processes and that, through intervention in education, behavioral changes in the brain may be induced in normal or disordered children [46].

There are many efforts from different perspectives in *reverse engineering of the mind* [128] –in words of Josh Tenenbaum–: understand human learning and inference. This knowledge can be used to build new and more accurate learning methods and other systems used in the common life.

The internal cognitive structures of the mind –metacognition [31, 39, 78], decision making [72, 105, 74], fluid reasoning [37], working memory [137, 22, 35]– are not directly observable: they had to be explored via an indirect method. The analysis of gazed elements of the visual scene is an indirect technique. It consists in using the positions of the screen where the participant focused to explain why she decided on the answer [127, 53, 34]. Since eye tracking studies require special equipment to follow the gaze, scaling up this kind of studies is too expensive.

In a bigger scale, some studies analyzed cell-phone calls as large-scale networks, describing communication patterns and network structure [73, 97, 71, 98, 61]. Communication

data has been studied to understand human mobility patterns [45, 122, 85]. But it also offers an indirect measure for social interactions where relationships between people may be inferred from reciprocal calls patterns [139, 118, 64, 32, 16, 70].

Another possibility is observing human behavior when interacting with something. Observing someone playing chess is a way to understand how she plays and think: If she is aggressive; if she prefers moving the knights before the bishops; which are her favorite openings. This exercise could also be done observing data registered with a computer. i.e., someone opened a text processor and wrote: first the date in the right margin; then she wrote *To whom it may concern.* Just with these two lines, we could guess that she was writing a formal letter. The analysis of human interactions with software can reflect their thoughts. This becomes more achievable with access to technology –and access to Internet– . Other technique is to study log files or click streams of software usage [82, 19, 132, 80]. This strategy allows to detect subtle changes due to the analysis of a massive dataset. The challenge is how to build activities to answer research questions, and how to engage participants to use these activities.

Video games can get the attention of children and adolescents [89]. However, it is important to assess to which extent video game technology has an impact on childhood education. Video games have the capacity to engage children in learning experiences, which make them a powerful educative entertainment.

Several projects propose the introduction of games in the classroom as part of education activities [76, 109, 47, 80]. Some authors propose competition-based games for stimulating learning [12]; even for university [33, 15]. In summary, many initiatives introduce different games as a strategy to stimulate learning in students.

Games with a Purpose is one of most inspirational works of this thesis. These kinds of games were developed by Luis von Ahn under a great idea of using people doing different kinds of stuff with a computer to solve hard computational tasks. He presented the novel idea in his PhD Thesis [131]:

> *Construction of the Empire State Building: 7 million human-hours. The Panama Canal: 20 million human hours. Estimated number of human-hours spent playing solitaire around the world in one year: billions. A problem with today's computer society? No, an opportunity.*
>
> *What if this time and energy could be channeled into useful work? (...) Although computers have advanced significantly in many respects over the last 50 years, they still do not possess the basic conceptual intelligence or perceptual capabilities that most humans take for granted. (...)*
>
> *In this paradigm, we treat human brains as processors in a distributed system, each performing a small part of a massive computation. Unlike computer processors, however, humans require an incentive in order to become part of a collective computation. We propose on-line games as a means to encourage participation in the process.*

Some of the projects he developed include: The ESP Game, GWAP, reCAPTCHA, and Duolingo. With this projects, he was able to tag objects in images and locate those objects within the images. He also digitalized books that could not be digitalized by a computer. Other task solved: recognition and reading road signs; translation of documents into several languages; music styles tagging; collection of common definitions of words used by people to describe concepts, as an alternative to dictionary definitions; among others.

The way von Ahn collected human thoughts while solving tasks in an ecological environment can be defined as fun, and it showed how participants can engage while *solving tasks*.

The capitalization of these ideas into games and computational interventions based on the knowledge of the architecture of human processor – the brain – has been part of a growing effort [103].

Efforts to include ICT in education started more than 20 years ago. In the 80's, ideas of teaching children and adolescents programming skill were introduced [100]–an effort to consider ICT part of the new digital age–. Another proposal was the development of computerized tutors. They have supported individual students for guided learning, helping teachers that have very little time to spend with each one [5]. Traditional education in schools consists of supervised learning, where teachers communicate knowledge in classrooms and attend individual children questions. Classrooms are populated by many students per teacher, and thus teachers have little time slots with each child.

There is evidence showing experience where tutors had comparable results to human tutoring [1]. This implementation relies on the installation of the corresponding software on each computer or a constant Internet access. Thus, this solution is suitable for using the tutors in computer labs at schools or small-to-middle scale experiences [67].

Intelligent tutoring systems (ITS) were introduced around the 80s. It was software designed to provide scaffolding to the user to learn some specific domain, providing customized instruction and feedback to each student. Anderson et al made an ITS for students learning to program LISP language [4].

Building tutors requires a huge effort. Given that it has been estimated that 300 hours of expert design effort are required to design a single hour of content [2], the scalability of ITS is still poor. There were lots of efforts to make it scalable by reducing the complexity of the system to build the tutors (e.g. CTAT) [2] However, systems must still contemplate a huge number of possible interactions.

One of the well known ITS is the Cognitive Tutor [5], which uses the knowledge tracing [21] idea: model knowledge as a set of rules. Each one represents what the students should do when they are in a particular state with a certain goal. The system decides which specific problems show to a student based on her performance, and it gives certain problems until it considers the student has learned the involved rules.

The present thesis combines ITS principles and von Ahn's methodology to achieve a twofold objective: 1) to increase the understanding of the human mind through collected data of humans solving particular tasks; 2) use this knowledge to design and build cognitive interventions in large scale scenarios. The thesis aims to contribute to the state of the art with the study and development tools to favor equalizing academic outcomes across children closing gaps that exists due to the cultural and socioeconomic development.

Resumen en español

Introducción

El mundo atraviesa una de las transformaciones más importantes de la historia: dados los avances de las Tecnologías de la Información y la Comunicación (TIC) éstas se están convirtiendo en una parte esencial de distintos ámbitos de la vida cotidiana: la educación, el trabajo y el ocio.

Dado que las TICs se despliegan en una gran variedad de contextos sociales, económicos y culturales, la implementación de la formación cognitiva basada en ellas plantea múltiples desafíos: la cuestión de la generalización de los resultados de las investigaciones, la evaluación de eventuales efectos negativos, la replicabilidad de los hallazgos, entre otros. Estos retos requieren un enfoque interdisciplinario para explorar las distintas dimensiones involucradas en este fenómeno. Incorporar la tecnología en diferentes etapas de la vida, genera cambios en las competencias cognitivas y de aprendizaje. Por ello, son necesarios enfoques multidimensionales para estudiar la complejidad de estos procesos.

La incorporación de los cambios generados por las TIC permiten la construcción de nuevos conocimientos, aprendizajes y patrones de comunicación. Sin embargo, la comprensión de estos cambios y su impacto en la sociedad no es claro ni obvio. Por el contrario, existe una extensa y diversa discusión a través de libros, artículos periodísticos y revistas, televisión e incluso la opinión de profesionales de diferentes disciplinas. Este contexto contribuye a generar un escenario en el que coexisten opiniones divergentes a favor y en contra de las TIC, pero que carecen de evidencia empírica construida adecuadamente y que pueda contribuir a un debate serio sobre esta problemática.

En la Argentina, hoy tenemos una oportunidad única: la mayor parte de los estudiantes de la escuela primaria y secundaria pública tienen su propia computadora –provista por programas estatales como Conectar Igualdad, Joaquín V. González (en la Provincia de La Rioja) o el Plan Sarmiento (en la Ciudad Autónoma de Buenos Aires)–. En el año 2015, se estima que el 70 % de la población del país contaba con acceso a Internet [7]. El hecho de que los estudiantes posean su propia computadora con conectividad, habilita el uso de nuevas aplicaciones educativas e intervenciones a gran escala.

Los esfuerzos por incluir las TIC en la educación comenzaron hace ya más de 20 años. Fue en la década de 1980, cuando surgieron las primeras ideas de enseñar programación a niños y adolescentes como parte del esfuerzo por adaptar los aprendizajes a las necesidades de la nueva era digital. También se desarrollaron tutores informatizados capaces de guiar al estudiante de forma personalizada en su proceso de aprendizaje, y de este modo, complementar la tarea de los maestros frente a cursos numerosos.

A medida que pasó el tiempo y la digitalización de las prácticas de la vida cotidiana se extendieron, surgieron nuevas metodologías de aprendizaje basadas en las TIC. Por ejemplo, los cursos masivos *online* (MOOCs - Massively Open Online Courses) gratuitas en línea que cuentan con un gran número de estudiantes.

Estos cursos se encontraron con una serie de desafíos, entre ellos, la deserción y la eficacia del aprendizaje [77, 54, 140, 50].

El compromiso de los estudiantes con el curso y el material de estudio es clave para que estas nuevas metodologías funcionen. Un modo de atraer a los estudiantes es mediante el uso de videojuegos que se incorporan al aula como herramientas pedagógicas. Estos juegos buscan que el proceso de aprendizaje sea lo más estimulante y agradable posible. A pesar de las controversias respecto de las consecuencias del uso de videojuegos, existe evidencia que demuestra que éstos pueden entrenar capacidades cognitivas que pueden ser transferidas generando mejoras en el rendimiento escolar. [43, 69, 49, 80].

En este marco, la presente tesis tiene como objetivo el desarrollo de herramientas tecnológicas y el análisis del comportamiento humano a fin de generar claras evidencias empíricas que puedan contribuir al debate sobre la pertinencia, las formas y los momentos adecuados para la incorporación de tecnologías en los procesos de aprendizaje.

¿Qué tipo de impacto tiene el uso de las TICs en contextos educativos formales e informales? ¿Cuáles son las variables que modifican sus efectos a nivel individual y contextual? ¿Cuáles son las características intrínsecas de las TICs? Estas son algunas de las preguntas que guían esta investigación.

Esta tesis es un esfuerzo combinado de los campos de Ciencias de la Computación, Ciencia Cognitiva y Neurociencia que se propone contribuir al diseño de software –"disfrazado" de videojuego– para el entrenamiento cognitivo, buscando algoritmos adaptativos incorporados –individualidad de las personas– y herramientas de análisis para estudiar los procesos cognitivos.

Durante las dos últimas décadas, la neurociencia y la psicología del desarrollo han diseñado e implementado una serie de intervenciones dirigidas a optimizar el rendimiento cognitivo de diferentes poblaciones de niños. En este contexto, la formación cognitiva se define como el proceso de mejora a través de la práctica proporcionada por ejercicios o propuestas de enseñanza diseñadas para tal fin. En líneas generales, estos estudios se centraron en dos objetivos. Uno de ellos busca responder a preguntas empíricas sobre los procesos cognitivos que pueden ser entrenados y los mecanismos que generan cambios. La otra, se orienta al diseño de las intervenciones.

Trabajos interdisciplinarias que combinan las ciencias cognitivas y de la computación, han llevado a cabo un programa de ingeniería inversa para detectar las operaciones involucradas en los cálculos aritméticos. Algunos de estos estudios han demostrado la existencia universal de un sistema de cálculo aproximado que está presente desde edades tempranas y se encuentra presente en la mayoría de las culturas. Descubrir este tipo de características permite intervenciones eficientes. Como conocer la arquitectura del procesador ayuda a la optimización del código, el conocimiento profundo detrás de los cálculos humanos podría permitir la optimización de las actividades educativas.

Las estructuras cognitivas internas de la mente –metacognición [31, 39, 78], toma de decisiones [72, 105, 74], razonamiento fluido [37], memoria de trabajo [137, 22, 35]– no son directamente observables, por ello sólo pueden ser estu-

diadas mediante un método indirecto. Por ejemplo, puede captarse a través de la observación del comportamiento humano durante la realización de alguna tarea. Ver a alguien jugando al ajedrez, puede ser un modo de entender cómo juega y piensa: si la persona es agresiva; si prefiere mover a los caballos antes que a los alfiles; cuáles son sus aperturas favoritas. etc. Este ejercicio también podría realizarse observando los datos registrados con una computadora: alguien abre un procesador de texto y escribe primero la fecha en el margen derecho; luego "A quien corresponda:". Sólo con estas dos líneas, podemos arriesgar que está escribiendo una carta formal.

El análisis de las interacciones con software puede reflejar los pensamientos de la mente. Esto se hace más factible con el acceso a la tecnología –y a Internet–. Esta estrategia favorece la detección de cambios sutiles a través del análisis de un conjunto de datos masivo. El desafío consiste, entonces, en diseñar actividades que respondan a la preguntas de investigación y lograr involucrar a los participantes para que la lleven a cabo.

Los videojuegos pueden llamar la atención de los niños y adolescentes [89]. Solo observándolos jugar, resulta claro la preferencia por esta propuesta de aprendizaje y la capacidad de implicarlos en el proceso. Esto los convierte en una poderosa herramienta para el entretenimiento educativo. De allí la importancia de evaluar hasta qué punto y de qué forma la tecnología del videojuego puede potenciar la educación infantil.

Games with a Purpose es uno de los trabajos más inspiradores de esta tesis. Utilizando juegos, su inventor logró canalizar esa actividad en la resolución de tareas computacionalmente difíciles. Con estos proyectos fue capaz de etiquetar objetos en imágenes y localizarlos dentro de ellas. Con esta técnica, logró digitalizar libros que no habían podido ser pasados al formato digital por una computadora. Asimismo, permitió el reconocimiento y la lectura de señales viales; la traducción de documentos en distintos idiomas; el etiquetado de estilos musicales; la recolección de definiciones usadas habitualmente por distintas personas para describir conceptos, que pueden reemplazar a las definiciones tradicionales del diccionario; entre otros.

La presente tesis combina la incorporación de la tecnología y las ideas de Games with a Purpose para lograr un dos objetivos: 1) aumentar la comprensión de la mente humana a través de datos durante la resolución de tareas particulares; 2) utilizar este conocimiento para diseñar y construir intervenciones cognitivas en escenarios de gran escala. La tesis pretende aportar con el estudio y desarrollo herramientas que puedan en el futuro realizar intervenciones educativas que, mediante el entrenamiento cognitivo, permitan igualar las oportunidades de niños y cerrar las diferencias que existen debido a las desigualdades del desarrollo cultural y socioeconómico.

2. PLANNING GAMES FOR INTERVENTIONS

Mi único héroe en este lío.
–Patricio Rey y sus Redonditos de Ricota

Educational neuroscience investigates the development of mental representations, learning approaches and aims to improve different teaching approaches [103]. This discipline requires to collect data from children, which can be done in different manners.

The strategy used in the present work is to collect data from videos games. Why video games? Because playing is ubiquitous among humans since prehistoric times [134, 51]. Furthermore, playing is a significant contributor to the child's cognitive, physical, emotional, and social development [9]. And video games are not the exception, most of children like playing with a computer, and sometimes appear to maximize children's attention and prevent them from being distracted [143].

Since it is a common activity nowadays, analyzing video games playing behavior is one of the most ecological ways to collect data which can be use to study and improve the way children learn, solve problems and think.

Along their lives, humans face different problems. Some of the require planning ahead the steps needed to solve it. In Compute Science *planning is the model-based approach to autonomous behavior where the agent selects the action to do next using a model of how actions and sensors work, what is the current situation, and what is the goal to be achieved* [41]. Planning algorithms include some elements such as a value function, search depth and, stochasticity. We set to study planning strategies in children playing video games. To address this issue, this chapter focuses on understanding the problem structure –sequence of trials, difficulty progression, characterization of trials–. With a good understanding of the structure of a game, collected data will be more likely to be useful to explain human strategies.

We use two planing tasks: Tower of London and Chocolate Fix. The exploration of different sequences of trials based on users performance are enumerated, and tools and proposals are detailed to be useful in future educational interventions.

2.1 Planning game I: Tower of London

Tower of London is a planning task designed in 1982 by Tim Shallice [117]. It was originally aimed to measure the cognitive ability of planning [95, 68].

The rules are as follow:

- there are three beads: one red, one green and one blue;
- the beads are placed in sticks of unequal length. There are three stick which can handle different amount of beads (1, 2 and 3 respectively);
- the participant have to move the beads one at a time from a starting configuration to a target position in a minimum number of moves (Fig. 2.1 present a complete trial).

Since the objective is to solve it in the minimum amount of moves, a trial-error strategy does not work. The participant has to plan before hand what to do in each move, because any wrong move will lead to a failure in the trial.

To solve the puzzle of Fig. 2.1, if we consider the configuration of the left (S) as the starting position and the right (T) as the final configuration, the trial must be solved in five movements: the unique solution is Red from peg 3 to peg 1, Blue from peg 3 to peg 2, Green from 3 to 2, Red from 1 to 3 and Green from 2 to 3).

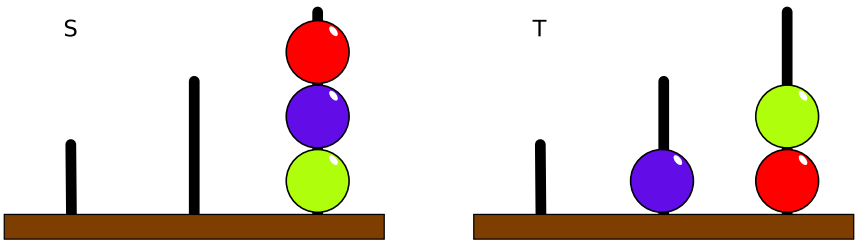


Fig. 2.1: Example of a trial of Tower of London. S: starting configuration, T: target configuration. This puzzle has a solution of five movements

The protocol used in previous works in our lab was made of 40 trials ordered by the amount of movements required to solve it. When inquiring how the sequence of trials was designed the only one dimension mentioned was the *number of moves* needed to solve each trial (distance among boards in the problem space, see Fig. A.26). This protocol used a cut-off of three wrong trials in a row (i.e. when the participant makes three consecutive wrong trials the test ends).

Data from an intervention on 2005 –519 children of an intervention were tested with Tower of London– showed an increase in the difficulty of the test along the trials (blue line in Fig. 2.2). Meanwhile, if the trials are grouped by the same distance, each group does not showed an increasing difficulty (red lines in Fig. 2.2), just the opposite.

This exploratory analysis arises two possible hypothesis:

- The participants learn how to solve puzzles of the same distance
- The list of trials is not properly arranged with monotonous increasing difficulty

This is sound with Newman and Pittman, who suggested that *using simply the minimum number of moves in Tower of London problems to characterize problem difficulty has to be reevaluated. This may be an important consideration given that many of the Tower of London tests used to assess executive functioning in clinical populations simply use the minimum number of moves as a measure of difficulty* [96].

To test the two hypothesis new data must be collected. Although this study is out of the bounds of the present work, the reason to include this section is to remark the importance of thinking and planing the structure and administration of the tests.

In this direction and with the objective of adding information for future studies, a list of structural information is provided here:

- The beads are from tree possible colors (Red, Green, Blue in the Figures)
- The beads can be order by reading them from the smallest stick to the tallest from bottom to top. The possible orders are the permutations of the colors Red (R), Green

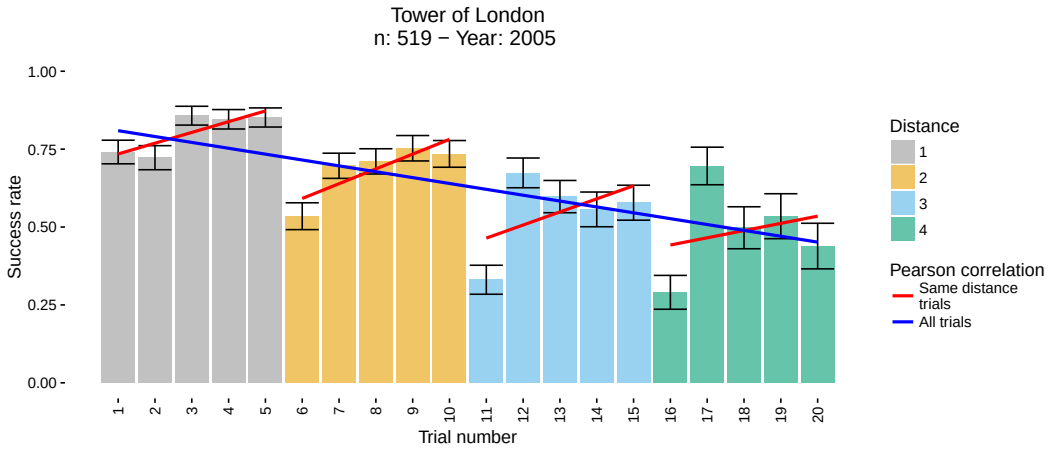


Fig. 2.2: Tower of London performance. Trials are colored by the distance needed to solve it. The blue line shows a negative linear relation among the number of trial and the success rate (Pearson Correlation $r(7747) = -0.212, p = 7.812E - 80$) –the global difficulty is increasing–. The red lines show the positive relation among the number of trial and the success rate for trials of the same distance (Pearson Correlations: distance 1: $r(2591) = 0.122, p = 3.528E - 10$; distance 2: $r(2345) = 0.142, p = 3.730E - 12$; distance 3: $r(1726) = 0.118, p = 7.193E - 07$; distance 4: $r(1079) = 0.065, p = 3.070E - 02$).

(G) and Blue (B): RGB, RBG, GRB, GBR, BRG, BGR. Six possible assortments of colors

- The board allow 6 possible layouts which can be classified in three groups: Tower (all the beads in the same stick), Partial (using two out of three sticks to place all the beads) and Flat (every sticks has a bead) [26]. See Fig. 2.3

- The number of possible boards are 36 (6 layouts $\times$ 6 color assortments)

- The combination of an assortment of colors and a layout gives a unique codification for each board. E.g. in Fig. 2.1 the board S has layout A and color: RBG, meanwhile board T has layout B and color BGR. For practical purposes a code number was assigned to each board: $layout \times 6 + colors\_order$, where layout are numbers from 0 to 5 in the order exposed in Fig. 2.3, and $colors\_order$ also numbered from 0 to 5 in the order mention in the previous item (In the example of Fig. 2.1, code for S is $33 = 6 \times 6 + 3$ because it is a tower and the arrange of colors is GBR, meanwhile code for T is $28 = 4 \times 6 + 4$ because it is a Partial layout and the arrange of colors is BRG)

- The number of possible trials is the combination of choosing 2 different configurations (source and target boards) out of 36 boards in both possible directions, i.e. $\binom{36}{2} \times 2! = 1260$ different trials

- If considering only one assortment of colors for the starting configuration, the number of boards is $6 \times 35 = 210$. Any other trial is isomorph to one of this initial set [65]

- The minimum distance among two boards is less or equal to 8. The number of different puzzles for each distance is summarized in Table 2.1

- Trials can have unique path solution or more than one way to reach the target in the minimum amount of moves. The number of solutions per distance is summarized in Table 2.1. The number of possible paths to solve a trial may influences its difficulty.

- The problem space has planar representation which can be observed in Fig. 2.4. Each node represents a board. Each edge represent a movement which transform a board into other. A trial is a pair of nodes from this graph. In the same Fig. the solution to the trial exposed in Fig. 2.1 is shown.

- The starting layout changes the number of possible first moves, i.e. starting a trial with a Tower layout allows the participant to pick rightly the first bead and then she has a 50% chance of placing it correctly, meanwhile in a Flat configuration of a trial with unique solution the participant has 33.33% of chance of picking correctly the ball, and then another probability of placing correctly. Starting with tower seems to be less likely to make wrong the first move using similar conditions

- The target layout influences the solution order: finishing in a Tower makes an unambiguous order to place the beads (firstly the one in the bottom, then the one in the middle and the last one is the one on the top), meanwhile Flat make an ambiguous order, there is no first look evidence that indicates which should be place first [26]

- Each movement can be characterized as: Goal (G), when placing a bead in its final position; Counterintuitive (C), when a bead has to be removed from its final position and; Intermediate (I), moves that are none of the previous categories. Labeling the solution with the kind of moves can also be used to characterize the difficulty of the trial [3, 63]

- The number of intermediate moves that have to be considered before the first goal can be reached has influence in the performance [124]

- The number of series of successive moves that are not goals can be consider as a difficulty predictor (number of chunks [3])

		Distance								Total
		1	2	3	4	5	6	7	8	
# Solutions	1	108	168	180	132	84	96	84	-	852
	2	-	-	36	42	84	66	30	12	270
	3	-	-	-	-	-	24	60	24	108
	6	-	-	-	-	-	-	-	24	24
	8	-	-	-	-	-	-	-	6	6
Total		108	168	216	174	168	186	174	66	1260

Tab. 2.1: Tower of London: number of available trials for different distances and number of possible solution paths

In the example of trial exposed in Fig. 2.1, considering the trial as exposed from board S to board T, may not have the same difficulty playing it backwards (from S to T). In table 2.2 the comparison among forward and backwards trials is detailed.

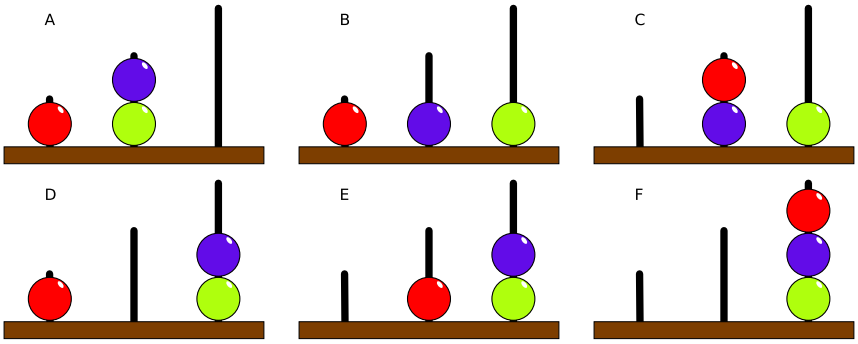


Fig. 2.3: Possible layouts of Tower of London with all the colors ordered in a sequence of Red-Blue-Green if read by stick (from smallest to the tallest) from bottom to top. There can be consider in three categories: Tower (F), Partial (A, C, D, E) and Flat (B).

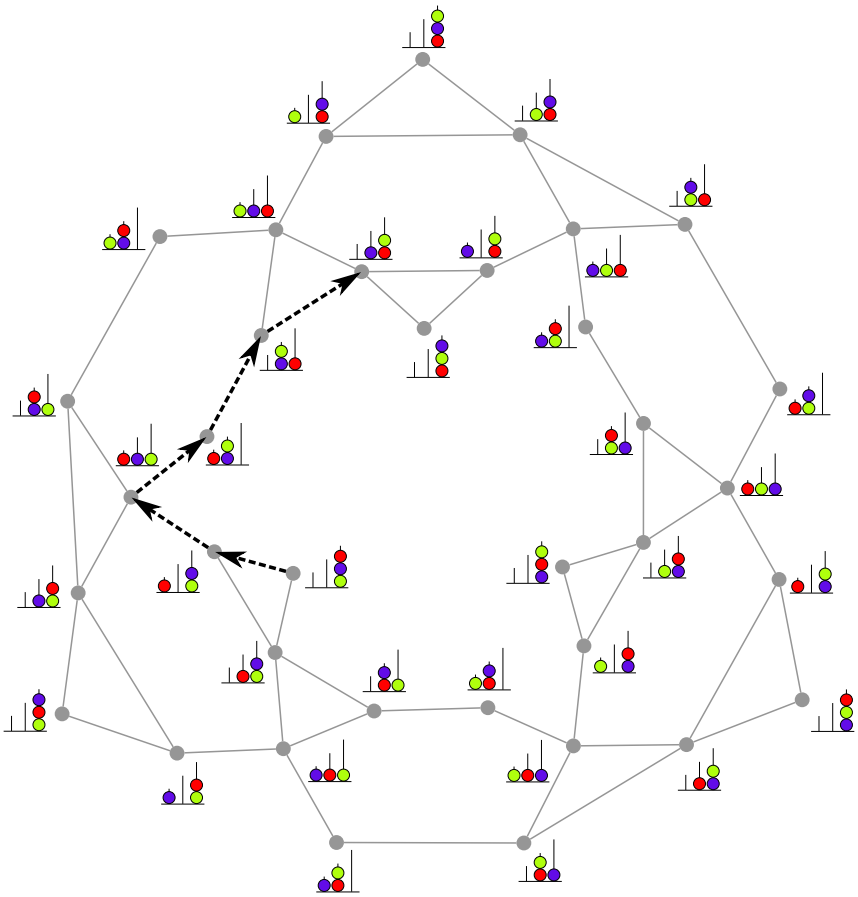


Fig. 2.4: Problem space of Tower of London. Each node represent a board, and there is an edge between two boards if one valid move changes form one to the other. The dashed line represent the sequences of moves to solve the trial of Fig. 2.1

As a first and quick answer to the question "Are both trials, from S to T and from T to S, of the same difficulty?", the table shows some coincidences, but also some differences. It might be needed to reflect a bit before getting an answer.

The following testimony of an operator who administered the test can help to under-

Characteristics	From S to T	From T to S	Are equal?
Length	5	5	Yes
# different paths	1	1	Yes
# first possible ball	1	2	No
# first move possible results	2	4	No
Initial configuration	Tower	Partial	No
Final configuration	Partial	Tower	No
Path characterization	IGIGG	IIGGG	No
# moves to first goal	1	2	No
# chunks	2	1	No

Tab. 2.2: Tower of London: trial analysis of the example of Fig. 2.1 forward (from S to T) and the difference with backwards (from T to S)

stand:

> *When the child is very impulsive (with low inhibitory control), she has to take a ball in her hand to start "playing" and then think ahead where to place it. When multiples balls are available to pick in the first movement, sometimes she picked the wrong one.*

Although this is a one-subject experience, it throws some light in the topic of this section. This test might be problematic to asses children with low inhibitory control in a planning task. This issue could be solved if the development of inhibitory control is considered when choosing the sequence of trials of the protocol.

In addition, the *three in a row wrong* cut-off seems to be a fair rule to avoid frustration, but it does not let the participants complete the whole protocol. Is this correct? How can this be fixed to avoid frustration, asses planing and have the same amount of trials for each participant?

To conclude this section, a proposal of a new protocol is presented. Unfortunately there is no collected data to analyze the performance of this protocol. This sequence aims to test if children's performance have similarities with a random agent play.

The protocol has the next properties:

- It will include 5 trials of each distance: 1, 2, 3,...,8 (40 trials)
- It will include trials with only one possible path, except for those of distance 8 which do not have trials of unique solution
- To maximize the exploration of the problem space, only 6 boards will be consider as source (one of each layout), leading to $6 \times 35 = 210$ puzzles
- For each trial colors will be assigned randomly as a transformation to avoid visual repetitions in the source board. A color transformation could be the identity function, or swap Red with Blue, or all the colors. The properties of the board will be the same.
- For each possible trial, the win probability of a random agent is calculated. The agent has memory: it does not visit the same board twice in one solution.

- Each trial belongs to a pair (distance, probabilities of winning).
- The 5 trials are split in subsets of decreasing probability of winning by chance. The subset size are as similar as possible, if they can not be equal, higher chance are preferred. For each subset trials are randomly selected and a color transformation is selected. I.e. distance 1 has three probabilities: 0.5, 0.33 and 0.25. So, there will be 2, 2 and 1 trials of each. The available boards for probability 0.5 are 4, for 0.33 are 6 and, for 0.25 are 8.

An example of the protocol can be observed in the Table A.1 of Appendix A.1.

Future iterations of the protocol will include *relax trials* (easier trials in the middle) to avoid frustration of the participants and let them get further reducing the cut-off criterion usage. It also should include the characteristics not included in this one but listed before. Each of this improvements must include data collection and analysis. A better understanding of the tool will allow researchers to have a better assessment of the population.

2.2 Planning game II: Chocolate Fix

Chocolate Fix is a board game of the company Think Fun[1] where the player works in a candy shop completing orders from customers who request different assortment of chocolates (this game was used in the study included in Chapter 6).

The game has a story: the player works in a Candy shop where customers have strange wishes for their boxes of chocolates. The tray to be filled is a 3 by 3 grid, and it will contains always the same 9 chocolates arranged in some way. The chocolates are all the combinations of 3 flavors (chocolate, strawberry and vanilla) and 3 shapes (circle, square and triangle).

Each trial has a unique solution which consists in filling the tray with an assortment of all the nine chocolates that accomplish every customer requirement. The requirement is presented as one or more fragments –hereinafter referred as cues– of the tray where each position could be: a piece, an attribute of a piece (flavor or shape) or an empty spot. A cue is consider accomplished if it can be place over the board in a position where the pieces match with the information of the cue. i.e. over a strawberry triangle chocolate a cue can either have the same piece, a triangle shape, a strawberry flavor sign, or an empty spot to be considered a match. A cue cannot be rotated, but it can be placed in any place of the tray, e.g. a 2 by 2 cue has 4 possible positions in the board, a 3 by 2 cue only 2, and a 3 by 3 cue has a unique position.

Since the game is composed by a board, pieces (made of two attributes, flavor or color and shape) and cues (which are subpart of a board with pieces or attributes in it), it can be translated to numbers for the complete pieces and letters for individual attributes (see Fig. 2.5 for details). Using these codes, a trial was defined as a 9 characters string for the board (`bbbbbbbbb` for the empty one or with other characters for non empty ones) and as many sequences of 9 characters as cues has the board.

Using this string representation for the puzzle it was possible to develop a puzzle builder in a web page (see Fig. 2.6) which allows researchers to design new puzzles and check if they have a unique solution (or more than one or none). This tool also includes the

[1] Used with the permission of Mark Engelberg and Bill Richie.

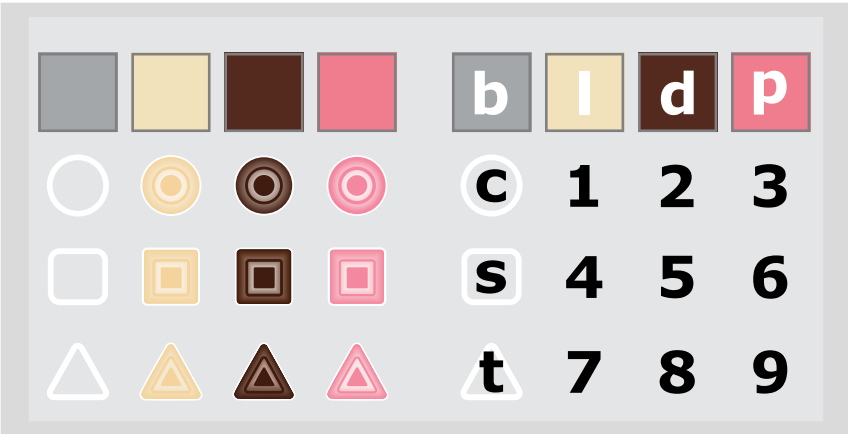


Fig. 2.5: The codification of the pieces and attributes. Complete pieces are numbered from 1 to 9. Color attributes are l (light), d (dark) and, p (pink). Shapes are c (circle), s (square) and, t (triangle).

option of transforming puzzles to generate isomorph puzzles. The transformation included are:

Transformation: Rotate 90°, Rotate 180°, Rotate 270°, MirrorH (Horizontal Mirror, exchange top row with bottom row), MirrorV (Vertical Mirror, exchange left column with right column) and, Swap Shape and Color (exchange every shape for a certain color, and every color for a certain shape, it can be visualize as the transpose of the numbers of the matrix on the right of Fig. 2.5)

Swap shape: Circle-Square (exchange every circle for a square, and every square for a circle), Circle-Triangle and, Square-Triangle

Swap color: Light-Dark (swap light attributes with dark), Light-Pink and, Dark-Pink

In the way to study the algorithms used by the participants to solve the puzzle, one possible way is observing trajectories [14]. In our game, as an exploratory analysis we choose to develop visualizations of the movements used to solve each trial.

The puzzle of Fig. 2.7 was explored in a dataset of children from 6- to 10-years-old and produced the trajectories exposed in Fig. 2.8. Each square is one of the different possible board states while solving the puzzle (initial board in the first row, boards with one piece place in the second row and full-filled boards in the third row), the boards are linked by edges, as a direct graph, labeled by a piece and proportion of how many of the participants who where in that state decide to took a certain transition –which is also reflected in the width of the line–. The visualization of this puzzle lets the observer notice (see Fig. 2.8) that the preferred first move is with the strawberry circle (pink circle): 144 out of 156 made this choice. The first move with the preferred piece lead to two different board states: 75 participants place it correctly (third board of second row) meanwhile 69 place it wrong (first board of second row).

This technique was particularly useful to get insight of the data in the study presented in Chapter 6, where we analyze small space problems like the one of Fig. 2.8. The visualization was possible because the possible board states are reduced to 7 possible boards: the initial board, 4 possible boards with one piece placed (each of the two pieces can be

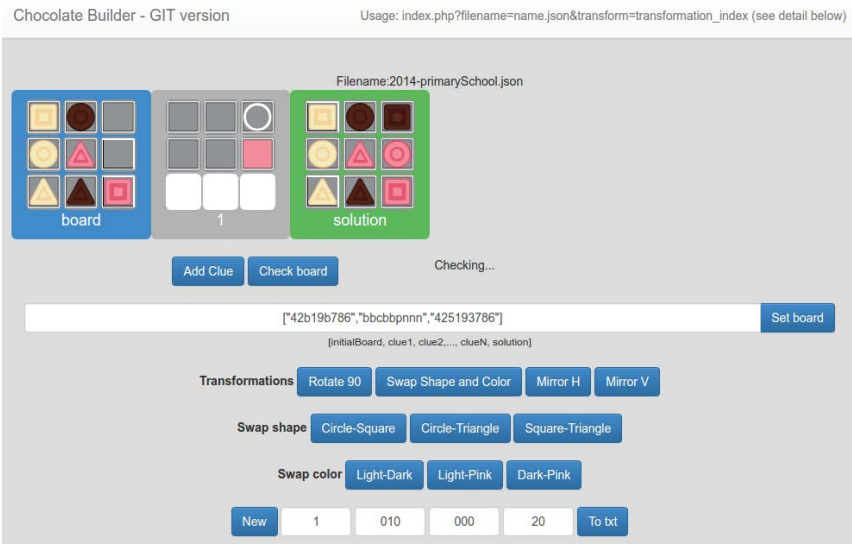


Fig. 2.6: Screen shot of the puzzle builder app. It includes buttons for manual transformation and the string representation of the board.

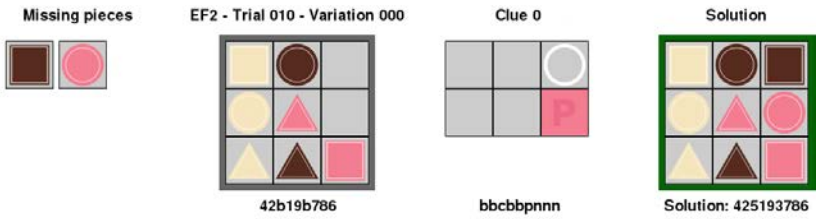


Fig. 2.7: Chocolate Fix puzzle with only two missing pieces (in the right), the initial board, the cue and the final solution. To check the solution the cue must be place in the bottom of the board, where the circle attribute it is accomplished by the strawberry circle and the strawberry flavor is accomplished by the strawberry square.

placed in two positions) and the board with all pieces placed (the correct solution and a wrong solution).

However, not all the puzzles can be visualized in the same way. For example, the one of Fig. 2.9 produced the trajectories of Fig. 2.10 which is much bigger than the previous one. Since this trial starts with an empty board, in the second row there are 81 potential boards –the 9 possible first pieces in the 9 possible spots of the tray–. After this first move, the number grows which made it harder to extract ideas on the first sight. Although some things can be observed: there is only one full-filled board, and it is correct –no one full-filled wrong, it seems to be an easy puzzle–; the two most visited paths seem to be one with around 50 participants –starting with the vanilla circle–, and the other with around 30 out of 150 participants –starting with the vanilla triangle–, both initial pieces are the top most piece of each cue.

This huge board space will be interested to study when collected data is not limited to a few hundreds of participants. It is interesting to see which percent of the solution path space is cover while solving trial of Fig. 2.10 in comparison with all possible trajectories states (see in Table 2.3).

Along this work, many puzzles were designed to generate a sequence of levels according

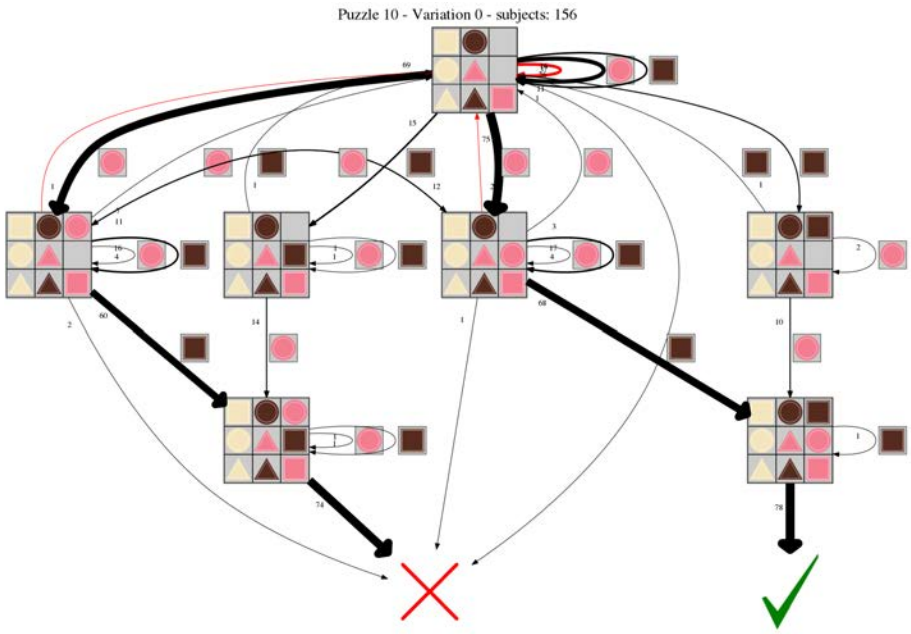


Fig. 2.8: Chocolate Fix solve tracing graph example. The top board is the initial one, it starts pre-filled with only two empty spots. The second row of 4 boards represent the result after placing the first piece. The third row represent the two possible full filled. The fourth row are the possible results of the trial, participants receive the result when they click a *Check* button. In the bottom of the image the puzzle is shown: the missing pieces, the initial board, the cue and the solution. Red lines represent the *Reset* button click.

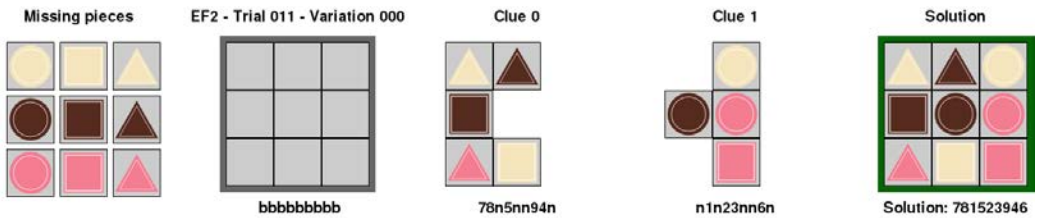


Fig. 2.9: Example of puzzle which visualization size explodes due to the number of possibilities.

to the different populations. Next a summary of the levels designed is presented:

Primary School 2014: 6- to 10-years-old participants in a 6 session experiment. 64 trials ordered by increasing complexity. The dimensions of complexity had different ranges (ordered from simpler to complex):

- Board pre-filled to empty
- Missing pieces from 1 to 9
- Cues of 3-by-3 to smaller cues
- Cues information disjoint to cues that must be integrated
- Cues with full pieces to cues with partial attributes
- Puzzles with information to place every piece to puzzles were pieces are place by discard

Fig. 2.10: Example of puzzles whose visualization size explodes due to the number of possibilities.

Row/Number of pieces placed	Number of boards	Possible number of boards
0	1	1
1	9	81
2	20	2592
3	20	42336
4	23	381024
5	22	1905120
6	25	5080320
7	26	6531840
8	10	3265920
9	1	362880
Total	157	17572114

Tab. 2.3: Analysis of the visualization of the data solving the puzzle of Fig. 2.9 and the comparison with the possible full cases. Each row of the visualization correspond to a different number of pieces placed. The amount of possible boards for each row is calculated as $\binom{9}{row} \times \binom{9}{row} \times row!$ (with row as the number of row).

- Cues can be selected in any order to cues that must be selected in a certain order

To generate engagement in the participants a score was included. The trials were grouped in levels of 6 trials each. And the score of each level gives $10 \times fibonacci(level\_number + 1)$. i.e. level 1 gives 10 points, level 2 gives 20, ..., level 4 gives 50, and so on...

Berkeley Eye Tracking 2014: undergraduate students of UC Berkeley. 48 trials in 1 session. The objective was to explore how small manipulations of the puzzles impact on the performance of the participants. Using cues with only partial attributes, the manipulations were:

- Integration: 2 cues, 3 pieces missing, 1 piece place by process of elimination. The small manipulation was produced by moving an attribute in one cue, re-

ducing the ambiguity. Two kind of puzzles: high and low integration

- Distraction: 2 cues, 2 pieces missing, 1 piece place by process of elimination. The small manipulation was produced by a cue which is similar to the empty spots of the tray and it as attributes of the missing pieces (high distraction), the control condition has a cue already accomplished in the tray but not similar to the empty spots.
- Redundant: 3 cues, 3 pieces missing. One condition's puzzle has one of the cues which is enough to solve it, meanwhile the other conditions at least two out of the three cues were needed to solve it.

A score of cues was designed with the objective to measure the information provided in each one. The formula to get the score is:

$$cue\_score = (5 \times \#pieces + 2 \times (\#color + \#shape)) \times (width + height - 2)$$

with #pieces being amount of complete pieces of the cue; #color the number of color cells in the cue; #shape the number of shapes of the cue and; width and height which refer to the size of it.

An example of each category of puzzle is included in Appendix A.2

Primary School 2016: 4 sessions for children from 5- to 7-years-old in Argentina and from 6- to 9-years-old from Colombia. Similar to the previous primary school edition but adapted to be play in 4 sessions. Around 80 participants.

Uruguay 2016: 12 session for 5 years old children (around 200 students). Trials are easier:

- Level 1 (trial 1 to 8): 1 cue, 1 missing piece
- Level 2 (trial 9 to 12): 1 cue, 2 missing pieces
- Level 3 (trial 13 to 16): 1 cue, 2 missing pieces, 1 piece place by process of elimination
- Level 4 (trial 17 to 20): 2 cues, 2 missing pieces
- Level 5 (trial 21 to 22): 2 cues, 3 missing pieces

from session to session a transformation was applied to make the trials seem different.

The last 3 set of puzzles were not analyzed yet.

The sequence of trials a participant faces and the order they are presented was study from game perspective [14] to Intelligent Tutor Systems [18]. A long the last 5 years the following strategies were explored:

- Sequence of trials: when the participant presses check and the trial is incorrect the game let the player fix it and check it until it is correct. When it is correct the game moves to the next sequence. Checking in wrong configurations can be used as a strategy to game the system, but this could be avoided if points are scored on perfect solutions, or points are subtracted when checking wrong solutions. This strategy is easy to read it in the pseudo-code 2.1
- Sequence of trials with no fixing: the same as before but only one possible fix (or none) is allowed.

- Grid of trials: in a graphical representation, the first row is a sequence similar to the previous ones, the columns are the result of applying some randomly selected transformations applied to the trial in the first row. The participants will start playing in the first row, if the trial is correct, next trial is the first row of next column. If it is incorrect it goes back one column, but it moves to a row she has not played yet. In other words: if it is correct it moves to the next (right) column to the highest un-played cell, if it was incorrect it moves to the previous (left) column to the highest un-played cell. For a graphical example see Fig. 2.11 where two chances per trial are allowed.

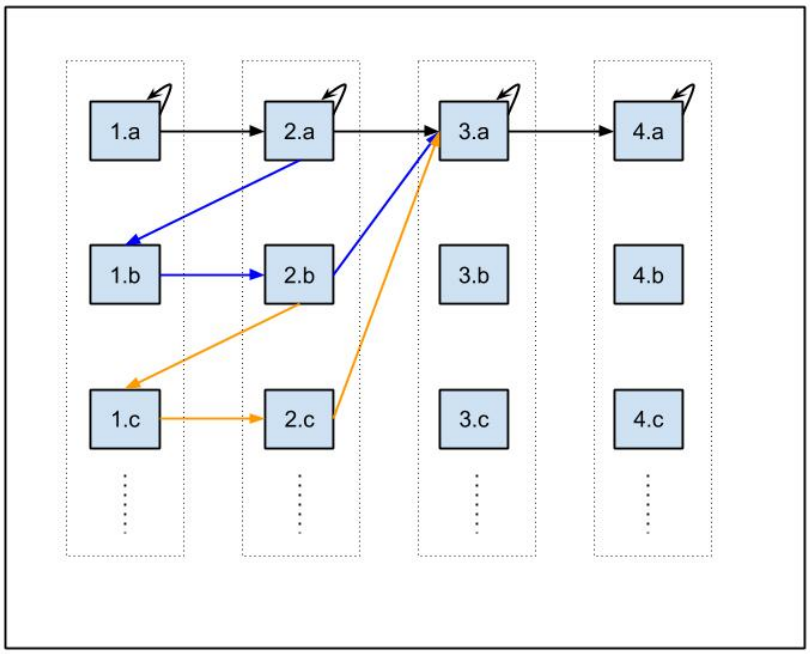


Fig. 2.11: Sequence of levels. Each column has similar puzzles, each one was made with a transformation (like rotation, changing colors, etc). Each time the participant lose a trial, the next trial is a variation of the previous one that she had not play yet. In the figure, the black arrow represent a player who never loses. The blue one, is the variation if the player loses in trial 2.a. And the orange arrow represents a player that loses trial 2.a, then win 1.b and then loses 2.b, then win 1.c and 2.c and then she join the original path. Every trial has 2 chances before going back to the previous one. All the player played always the trials with letter "a".

```
cur_trial = 1 # current trial number
On Check trial:
    If correct:
       score = score + level * 10
       cur_trial = cur_trial + 1
    Else:
       score = score - level
       cur_trial = cur_trial
```

Code 2.1: Next trial selection for the straight forward strategy.

```
cur_trial = 1 # current trial number
cur_variation = 0 # current variation number
On Check trial:
  If correct:
    cur_trial = cur_trial + 1
    cur_variation = least_times_play_variation(cur_trial)
  Else:
    If it is first try:
      // try again
    Else:
      cur_trial = cur_trial - 1
      cur_variation = least_times_play_variation(cur_trial)

# least_times_play_variation has access to what the player
# have played and returns un-played variation of the number
# passed as parameter
```

Code 2.2: Next trial selection for the grid of trials strategy.

Unfortunately we could not study differences among these strategies due the population where each strategy was try were different. It will be part of future work to explore if these strategies change results within the same population.

Resumen en español

Juegos de planeamiento para intervenciones

La neurociencia educativa investiga el desarrollo de las representaciones mentales, los enfoques de aprendizaje y apunta a mejorar diferentes enfoques de enseñanza [103]. Esta disciplina requiere recolectar datos del modo en que los niños aprenden, y esto puede llevarse a cabo de diferentes formas.

La estrategia utilizada en el presente trabajo es recolectar datos de la interacción de los niños con videojuegos diseñados para tal fin. ¿Por qué los videojuegos? El juego es parte de la cultura desde tiempos prehistóricos [134, 51]. Además, los juegos son un factor importante que contribuye al desarrollo cognitivo, físico, emocional y social de niños y niñas [9]. Los videojuegos no son la excepción, a la mayoría les gusta jugar con una computadora, llegando incluso a niveles de atracción en los que los videojuegos parecen capturar toda su atención y hacerlos ignorar lo que sucede a su alrededor [143].

Como los videojuegos pueden ser catalogados como una actividad cotidiana, se transforman en una de las formas más ecológicas de recopilar datos para estudiar y mejorar la forma en que los niños aprenden, resuelven problemas y piensan.

A lo largo de sus vidas, los seres se humanos enfrentan problemas diferentes. Algunos de ellos requieren la planificación por adelantado de los pasos necesarios para resolverlos. En Ciencias de la Computación, *la planificación es el enfoque basado en modelos de comportamiento autónomo en el que el agente selecciona la acción a seguir usando un modelo de cómo funcionan las acciones y los sensores, cuál es la situación actual y cuál es el objetivo a alcanzar* [41]. Los algoritmos de planificación incluyen algunos elementos como una función de valor, profundidad de búsqueda y aleatoriedad. En este capítulo nos enfocamos en las estrategias de planificación que despliegan los niños y niñas jugando videojuegos. Para ello, nos centramos en la comprensión de la estructura de los problemas a estudiar –secuencia de ensayos, progresión de la dificultad y caracterización de las secuencias de niveles–. Una buena comprensión de la estructura de un juegos permite que los datos recopilados tengan más probabilidad de ser útiles para explicar estrategias y métodos humanos para resolver estas tareas.

Concretamente, se analizan dos tareas de planificación: Torre de Londres y Chocolate Fix. A partir de ello, se enumeran las secuencias de ensayos basadas en el rendimiento de los usuarios. Finalmente, se desarrollan herramientas y propuestas para futuras intervenciones educativas.

3. CELL-PHONES INFLUENCES IN FACE-TO-FACE INTERACTIONS

Y aquí gracias a Dios uno no cree en lo que oye.
–Patricio Rey y sus Redonditos de Ricota

The world has changed since cell-phones became popular. In the last 20 years, Argentina changed from having less than three cell-phone every 100 people to more than 1 cell-phone per person in 2015, according to a report of the International Telecommunication Union, World Telecommunication [57]. This is a world phenomenon as shown in Fig. 3.1 and Argentina's ratio is bigger than World's ratio since 2003.

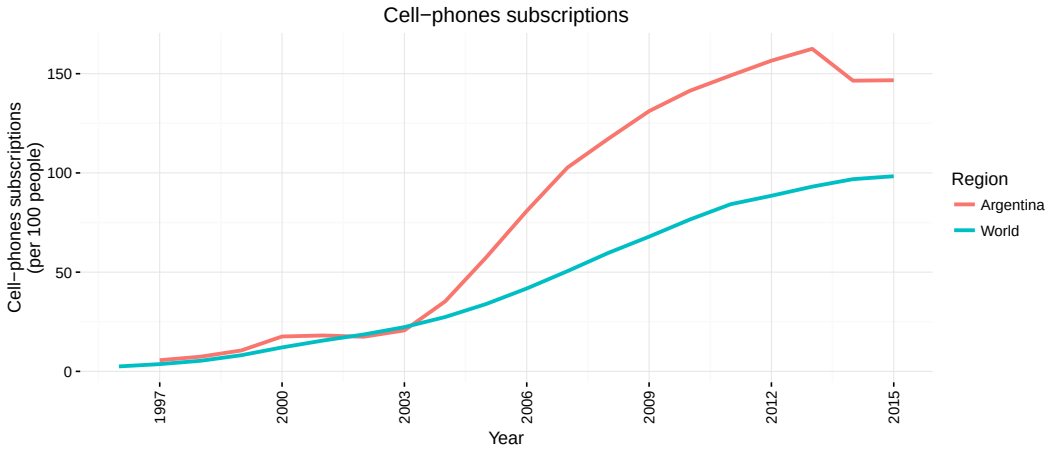


Fig. 3.1: Number of cell-phones every 100 people. Argentina's proportion of phones and people is over 1 since 2007. World's proportion is almost 1 to 1 for the year 2015.

In this chapter we explore how cell-phones affect the face-to-face talk. We designed a protocol where in a few minutes of interaction dyads had to interact under different conditions and report perceptions and feelings. The experiment was done in two editions of TEDx Talks (TEDxRíodelaPlata and TEDxRosario) and it was a joint effort of biologists, computer scientists, designers and hosts and volunteers of TEDx Talks.

This work contributes in to the assessment of the impact of including ICT in informal contexts. Further work should study if the impact is the same in formal environments, such as schools.

This chapter was partially published in the *Journal Plos One* on June 3, 2015 [83].

3.1 Introduction

Cell-phones are ubiquitous, with more than 300 million users in the United States and more than 6.800 millions world-wide [135]. This pervasive device opens a window to understand multiple aspects of human society. Several studies have analyzed cell-phone calls as large-scale social networks, describing communication patterns and network structure e.g. [73, 97, 71, 98, 61]. Communication data has been studied to understand human mobility

patterns [45, 122, 85]. These data also offers an indirect measure for social interactions – showing non-Poissonian bimodal interevent distributions [139, 118, 64] – and where relationships between people may be infered from reciprocal calls patterns [32, 16, 70].

These studies focus principally on cell-phones as a tool for quantitative sociological research to understand different aspects of human behavior. Instead, we investigate the impact that cell-phone distraction may have in human relations. Cell-phone use impairs attention even to gazed elements of the visual scene [127], which has been recognized as a major risk factor during driving [90]. Because this has become a major safety issue, the vast majority of research on inattention due to cell-phone use has concentrated on its implication on driving deficiencies.

However, the consequences of inattention are obviously not only specific to driving. For instance, cell-phone users walk more slowly, change direction more frequently, and are less likely to acknowledge other people [56]. In fact, many people sense that using cell-phones and other electronic devices may have a strong cost on how we communicate and relate to each other. However, this cost has not been thoroughly studied empirically. The goal of this paper is to solve this empirical gap determining - in a quantitative manner - how human social interactions are affected by frequent interruptions based mostly on the use of electronic devices.

To this aim, we performed a two-player social game, with a large sample (N=713 couples) playing simultaneously in two different experiments performed in theaters, during TEDx events. Each player of the dyad was assigned a different role. The speaker was asked to narrate a very engaging four-minutes story. Some listeners were given instructions to pay full attention to the speaker. Other listeners had to ignore the speaker mainly using his cell-phone (text messages, Twitter, etc.). Finally, others listeners had to change their attitude from full attention to no attention for different time periods during the four-minute exercise. The speaker and the listeners only read the instructions of their role. However, the instructions of the listener included also a description of the speaker task. Thus, the listener is aware of all aspects of the experiment, while the speaker is not informed of the listener's role. Following this treatment, we measured participant's beliefs about the quality of the story and about the conversational partner.

3.2 Methods

Experiments were performed on a large audience in two different theaters, in TEDx events. This experiment is part of an initiative referred to as TEDxperiments which aims to capitalize on TEDx events to construct knowledge on human communication. The first experiment was performed on September 27, 2013, in Buenos Aires, with an audience of 1200 people at TEDxRíodelaPlata (`http://www.tedxriodelaplata.org`). The second experiment was performed on October 9, 2013, in Rosario, with an audience of 900 people at TEDxRosario (`http://www.tedxrosario.com.ar`).

The theater research assistants handled the material (paper and pencil) to participate in the experiment. Participants were informed that participation in the experiment was completely voluntary and they could simply choose not to participate. Participants provided a verbal consent. They simply responded to the research assistants the willingness to participate. Due to the brevity of the experiment, participants did not sign a written consent form. Participants were explicitly assured that 1) they participation in the experiment was completely voluntary and that they could leave the experiment at any time

and 2) that all the data was completely anonymous. Anonymity was assured since the questionnaires filled by participants did not have any personal information (name, age, gender) and were dropped in a common box. The procedures of the experiments described here were approved by the ethics committee of CEMIC (Centro de Educación Médica e Investigaciones Clínicas Norberto Quirno).

Each player was paired with the person sitting in the next row of the theater (to make it more probable that people would play with someone they did not know beforehand) and was given an envelope containing the instructions for two roles that were assigned randomly (instructions are presented in Appendix A.3. Players were asked to open their role and then the game commenced. Videos of the game (see `http://www.tedxriodelaplata.org/videos/tedxperiments`) reflect a very strong commitment of players, in both theatres. Participants were assigned randomly to 6 different groups: 10% of the players were in the full attention group (the instruction was to pay attention all the time), another 10% were the none attention group (they were asked to ignore their conversational partner all the time). 80% of the participants were uniformly divided in four groups in which the listener ignored the speaker for a total of two minutes (half of the total story duration). The four groups were assigned different temporal patterns labeled $[(++--), (--++), (+-+-), (-+-+)]$, where each symbol denotes a minute of the dialog and plus and minus signs index whether the listener attends the speaker $(+)$ or not $(-)$. After playing the game, players had one minute to fill an anonymous questionnaire which responding in range of 1 to 10 (1 is the minimum and 10 the maximum)

Speaker Questionnaire:

- About the story: whether they believed the story they told was entertaining (Question 1, Q1) and emotive (Q2).
- About the conversational partner: whether their partner was an interesting (Q3), attractive (Q4) and enjoyable (Q5) person; whether the way they told the story was effective (Q6), recited fluidly and with good rhythm (Q7) and well perceived by the listener (Q8).

Listener Questionnaire:

- About the story: whether they believed the story they heard was entertaining (Question 1,Q1) and Emotive (Q2) and told fluidly (Q3).
- About the conversational partner: whether their partner was an interesting (Q4), attractive (Q5) and enjoyable (Q6) person.

The speaker and the listeners only read the instructions of their role. However, the instructions of the listener included also a description of the speaker task. Thus, the listener is aware of all aspects of the experiment (that the speaker is asked to tell a very important story and that they should ignore the speaker during specific times). Instead, the speaker is not informed of the listener's role and has no way to know it. To assure that they did not guess that the listener was acting a role, we asked a random sample of participants (N=170) whether they realized that the listener was acting a role. Only 3 participants ($< 2\%$) responded positively.

Only pairs of players for whom we had all responses complete were considered for analyses, to assure that all comparisons could be paired. This left us with a total of 414 pairs in the first experiment and 299 in the second experiment.

3.3 Results

After playing the game, the players completed a questionnaire responding several questions (on a 1-10 scale) about how they perceived their conversational partner and the story (see methods for a full description of the questionnaire).

Is there a bias such that either speakers or listeners tend to judge the story and the partner differently?

Listeners had a better opinion of the speaker than vice versa (average partner perception: speaker 6.68 ± 0.08, listener 7.53 ± 0.07). Similarly, the story (its quality, whether it was entertaining, recited fluidly) was better ranked by the receiver (average story perception: speaker 5.85 ± 0.09, listener 6.67 ± 0.09). These results were confirmed by a paired t-test comparing the scores within each speaker-listener pair which showed a highly significant difference (story: $t = 9.08$, $df = 665$, $p < 1.18 \times 10^{-18}$, Cohen's $d = 0.42$, partner: $t = 7.91$, $df = 615$, $p < 1.18 \times 10^{-14}$, Cohen's $d = 0.41$).

Are speaker judgments determined by attention time?

The speaker's judgments of the quality of the story they told and about their conversational partner increased markedly as the time of attention by the listener augmented (Fig. 3.2, top panels a and b).

Are listener judgments determined by attention time?

Analogously, the listener's judgments of the story they listened to and their opinion about the speaker increased with the time of attention (Fig. 3.2, bottom panels c and d). This effect was more surprising since the listener was aware that ignoring the story was only due to task instructions and not because she considered the story or the conversational partner uninteresting. To confirm these results we submitted the data to two independent ANOVA, one for the story and one for the partner scores, with role (Speaker or Listener) and time of attention (0, 2 or 4 minutes as main factors). Both ANOVAS showed highly significant effects of the main factors (time and role) without interaction (Table 3.1). These results were replicated when analyzed for each of the two independent experiments (Table 3.3)

	Story			Conversational Partner		
Factor	F	df	p	F	df	p
Role	16.47	2	$\mathbf{< 10^{-8}}$	12.81	2	$\mathbf{< 10^{-8}}$
Attended Time	26.53	1	$\mathbf{< 10^{-8}}$	34.04	1	$\mathbf{< 10^{-8}}$
Interaction	1.89	2	0.15	1.5	2	0.22

Tab. 3.1: ANOVAs for the story and for the partner scores, with role (Speaker or Listener) and time of attention (0, 2 or 4 minutes as main factors).

Which judgments are affected by attention time?

The questionnaire asked participants about different dimensions of the story and how it was told; including its flow and rhythm, whether it was entertaining and how emotional was its content. To investigate the effect of attended time in each question, we performed independent linear models (one for each question) with score as dependent variable and

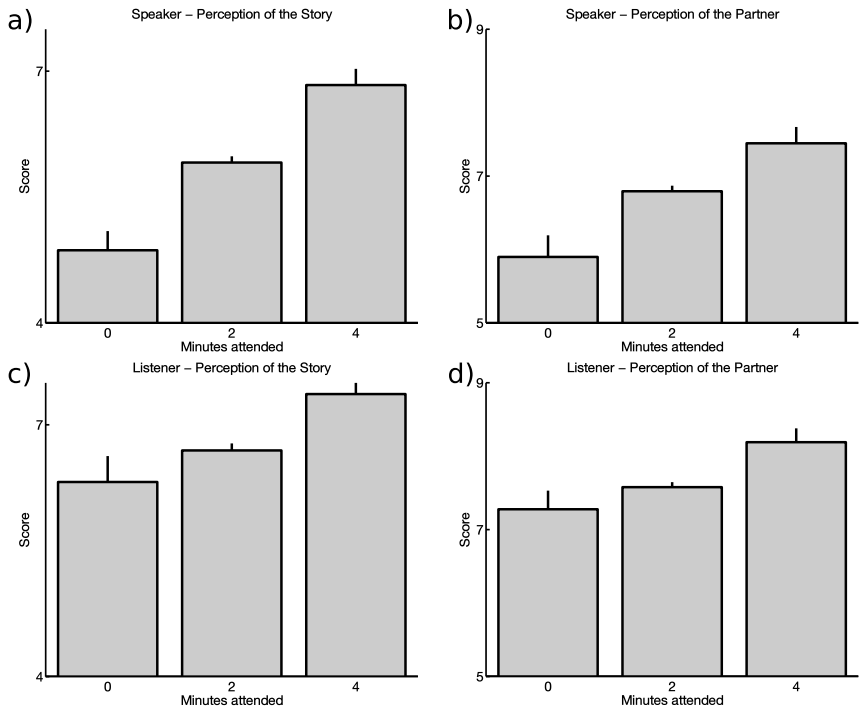


Fig. 3.2: **Speakers and Listeners judgments on quality of story and conversational partner increase as minutes of attention augment.** a) The speaker's judgments of the quality of the story. b) The speaker's judgments of their partner. c) The listener's judgments of the story. d) The listener's opinion about the speaker.

	Speaker					Listener		
Params	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3
α	4.74	5.35	4.72	5.23	4.19	6.14	5.67	6.37
β	0.29	0.18	0.56	0.51	1.02	0.44	0.10	0.36
PP	**0.01**	0.16	$\mathbf{< 10^{-8}}$	$\mathbf{< 10^{-8}}$	$\mathbf{< 10^{-8}}$	$\mathbf{< 10^{-8}}$	0.45	**0.0041**

Tab. 3.2: Linear regression of questions only about the story for speakers and listeners.

attended time as the main regressor. The sole score which did not increase monotonically with time of attention was the emotional content (Question 2, see Table 3.2).

Are judgment affected by whether attention is deployed at the beginning, at the end, or alternating throughout the story?

The condition in which the listener ignored the speaker for two minutes (half of the total story duration) was organized in four different temporal patterns labeled $[(++--), (--++), (+-+-), (-+-+)]$, where each symbol denotes a minute of the dialog and plus and minus signs index whether the listener pays attention to the speaker $(+)$ or not $(-)$. The aim of this factorial design is to investigate whether the dynamics of deployed attention (for fixed attended time) is pertinent for subjective constructs of the success of human communication. This experimental design controls in a factorial manner the durations of moments of attention and inattention (interval duration 1 or 2 minutes) and whether attention is deployed at the beginning and then fades out or conversely whether the listener first ignores the speaker and then attention grows (interval order: pay attention first or

pay attention last).

We submitted the subjective scores of conversational partner and story perception to independent ANOVAs with order and duration and their interaction as independent factors. The data consistently showed that none of the subjective scores were sensitive to the effect of order or duration (Table 3.4, none of the eight effects of interest reached significance or even reached marginal significance). Of course, it is impossible to discard a residual small effect size. However, the fact that this experiment is performed over a very large sample ($N = 713$ couples), and that the effect of attended time reached very high levels of significance, consistently in two independent experiments, suggests that overall the variance in subjective perception of human communication is largely determined by the total amount of attention and not on how it is distributed.

3.4 Discussion

We investigated how neglect resulting from use of mobile devices affects human social perception. Our results show that speakers grade the quality of the story and the sympathy towards the listener in proportion to the time they are paid attention. The quantitative measure of this results points out that 1) this is a highly significant effect (perception of the story by a speaker changes from $\approx 4.86/10$ to $\approx 6.83/10$ from no to full attention) which is consistent across experiments and reaches very high levels of significance and 2) the effect is mostly independent on whether inattention alternates in time or is collapsed in long lasting episodes.

More surprisingly, the listener qualification of the story and the sympathy they report towards the speaker also varied with the amount of attended time. This result was less expected to us prior to the experiment because the listener is fully aware of the fact inattention solely responds to task instructions and not to the merits of the story. The current data cannot fully narrow the principles leading to this observation, but there are two parsimonious (and not mutually exclusive) explanations. First, this observation naturally results as a way to avoid cognitive dissonance [38]. Ignoring and praising someone at the same time (even if ignorance is presumably not related to the merits of the speaker) yields to two contradictory beliefs and values, which is known to be a source of cognitive stress and discomfort that tends to be avoided implicitly [38]. A similar interpretation of this finding is that it may be affected by demand effects, usually defined in economic experiments to refer to changes in behavior by experimental subjects due to cues about what constitutes appropriate [142]: the listener may think that if the experimenter asks him/her to ignore the speaker, a response of less interest is expected.

Second, it is possible that upon being ignored, the speaker changes his attitude, gives up on the story and makes it worse in which case the receiver's scores would reflect a genuine deterioration of the quality of the story (and probably of the way the speaker looks) as they progressively remove attention.

Studies examining children raised in severe cognitive neglect have consistently shown marked cognitive development deficits related to major decreases in individual attention and emotional affection [101]. Our study can be seen as a way to address the effect of inattention in subjective beliefs in a much more frequent and less extreme condition: a person being partially ignored while telling a very meaningful story.

One consequence of inattention and distraction during human communication is the disruption of ostension. Ostensive signals, which include among others directing gaze,

raising the eye-brows and changing the tone of voice, constitute a natural protocol to convey pertinence in human communication [24]. In natural human communication, ostensive signals index the reliability and trustability of the communicator as well as the pertinence of the communicated message [24]. Hence, a natural prediction of the theory is that a person ignored while telling a story would generate negative beliefs about the recipient (the person breaking ostension) and also about the story being told. Our work confirms this hypothesis both from the point of view of the emitter and the receiver of the communicated message, quantifying the cost of inattention.

Finally, our analysis showed that the temporal distribution of attention has a negligible effect on the communicators' beliefs. Different theories predicted opposing outcomes from this temporal manipulation of attention.

One principle is that communication begins with a handshaking protocol, a statement by which the agents agree upon their intention to communicate [93]. From this principle it derives that inattention in the first moment of the dialog should lead to a worse perception of the receiver (the lack of politeness of sustaining a hand-shaking protocol). Instead, inattention during the last moments of the dialogs does not break the protocol of communication and may be instead attributed by the emitter to the fact that the story was uninteresting. A dancing metaphor may help anchoring this idea. If a person invites another to dance and the invitee (the recipient) rejects the invitation, the emitter (the person making the invitation) may make the inference that the invitee was not polite. Instead, if the receiver accepts the invitation but breaks the dance by the end of it, the natural inference is that the dance was not good enough to sustain attention and hence blame is on the message (the dance, or the story) and not on the conversational partner.

Alternatively, it may be reasoned that – as observed in other domains of cognition – retrospective beliefs are dominated by the perception at the last episodes [62]. For instance, Kahneman (1993) found that subjects retrospectively prefer a treatment in which a fixed amount of pains is followed by a lower dose than when it is followed by no pain. This implies that participants do not accumulate the total amount of discomfort but instead generate beliefs based on fragments of the experience, largely dominated by the last episode. In our experiment, if the data were dominated by this principle we would expect that participants' beliefs would be affected by the temporal distribution of attention (experience would be ranked worse when the listener neglects the speaker during the end of the story) and less so by the total amount of attention. Instead, our data show that the distribution of attended time plays a negligible role in the variance of subjective perception of the story and of the conversational partner.

In summary, a quantitative analysis following a face to face brief (4 minutes) communication shows that the total amount of attention is the major factor driving subjective beliefs about the message (the story being told) and of the conversational partner. The effect is observed on both the emitter and the receiver and is mostly independent on how attention is distributed in time. Interruptions during day to day communication between peers and also with children are extremely frequent. Our data should provide a note of caution, by signaling the consequences of these windows of neglect on the teller and receiver of a story.

3.5 This experiment in the media

- `https://www.tiempoar.com.ar/articulo/view/57222/las-apps-no-hacen-la-felicidad-`

el-80-se-borran-en-el-dia

- https://www.clarin.com/sociedad/ciencia_argentina-celulares-comunicacion-conversacion_cara_a_cara_0_rJiXVJKDQg.html
- http://www.perfil.com/ciencia/muestran-que-el-smartphone-afecta-el-dialogo-cara-a-cara-0219-0080.phtml

3.6 Supporting Information

Buenos Aires						
	Story			Conversational Partner		
Factor	F	df	p	F	df	p
Role	11.28	2	$\mathbf{< 10^{-8}}$	6.25	1	**0.002**
Attended Time	14.64	1	**0.0001**	10.4	1	**0.0013**
Interaction	0.72	2	0.49	0.45	1	0.64
Rosario						
	Story			Conversational Partner		
Factor	F	df	p	F	df	p
Role	4.82	1	**0.0084**	8.47	1	**0.0002**
Attended Time	12.08	1	**0.0005**	27.69	1	$\mathbf{< 10^{-8}}$
Interaction	1.47	1	0.23	3.44	1	0.03

Tab. 3.3: **Story and the partner scores.** ANOVAs for the story and for the partner scores at each experiment location (Buenos Aires and Rosario), with role (Speaker or Listener) and time of attention (0, 2 or 4 minutes as main factors).

Speaker						
	Story			Conversational Partner		
Factor	F	df	p	F	df	p
Order	1.54	1	0.21	0.2	1	0.66
Pattern	3.14	1	0.07	0.76	1	0.38
Interaction	0.28	1	0.6	2.87	1	0.09
Listener						
	Story			Conversational Partner		
Factor	F	df	p	F	df	p
Order	2.92	1	0.09	5.89	1	**0.01**
Pattern	4.25	1	**0.04**	2.92	1	0.09
Interaction	0.14	1	0.71	0.11	1	0.74

Tab. 3.4: **Subjective perception of the story and partner.** ANOVAs for the subjective perception of story and partner for each role, with order (whether listener finished paying attention or not) and temporal pattern of attentions (2 interleaved blocks of 2 minutes or 4 interleaved blocks of 1 minute).

Resumen en español

La influencia de los celulares en las charlas en persona

El mundo ha cambiado desde que los teléfonos celulares se hicieron populares y accesibles. De acuerdo a un informe de la Unión Internacional de Telecomunicaciones (World Telecommunication) [57], en los últimos 20 años, la Argentina pasó de tener menos de tres teléfonos celulares cada 100 habitantes, a más de uno por persona en el 2015. Esta misma tendencia se extiende a nivel mundial.

En este capítulo se explora cómo los teléfonos celulares pueden afectar la percepción de los participantes de una charla cara a cara. Para ello, se utilizó un diseño experimental en el que dos personas interactuaban unos pocos minutos en los que uno le contaba al otro una historia personal. Esto se realizó indicado a quien escuchaba esa historia que interrumpa en mayor o menor medida su atención hacia el otro para mirar su celular (escuchar y mirar todo el tiempo al otro, ignorarlo todo el tiempo mirando el celular, prestar atención solo la mitad del tiempo, etc). Luego de ello, ambos debían reportar lo que habían percibido del otro y sus sentimientos respecto de la interacción. Este experimento fue llevado a cabo en dos ediciones de las charlas TEDx (TEDxRíodelaPlata y TEDxRosario).

Una vez recabada la información de los reportes, se realizó un análisis cuantitativo que mostró que el tiempo de atención es el factor principal que impulsa las creencias subjetivas sobre el mensaje (la historia que se cuenta) y el compañero. El efecto se observa tanto en el emisor como en el receptor y es independiente de cómo se distribuye la atención en el tiempo de interacción. Las interrupciones durante la comunicación diaria entre compañeros y también con los niños son extremadamente frecuentes y tal como muestra este estudio, afectan el vínculo entre las personas. Por ello, podemos afirmar la importancia de considerar las consecuencias de las interrupciones durante un diálogo, a fin de evitar que el mensaje y las percepciones de los participantes se vean afectadas.

En síntesis, este capítulo contribuye a la evaluación del impacto de la inclusión de las TIC en contextos informales. Queda pendiente avanzar en estudios de este tipo en entornos formales, como las escuelas.

4. CROWDSOURCING EXPERIMENT: TRANSMISSION OF AN IMAGE THROUGH TEXT

No mires por favor, y no prendas la luz, la imagen te desfiguró.
–Patricio Rey y sus Redonditos de Ricota

Learning through Internet has become popular in the last years. MOOCs are an example of a new way of learning. This new educational setup has also different challenges to address, such as preventing drop-off, effectiveness of the learning and engagement. Some characteristics of these new environments are: less opportunities for direct interactions with the teachers, automated survey assessments and, students assign time and focus on their own. In the present chapter we address the following questions for digital environments: how good (or bad) can explanations in textual format be produced by the user of a digital environment and, how good (or bad) can these explanations be interpreted.

Using a specially developed cell-phone app, we collected data of people communicating simple geometrical figures in a written way, later reconstructing a figures from others explanations and, assessing the whole process.

The main contribution of this work is the measurement of the limits of communicating geometric concepts in a written way through an app, implementing a large scale digitalized experiment using crowdsourcing and gamification. Some of the observed strategies used to communicate worked out well while others did not work at all. This situation is an advice to be aware when including new technologies in educational processes: special attention must be taken into account when choosing communication strategies.

This work was possible through the participation of an interdisciplinary team with programmers, physicists, computer scientists, designers and organizers of TEDx talks.

This chapter was partially published in the *Journal Plos One* on November 10, 2015 [81].

4.1 Introduction

> *"Leaving hopes and utopias apart, probably the most lucid ever written about language are the following words by Chesterton: He knows that there are in the soul tints more bewildering, more numberless, and more nameless than the colours of an autumn forest... Yet he seriously believes that these things can every one of them, in all their tones and semitones, in all their blends and unions, be accurately represented by an arbitrary system of grunts and squeals. He believes that an ordinary civilized stockbroker can really produce out of this own inside noises which denote all the mysteries of memory and all the agonies of desire."*
>
> Jorge Luis Borges (The analytical language of John Wilkins)

Chinese whispers, also known as broken telephone, is a game in which a message is passed through a line of people. The last player announces the message which is often very different from the one that was uttered by the first. The game illustrates how different sources of noise accumulate, contaminating verbal communication.

However, even in the absence of noise, language poses a structural limitation in human communication. As expressed by JL Borges in the quote above, inner concepts are much richer than the words that describe them. Moreover, the same words often construct very distinct mental representations in different individuals. This is, of course, an essential problem of pedagogy: how to faithfully communicate concepts using words.

Here we designed a game inspired in Chinese Whispers, to investigate which attributes of a description affect its capacity to faithfully convey an image. This is a two player game: an emitter and a receiver. The emitter was shown a simple geometric figure and was asked to describe it in words. He was informed that this description would be passed to the receiver who had to replicate the drawing form this description.

Participants ranked different stages of the communication process: 1) the quality of the description of the geometric figure produced by the emitter ("*Description score*"), 2) the quality of the drawing produced by the receiver given the description that he obtained, ("*Drawing score*"), 3) the visual similarity of the original figure and the resulting drawing ("*Communication score*"). In addition, participants ranked five different aspects of each description: 1) Its coherence, 2) Whether it was procedural or not, 3) Its creativity 4) Whether it had mathematical content (mainly geometry and number concepts) and 5) Whether it used of metaphors.

Our aim is to investigate which of these aspects of the description predict its capacity to faithfully convey a message. Danielle McNamara and collaborators have shown that text comprehension is mostly determined by its cohesion [91]. Hence our first driving hypothesis is that the degree of coherence –which is a main factor driving cohesion– should be a very strong indicator of the communication score of a description.

Instead, our current knowledge does not make clear predictions on whether using more mathematical terms or procedural descriptions should be more effective to convey a geometric image. Since children from a young age have well developed intuitions of number, shape and other geometric features [36, 123], one can argue that using mathematical concepts may capitalize on an intuitive description system. Similarly, some pedagogical traditions have placed a strong emphasis in the description of procedures (for instance inclined to use logo like programming in school) [55]. Here we can settle empirically whether descriptions are more effective when weighted towards procedural terms (i.e. first go to the left corner, then draw a straight line...) or containing more mathematical concepts (it is a symmetric triangle, with a square on top) .

Finally, there is a growing belief in the general public that creativity [75], and to a lesser degree the use of metaphors, is a fundamental aspect of effective human communication. Here we examine this in a quantitative manner to investigate whether –for this narrow class of human communication– the perceived creativity of a description correlates with the success to effectively communicate a message.

4.2 Methods and procedures

This experiment was based in an Android app. The subjects downloaded the app.

They signed a consent to allow us to use the collected data for research. The procedures of the experiments described here were approved by the ethics committee of CEMIC (Centro de Educación Médica e Investigaciones Clínicas Norberto Quirno).

The experiment has 3 stages: 1) writing, 2) drawing and 3) rating. Each stage has a written explanation before the subject begins.

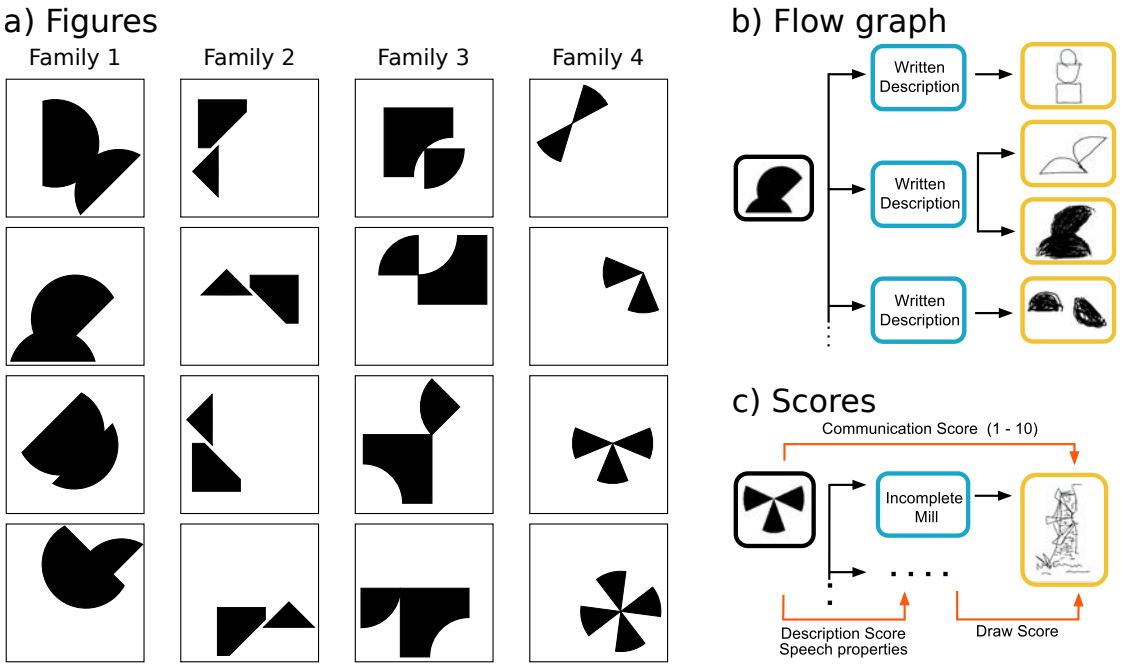


Fig. 4.1: **Flow graph of the game**. The panel a) shows a different possible flowgraphs. Starting with a real image example, from this image, many different subjects created different written descriptions, with one of these as instruction, other distinct subjects drew different pictures. Panel b) shows the three differents scores: 1) Communication Score: visual similarity among an image and a drawn picture. 2) Description Score: the quality of the written text as instruction to replay the image 3) Draw Score: the *power of the explanation* for the written description concerning the drawn picture. Beside these scores, the subjects rated written text *Speech properties*.

In the writing stage, the subject has the role of *emitter*, receiving a randomly assigned image (there are 4 different categories with 4 different images, see Appendix A.4) and writing a text to let another person replicate this picture later.

After writing, the second stage is the *description*. Here the subject receives a description produced by another participant. This text was randomly selected with the constraint that was created describing an image of a different category. The subject reads the instructions and then draws in a white canvas. The only two possible actions are drawing with black or erasing (drawing with white).

Finally in the last stage, the subjects rate the different parts of the process:

- Description score: they receive a picture and a description and rate the quality of the text.
- Drawing score: they receive a description and a drawing and rate the quality of the production.
- Communication score: they receive a picture and a drawing and rate the visual similarity of both elements.

Participants also rate description in different dimensions:

- Coherence. A description is very coherent when it makes sense and is well formed, allowing fluent reading.
- Procedural. A description is procedural if it explicitly presents the steps to get the result. In contrast, a description is not procedural when it does not present steps but a description of how it looks.
- Creativity. A description is creative when it relies on an original idea to describe the image.
- Mathematical. A description is very mathematical or geometrical when it includes terms of these areas.
- Metaphors. A description is very metaphorical when uses a different semantic context as resources for explanation.

In contrast to the description and drawing stages where subjects could only do these stages once, the rate stage could be completed as many times as the subject want. It seems like participants were engaged with this last stage: from the 800 subjects we got near 12.000 rates of the process and 3.500 rates of the dimensions of the descriptions.

To get a large amount of participants we capitalized on the TEDxperiment group in 2013. In the same year this group ran a live experiment in the TEDx Talks in human verbal communication [83].

4.2.1 The app

The app was developed in collaboration with the Grupo de Computación Movil (Departamento de Computación - FCEN-UBA). The original idea was tested many times using a paper version (in Appendix A.4 the material for the paper based is shown). This version was used as a game, and it glimpses the opportunity of characterize the communication process.

To scale-up an app was needed, and a non-bias qualification system was also required to have an accurate result. With the crowdsourcing idea, the app was designed with four stages to define the flow:

1. Login: the user has to accept to participate in the experiment and fill age, gender and occupation.
2. Write: user writes down the instruction
3. Draw: user draws according to the instruction of other participants
4. Assess: user assess different aspects of the process as explained before. This is the only one stage that can be done as many times as the user wants.

An schema of the flow can be observed in Fig. 4.2.

The aesthetics of the app was done by an external graphic design studio. Mockups are included in the Appendix A.4.

Unfortunately only an Android app was developed. There is no iOS, BlackBerry or Windows mobile version.

The data was registered by a web server hosted in the Departamento de Computación. The server was developed using Linux, PHP and MySQL (a schema of the database is included in the Appendix A.4).

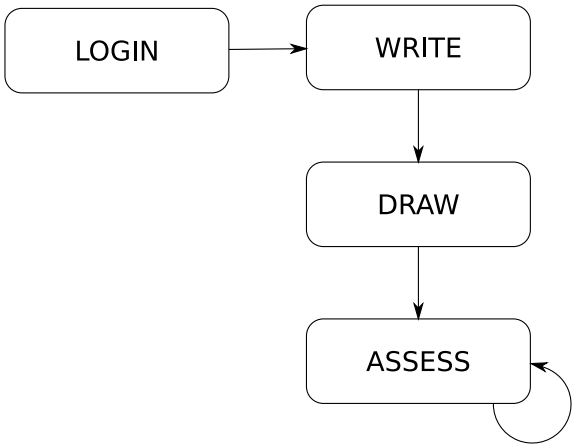


Fig. 4.2: app flow.

4.3 Results

Using a cell-phone app provides information about access to the application and usage.

In Fig. 4.3 a summary of the access to the Play Store page is shown. This information can be use to understand who are the members of the population that are using the app (Platform, Browser, Country) and how do they get to this page (mostly from a mailing list 57.3%, then Facebook 9.5% and Twitter 6.8%).

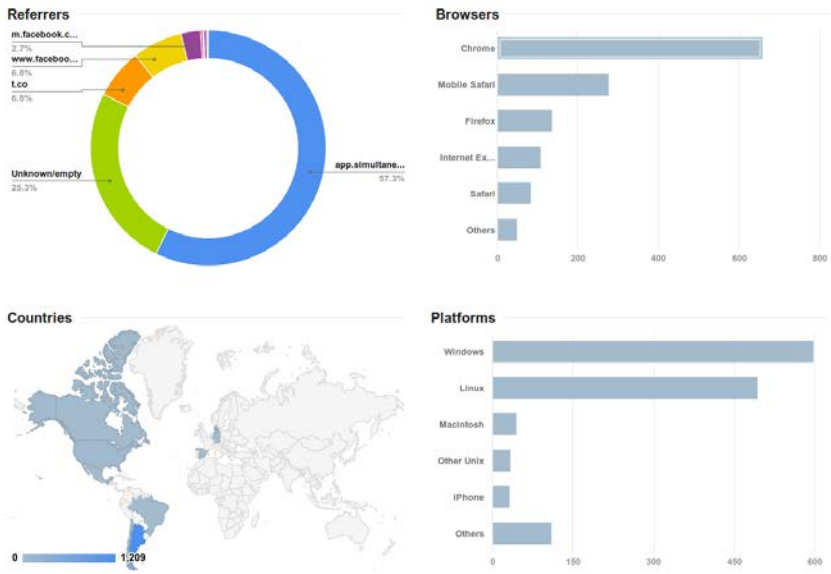


Fig. 4.3: Analytics from the Play Store access.

The second interesting result is the life cycle of the app which can be measure as the amount of users (or new users) by day. Three important moments are easily identify:

1. The midnight of 16th of September 2013: where the first massive email invited users to participate in the experiment

2. The TEDxRíodelaPlata event (September 28th), previous day are boosted by the

email remainder

3. The TEDxRosario event (October 9th), boosted as previous one but with a smaller population

This data is summarized in Fig. 4.4.

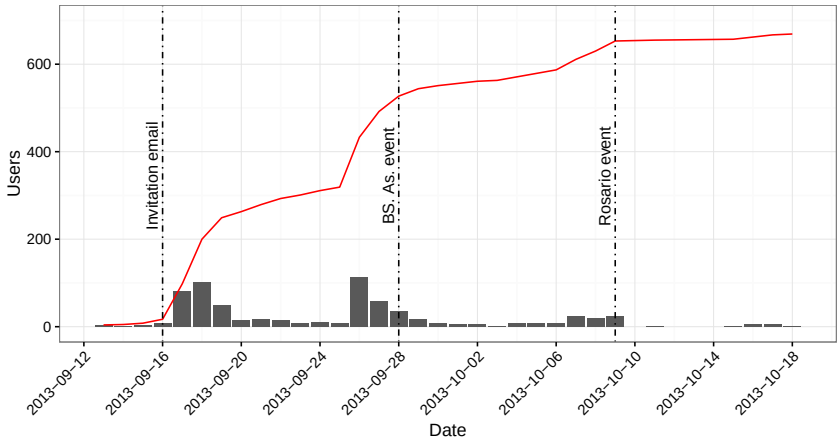


Fig. 4.4: New users register into the app day by day. The red line represent the total of users registered until each day. Three events are marked: first email invitation to join the app, and the two events who promoted the app.

In the Appendix A.4, Fig. A.12, A.13, A.14 and, A.15.

4.3.1 Correlational structure of the data

Naturally, the different measures are correlated as the communication score combines the precision of the description (*Description score*) and the capacity of the receiver to translate faithfully the description into a drawing (*Drawing score*). Here we do not seek to dig in detail how the description score combines from these two individual steps. Instead, we merely confirm the natural expectation that the quality of the description is tightly correlated to the quality of the entire communication process when the drawings are good. To this aim, we measure the correlation between *Communication score* and *Description score* for different values of the *Drawing score*. Results clearly show that the correlation is close to zero for the worse ranked drawings and grows to values close to 1 for the best ranked drawings (Fig. 4.5). For all subsequent analyses we filter out data with (*Drawing score* ≤ 4).

Next we investigate the correlation matrix of the five aspects of the description. This result showed that the correlation matrix is organized in two blocks: one corresponding to coherence, procedural and mathematical, and the other to metaphorical and creativity (Fig. 4.6 A). However, within these blocks there is sufficient dispersion in the data to assure a stable simultaneous regression to these five attributes (Fig. 4.6 B).

4.3.2 Attributes affecting description and communication scores

Our main aim was to investigate which attributes of a description affect its capacity to faithfully convey an image. To this aim we performed two independent linear regressions

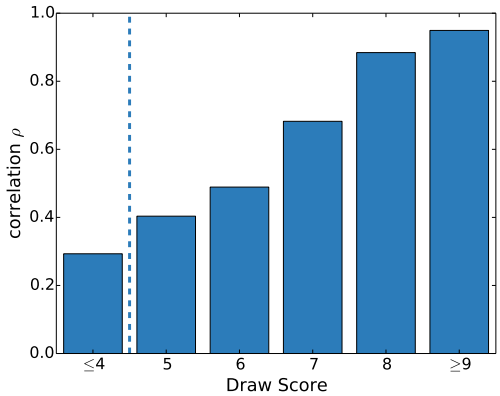
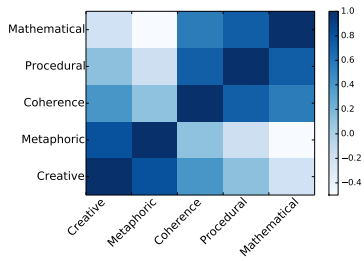
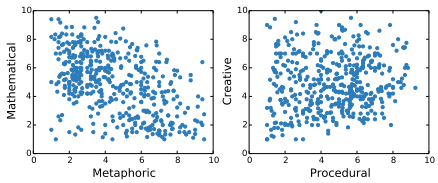


Fig. 4.5: **Correlation as function of the draw quality** Correlation between *Communication score* and *Description score* for different values of the *Drawing score*



(a) Correlation between all speeXch properties



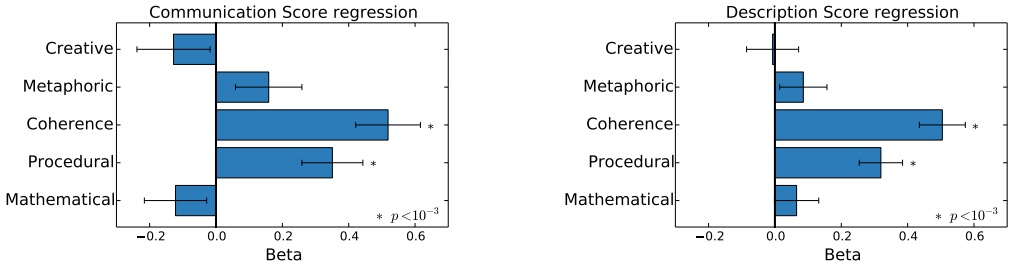
(b) Examples of correlations

Fig. 4.6: **Relation between speech properties** Panel a) shows the ρ of Pearsons correlations between every speech properties. Panel b) shows two example of correlation between some speech properties

(one for *Communication score* and one for *Description score*). Both analyses showed the same result: a highly significant effect of procedure and coherence and no effect of the other three attributes (Table 4.1, Fig. 4.7).

	Description score				**Communication score**			
	β	se	t	*pval*	β	se	t	*pval*
mathematical	0.0653	0.0668	0.9781	0.3287	-0.1220	0.0939	-1.2987	0.1949
procedural	0.3194	0.0653	4.8911	1.50×10^{-6}	0.3508	0.0919	3.8185	0.0002
coherence	0.5046	0.0693	7.2835	1.97×10^{-12}	0.5186	0.0975	5.3217	1.7855×10^{-7}
metaphoric	0.0856	0.0712	1.2026	0.2299	0.1589	0.1001	1.5869	0.1134
creative	-0.007	0.0784	-0.0891	0.929	-0.1280	0.1103	-1.1608	0.2465

Tab. 4.1: Linear regressions of speech properties



(a) Communication Score

(b) Description Score

Fig. 4.7: **Speech properties linear regression** Panel a) shows the β and the Standard error about the Communication Score. Panel b) shows the same but about Description Score

4.4 Discussion

Here we capitalized on vast data obtained from an android app to quantify the effect of different aspects of a description on communication precision. We show that descriptions more effectively communicate an image when they are coherent and whey they are procedural. Instead, the creativity, the use of metaphors and the use of mathematical concepts does not affect its fidelity.

The motivation for this experiment was to construct a minimal and controlled setup emulating variables of pedagogy amenable to quantitative analysis. This follows the strategy used by Shafto and Goodman [115, 116], to inquire whether learners adapt to make optimal inferences from the specific examples provided by teachers. Communication and pedagogy are intrinsically related with only subtle differences [126]. Compared to a typical pedagogical situation this experiment is simplified in four critical parameters: 1) the description is exclusively based on words (there are no gestures, prosody, ostensive indicators, etc. which are known to be fundamental for pedagogy [24, 44]. 2) There is no dialog or conversation, and hence no online feedback or opportunity to correct errors during the pedagogical process. Instead our setup builds on a one way message, resembling or modeling of an expository lecture. 3) The concept to convey is simplified to a simple geometric figure instead of a more elaborated concepts. Thus we propose this as a model of an extremely eroded and simplified version of a minimal pedagogical dialog. The great advantage of this design is that we can replicate many instances of the same pedagogic process and hence inquire which aspects of the message are relevant for the success of the pedagogical experience.

The observation that, when filtered for good drawings, the communication and description scores are highly correlated indicate that participants have a good understanding of what constitutes an effective description. This finding is not trivial since students often rely in heuristics to evaluate the effectiveness of a lecture, leads to misconceptions of the quality of a class. For instance, in the well known "Dr. Fox effect" students rely on teacher expressiveness (often more than on its content) to provide teacher ratings [94]. In our study we reasoned that judges may believe that creative or metaphoric explanations would be more effective than they are in reality. This hypothesis was rejected by showing that participants understand that coherence and procedural content are the most effective aspects of a description and that, conversely, highly creative or highly metaphoric descriptions are not very likely to be effective. It is important to mention that in our game

communication was considered effective when the resulting drawing replicated accurately the original image. In other pedagogical contexts where the goal may not be to optimize the precision of the communicated message, the results are expected to change.

Resumen en español

Experimento de *crowdsourcing*: transmitiendo una imagen a través de texto

El aprendizaje a través de Internet se ha vuelto popular en los últimos años. Los MOOC (Massively Open On-line Courses) son un ejemplo de esta nueva metodología de enseñanza-aprendizaje. Esta nueva modalidad educativa plantea nuevos desafíos: cómo prevenir la deserción, de qué forma lograr la eficiencia del aprendizaje, cuáles son las mejores maneras de incentivar la participación, entre otras. Algunas de las características de estos nuevos entornos son las menores oportunidades de interacción directa con los profesores, la necesidad de contar con correcciones automatizadas de las tareas y la participación de estudiantes que realizan los cursos con su propia dinámica.

Partiendo de estas consideraciones, en el presente capítulo analizamos el modo en que se produce una explicación escrita para trasmitir a otro una idea. Esta es una cuestión clave en los entornos digitales que, como ya mencionamos, no cuentan con la posibilidad de interacción cara a cara. Las siguientes preguntas orientan el análisis que presentamos: ¿qué tan buenas (o malas) pueden ser las explicaciones en formato de textos producidas por usuarios de entornos digitales?, y ¿de qué modo estas explicaciones son interpretadas por otros usuarios y cuán efectivas son éstas para la trasmisión de un concepto geométrico?

Para responder estos interrogantes, desarrollamos una aplicación para teléfonos móviles en la que los participantes tenían que escribir un texto destinado a que otro reconstruya una figura geométrica simple. Luego de esta primera reconstrucción, la persona debía repetir el procedimiento a partir de una nueva explicación propuesta por un usuario diferente. Finalmente, se le pedía que evalúe las figuras realizadas por otros participantes. Esta aplicación, especialmente desarrollada, nos permitió recopilar los datos de las personas que la utilizaron para luego analizarlos.

La principal contribución de este estudio es la medición de los límites de la comunicación de conceptos geométricos a través de la escritura, realizada mediante el uso de TICs. Por otra parte, la implementación de una experimento digitalizado a gran escala utilizando *crowdsourcing* y *gamification*. Del análisis de los datos surge que algunas de las estrategias para la descripción utilizadas por los usuarios fueron efectivas en la transmisión del mensaje, mientras que otras no funcionaron en absoluto. Los estilos elegidos para componer las descripciones afectan la efectividad de la transmisión del mensaje, siendo la descripción algorítmica el que dio mejores resultados, en cambio, las metáforas y la creatividad en las descripciones no produjeron el efecto deseado, es decir la transmisión fiel de la figura mediante el mensaje

A modo de conclusión, afirmamos la necesidad de presentar especial atención a la selección de las estrategias para la comunicación de las tareas y consignas, al momento de incorporar nuevas tecnologías en los procesos educativos.

5. MATE MAROTE

Fijate de que lado de la mecha te encontrás.
–Patricio Rey y sus Redonditos de Ricota

The Mate Marote project was created in 2008, designed as a cognitive training intervention tool based on video games. Along the years, it had many editions. There were pilots in the lab; later, there was an experience in a primary school where transfer from executive functions training to language and maths exams results was observed [43]; and finally an unsupervised version was deploy in La Rioja –using the One Laptop per Child (OLPC) platform– where the same behavior of game playing from previous schools was observed in children without tutor observation [82].

All these projects were developed by researchers attending to schools during the intervention and, thus, had problems to scale-up. In order to get a robust software which could scale-up to bigger interventions, a leader software factory was hired to address technical high performance issues, funded by a grant from Argentina's government. This Chapter is the report of this University-Industry interaction. The actors involved were programmers, scrum masters and, public agencies agents (Ministerio de Ciencia, Tecnología e Innovación Productiva and Fundación Sadosky).

This chapter was partially published in the *8 Jornadas de Vinculacion Universidad-Industria* in the *43 JAIIO* [121].

5.1 Introduction

Argentina is living a unique opportunity where a great amount of primary and secondary school students have their own laptop, as mentioned in the Chapter 1.

Concurrently, the advances in educational neuroscience have shown that research in cognitive neuroscience and behavioral methods – which investigate the development of mental representations – may help to improve different teaching approaches [103, 110]. Moreover, many initiatives introduce games as a strategy to stimulate learning in children by increasing motivation [76, 109, 47, 12, 33, 43].

Mate Marote is a flexible framework consisting of educational games and activities [82]. This tool was specially designed to implement interventions, allowing to measure the improvement of users while using the activities. Based on a pilot implementation in La Rioja province, Mate Marote was able to manage a few hundreds of registered users using the platform.

One of the motivations of developing Mate Marote was the absence of a free, unrestricted and open source tool for cognitive assessment and training. The scientific community used to pay licenses of private products to evaluate subjects, such as: E-Prime (Psychology Software Tools, Inc) and Wechsler Intelligence Scale for Children (WISC-V, Pearson Education, Inc. - PsychCorp).

In this digital era, where a significant fraction of students and teachers have their own laptops, Argentina is not alone. Worldwide initiatives are being implemented, delivering low-cost laptops to every primary and secondary school student, for example adopting the

One Laptop Per Child program (OLPC). Only in Uruguay, the CEIBAL program, provides the whole population of elementary and secondary schools – some 400,000 children and their teachers – with the same digital platform.

This new social phenomenon has the potential of making a profound change in the education of millions of children around the world. With the goal of re-implementing Mate Marote for massive scale interventions, University of Buenos Aires joined industry experts in architecture and scalable systems of EPIDATA to create a new partnership that may upgrade Mate Marote to the Big Data era. In this article, we describe the collaboration program started.

5.2 Previous Work

Pilot interventions of Mate Marote were successful in terms of educational issues, but did not scale well in technological aspects. Based on these results, a proposal of upscaling the system was presented to the Ministerio de Ciencia, Tecnología e Innovación Productiva (MinCyT). The main goal was to get the investment to re-development of the Mate Marote platform in order to be able to support millions of users.

The MinCyT accepted the proposal and forwarded it to the Fundación Dr. Manuel Sadosky (FS), which is a mixed (public / private) institution whose goal is to promote stronger and closer interaction between Industry and the scientific-technological system. Mate Marote research team and FS discussed the requirements needed for the upscaling of framework, and thus FS launched an open call for companies to develop the upgrade of the platform.

The company Epidata was the winner of the call. Epidata is an Argentinian company, which provides software development and maintenance services focused on: software architecture, agile methodologies, and open source technologies. Epidata has also a proven track record working on research and development projects along with several research departments in latin-america.

The software project management was based on the iterative and incremental agile software development framework Scrum, coordinated by a project leader offered by Epidata. The Scrum schedule was designed in ten-working-days sprints, with the review and planning meetings on the same day (after last day of current sprint and before the next sprint starts). The review meeting consisted of reviewing planned issues developed, where researchers had to analyze and eventually approve each individual issue. On the other hand, the planning meeting consisted of defining the issues to be developed during the next sprint, consensual by researches and developing team.

The development time was divided in two big sub-projects with equal dedication: the backend, which consists in a web server with resources and services; and the frontend, which consists on games and graphical resources.

5.3 Requirements

Due to previous experience some key aspects had to be faced with special attention. A list of these important issues were:

- API multimedia resources

  - Image: the games can use their own images or choose from the system some of this resources. i.e. characters, objects, landscapes, buttons, images with text in a certain language
  - Sound: as in the image, each game should be able to include its own sounds (i.e. a character voice, an special sound) or to use system sounds (i.e. background sounds, laughs, cheers, animal sounds).
  - Text: same as before.
  - Other resources: other items should be standardized and available to use in the games, such as: a level pointer (to show how far the child has advanced), a timer and a chronometer with a attractive representation, feedback for different trial results (a colored shadow, a character with a gesture), lives (maybe hearts, or mini-characters), awards.

- API state:
  - The games must be able to save and load easily the state to be able to recover if the session stops, or in checkpoint situations where the if the child makes a mistake the game restarts from that point.
  - The game must be able to save and load user personalization, such as avatar, selected character, favorite color, lives collected.

- API logging
  - The game must be able to store in a easy manner events of the game, information of the current state, and behavior. This information must include a timestamp and the play logged must be reconstructed from the collected logs.

- API gamification
  - The game must be able to use global rules for prizes, such as "the amount of lives needed at the end of the game to receive a medal is 3", or checkpoints, such as "every three trials there is a checkpoint".
  - The system must be able to control the game to allow its use under certain conditions. i.e. a game can only be played 20 minutes a day, a game can only be played in the afternoon, a game can not be used during weekends.
  - Tutorials and instructions, the system must provide flexible tools to implement animated tutorials, videos or text instructions under certain conditions, i.e. at the beginning, before certain trial, after a number of mistakes.
  - The games should include the characters of the system (actually a boy, a girl or a cat, but new characters will be included) and provide skins related to the character for aesthetics.

- Regionalization
  - The system must support multi-language configuration.
  - The system must support a multi-site configuration, different web address will go to the same system, but the content showed must be personalized for each site.

All these requirements allowed to build the initial backlog presented in the appendix (see section A.6).

5.4 Results of the first 6 months

This University-Industry interaction had as a first result documents where the games included in Mate Marote where formalized. This documents can be found in appendix A.5.

After 6 months of project execution, we completed 11 sprints where we had achieved, in this first stage, over 75% of backend development and the frontend for players. Five games were implemented on testing phase, with even more games and new features to develop in future stages. The development included more than 7,000 lines of code with over 95% of acceptance rules based on Sonar[1], with 86% of successful integrations based on Jenkins statistics[2].

The team consisted of three game-design, web-development and backend professionals who have so far spent about 2,500 hours of development. This number did not include the time spent in analysis, design, UI, testing and management that was done by different professionals in these areas, who helped to carry out the project and contributed their ideas and knowledge to reach the stated goal.

At the beginning researchers as product owners, developers and the Scrum Master did a planning meeting where priorities were defined and story points were assigned to each story that could be included in the next sprint. After that, according to the priorities, the candidate stories and the team work capacity (measured in story points per sprint), stories were selected to be done. Finally, with the revision and the approval of this selection by the product owners the sprint scope was defined.

During the sprint, stories were implemented following the specifications. The existed fluid contact between the developer team and product owners was important to clear up doubts and avoid rework. When a story, or part of a story, was completed the functionality was deployed in a testing environment to make an appropriate quality control check. If there was a bug, or something does not have the desired behavior the developers received feedback of the problem to solve it. If everything worked as expected the story was marked as done and ready to be checked at the end meeting.

At the end of the sprint there was a review meeting where the developer team showed project progress and stories implementation. The product owners approved, or disapproved, the implementation of the stories included in the sprint according to criteria of approval defined in each story. As a result of the product owner revision, the backlog was updated. Finally, a retrospective meeting was done where teams discussed the ups and downs of this sprint and they defined, if necessary, new work patterns to improve productivity, such as how many story points can be done in a sprint. After all this, a new planning meeting was done in order to start the next sprint.

Gained experience sprint to sprint, the use of this methodology had impact in the team production which raised the accuracy of estimating development times, improved the quality of the product and overall productivity.

[1] Sonar is an open source platform for continuous inspection of code quality to perform automatic reviews with static analysis of code to detect bugs, code smells and security vulnerabilities

[2] Jenkins is an open source automation server written in Java. It helps to automate the non-human part of software development process, with continuous integration and facilitating technical aspects of continuous delivery.

One of the remarkable facts of the scrum methodology is that at the end of each ten-working-days sprints a working software is delivered to the product owners which is incremented every sprint. This give the opportunity to show, use and test the project progress without unwanted detours.

We had finished programming almost the complete backend, including the admin interface that enables the creation of intervention programs (see a screenshot of admin interface in Fig. 5.1) and also the front end was nearly finished, (preliminary gameflow design where games are selected is shown in Fig 5.2a and a screenshot of the game Tower of London in Fig 5.2b).

5.5 Conclusions and Discussion

The massive availability of laptops owned by students opens up a new era in education, combined with recent advances in educational neuroscience. Argentina is leading this route with the one-to-one educational programs all over the country. Mate Marote has rised as a platform with educational games and activities, taking advantage of this new reality. However, pilot interventions were deployed in reduced scale scenarios.

Recent creation of Fundación Sadosky, has favored the collaboration between industry and academia in projects involving novel methods and ideas with big scale software platforms which require experience in architecture and high-end software development. The collaboration established showed exceptional conditions for the mutual benefit: the incorporation of academic research in educational neuroscience in the industry and the transference of well-established methodology and software development experience in the academic environment.

Moreover, the Scrum methodology adopted for the requirement definition was particularly successful as both sides dominated the technique. This particular (uncommon) situation helped the rapid development of the platform, with elevated scaling and the inclusion of several new games.

One of the biggest benefits for the industry side of this experience, i.e. Epidata, was the incorporation of state-of-the-art knowledge about educational games based on neuroscientific studies. Additionally, this kind of projects help reducing costs in I+D+i, while it harvests prestige because of linking with University.

From University's point of view, the development methodology introduced by industry and the short-term implementation times are very valuable, mostly due to its human resources with full-time dedication to the project. However, some disadvantages were observed, specially in the specific objectives of the platform and the quality-quantity tradeoff.

Government presence, through Fundacion Sadosky, facilitated the link between University and Industry, with the grant policy and acting as supervisor in the University-Industry link, towards the success of the project. This experience suggests that strategy adopted in this University-Industry interaction is advisable for the execution of future project of this kind.

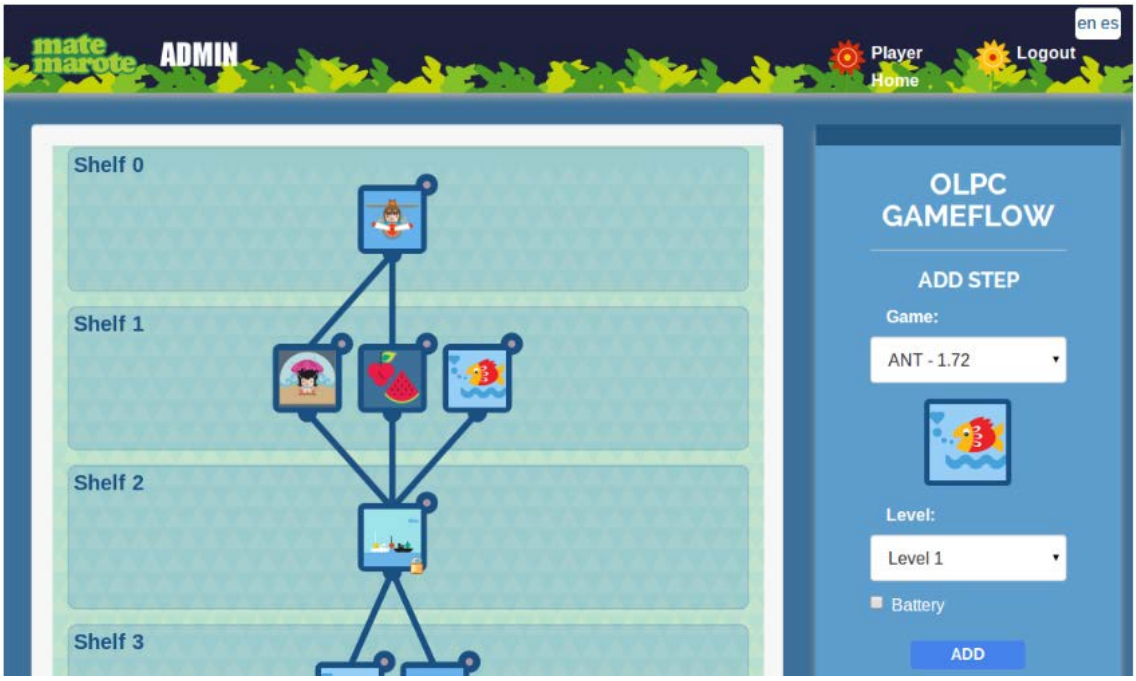


Fig. 5.1: Backend screenshot of Mate Marote new implementation. In this screenshot is shown the new tool designed by Epidata to elaborate interventions.

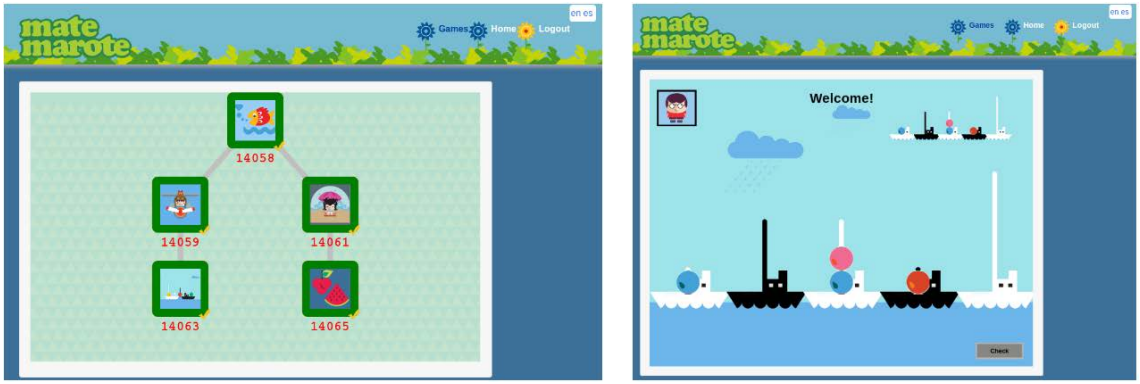


Fig. 5.2: Frontend screenshot of Mate Marote new implementation. In the left panel the Gameflow that the user will play is show, and in the right panel the game Tower of London is exhibited.

Resumen en español

Mate Marote

El proyecto Mate Marote fue creado en el año 2008 e ideado como una herramienta de intervención cognitiva basada en videojuegos. A lo largo de los años, tuvo varias ediciones. Primero, versiones piloto en el laboratorio; tiempo más tarde, se realizaron experiencias en escuelas primarias, donde los resultados mostraron que se transfiere el entrenamiento de las funciones ejecutivas a las notas de lengua y matemáticas [43]. Por último se realizó una prueba sin supervisión en La

Rioja, utilizando la plataforma One Laptop per Child (OLPC), donde se observó que los niños, jugando de manera no-supervisada en su tiempo libre, tenían el mismo comportamiento que el observado en las escuelas anteriores [82].

Todos estos proyectos fueron desarrollados por investigadores que debían asistir a las escuelas durante las intervenciones. Esto limitaba la posibilidad de implementación a mayor escala. Por ello, fue necesario cambiar la estrategia, mediante el desarrollo de un software robusto que pudiera ampliarse a intervenciones más grandes. Este trabajo se realizó mediante la contratación de una empresa líder en desarrollo de software que abordó los problemas técnicos de la alta performance. Esta etapa del proyecto fue financiada por un subsidio del gobierno nacional argentino.

El capítulo contiene la descripción de la interacción universidad-industria e incluye la documentación generada para el desarrollo de la aplicación, los juegos utilizados y las reglas de configuración de esta nueva plataforma.

6. CHANGE OF MIND

El lujo es vulgaridad... dijo... y me conquisto...
–Patricio Rey y sus Redonditos de Ricota

How children learn to solve a specific task? And how do they actually solve it? These questions can be labeled as an effort of reverse engineering the mind (i.e. understanding human learning and inference).

In this chapter we analyze children interactions with a video game to address this question. The video game is the digital version of a traditional board planning task.

The study explores how children learn under two conditions: learning from an adult and learning from a peer, in children from 6- to 10-years-old. We focus in how children identify their own mistakes and fix them (or not).

The video game stands on shoulders of Mate Marote platform detailed in Chapter 5. It was possible due to the joint effort of programmers, biologists, computer scientists, physicists, teachers and principals of a primary school.

The content of this chapter is currently in preparation to be submitted to an international journal.

6.1 Introduction

A child is playing with an assembling toy following instructions to build a car. When she finishes putting together all the pieces, she discovers a problem: one of the wheels was connected over the hood and, consequently, the car was not working. Despite having a great time, the outcome was not the one she expected. What should she do next? The child had to make an important decision: decide whether the car was correctly done or not. Even when we follow predefined instructions, outcomes may not be the ones we waited for. Following instructions can be sometimes tricky. What happens when instructions to follow are self-designed?

Observing humans behavior while working on difficult problems (but not unsolvable for them) may reveal certain broad characteristics of problem-solving [119]. In reasoning problems, the solver must find the solution to situations by herself, integrating available information in the best way. If the information available is complex, the accuracy may be affected [23]. The lack of capacity to build complete and correct solutions in the first shoot may lead to changes of mind, specially when the initial chosen actions were erroneous [107].

Decision-making has been studied in relying on binary choices by the Two-Alternative Forced-Choice paradigm. In addition, for simplicity, in the vast majority of the experiments, the decision variable is a single scalar (for example the luminosity of each patch, the number of dots in a number comparison task or the pitch in an auditory discrimination task).

This paradigm has a great benefit for computational and theoretical understanding of decision making. It can be fully expressed in a set of equations which have analytic solution [42]. Also, the functional dependence of behavioral observables such as response times, error rates or confidence can be described in detail by low dimensional models (e.g.

[104, 13, 129, 42]). While models differ, they almost always rely on a stochastic search process, in which the accumulation of the evidence is integrated over time, and whose crossing of a boundary represents the event of *reaching a conclusion* or making a decision [10].

Fluid reasoning –capacity to think logically and solve problems in novel situations, independent of acquired knowledge– changes over the childhood, affecting these decision processes [37]. Speed can be related to fluid reasoning: the better the player the faster she plays; or to lack of inhibitory control: the player plays quickly without an exhaustive analysis of the current state–inhibitory control is of particular relevance to cognitive development and processing capacity [108, 136]–.

The present work aims to measure the ability of change of mind in children from 6- to 10-years-old in the most ecological environment possible. For this purpose the study was developed in an elementary school, at the computers lab of school.

This study takes advantage of the motivation of playing video games, in this case implementing a reasoning-integration task. The information presented may lead them to wrong plans which they may fix. We analyze how schoolers – from 6- to 10-year-old children– solve puzzles, focusing in the decisions they take.

The game simulates a situation where the child arranges chocolates in boxes in a candy shop. To be attractive each new box has different requirements of assortments (i.e. all the dark chocolate in the left, and all the rounded chocolates in the right). The solution of each puzzle is unique.

The population was composed by 6- to 10-year-old children split in two groups: the first one learn the game from a researcher and, the second group learn the game from another child of the first group.

The present study inquires about three working hypothesis:

H1. the ability of changing the plan is affected by the age of the child, younger children are less likely to change,

H2. the ability of changing the plan is affected by who taught the game rules, those who received instruction from a peer will be more likely to rearrange the plan and,

H3. the slower the participant plays, the more likely to fix the wrong situation.

6.2 Materials and Methods

The subjects of the present work were 143 students (77 females) from a high socioeconomic status primary school from first to fourth grade –6- to 10-years old (mean age: 101.295 ± 14.108 months from 75 to 129 months)–. This study splits the population in two groups: the younger one, which includes first and second graders (n=61, 32 females, mean age: 87.733 ± 7.834 months from 75 to 101); and the older one, which includes third and forth graders (n=82, 45 females, mean age: 111.219 ± 8.113 months, from 94 to 129 months).

The hole population was also split in two groups across grades: one which learned the rules of the game from an adult; and one which learned from a child from the first group (peer teaching). All the subjects played 3 sessions of 15 minutes of different trials of Chocolate Fix in different day of the same week. The first group played first, then they taught the second group, which played the session afterwards.

Studies were conducted in a school of Argentina during Sep-Oct 2014. All children's parents or legal guardians gave signed voluntary consent. The consent form, presented

to the caregivers was previously authorized by an Ethical Committee: Comité de Ética de la Dirección de Investigación del Centro de Educación Médica e Investigación Clínica (CEMIC) "Norberto Quirno", Unidad Asociada del CONICET (Protocolo Nro. 683).

Chocolate Fix is a board game of the company Think Fun[1] where the player works in a candy shop completing orders from customers who request different assortment of chocolates (for extra details on Chocolate Fix, see Chapter 2, Section 2.2).

This game has a story: the player works in a Candy shop where customers have strange wishes for their box of chocolates. Each tray has to be filled is a 3 by 3 grid, and it will contains always the same 9 chocolates arranged in some way. The chocolates are all the combinations between 3 flavors (chocolate, strawberry and vanilla) and a shape (circle, square and triangle).

Each trial has a unique solution which consists in filling the tray with an assortment of all the nine chocolates that accomplish every customer requirement. The requirement is presented as one or more fragments –hereinafter referred as cues– of the tray where each position could be: a piece, an attribute of a piece (flavor or shape) or an empty spot. A cue is consider accomplished if it can be place over the board in a position where the pieces match with the information of the cue. i.e. over a strawberry triangle chocolate a cue can either have the same piece, a triangle shape, a strawberry flavor sign, or an empty spot to consider a match. A cue cannot be rotated, but it can be placed in any place of the tray. E.g. a 2 by 2 cue has 4 possible positions in the board, a 3 by 2 cue only 2, and a 3 by 3 cue has a unique position.

To win the participant must fill the tray and press the "Check" button to check the solution.

This study uses a digital version of Chocolate Fix (see Fig. 6.1) running in the web-based framework for interventions Mate Marote [82].

The present work aims to test what players would do when a puzzle presented to them includes some distractors and tricks which interfere with a straight forward plan to solve the puzzle. With this in mind, trials were designed with the following features:

- The tray starts partially filled with 7 pieces (participants can only move two pieces).
- One of the sides of the tray has the 2 empty spots that are adjacent.
- There is only one cue. It only has two spots with information: one with a features of shape and one with a feature of color. The information of the cue fits visually in the empty spots of the tray, but it also fits overlapping it with the piece already placed in the same side. The cue fits in 2 positions in the empty board (overlap with already placed piece or match visual shape).
- 2 pieces are missing: the *cued* piece–which shares the shape and the color with the cue– and the *uncued* piece –which does not share neither shape nor color with the information of the cue–.
- To solve the board, placing both missing pieces, the cue must be accomplished overlapping with the piece already in the board, in the same side of the empty spots. This piece shares a feature with the cued one. The uncued piece must be place by discard, because there is no information present about it is final position.

[1] Used with the permission of Mark Engelberg and Bill Richie.

Fig. 6.1 shows an example of a puzzle: the cued piece –pink circle– and the uncued –dark square– must be placed in the tray. The cue has the two features of the cued piece: a pink spot and circle shape. On the board a pink square is placed in the bottom right place. To solve the puzzle the cued piece must be place in the mid-right spot, next to the pink square. Meanwhile the uncued one must be placed in the remaining place, the top-right of the board.

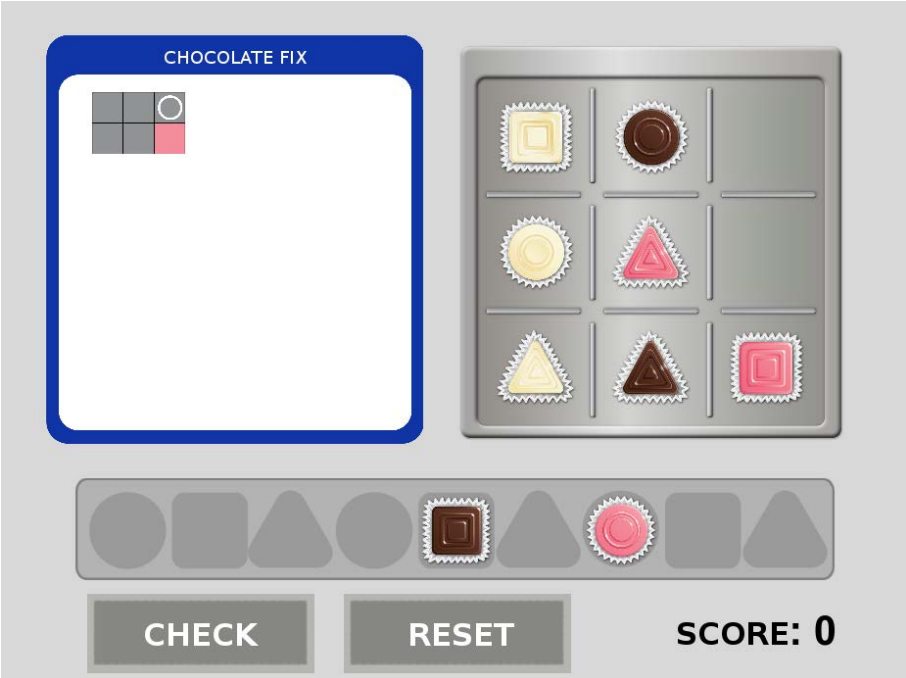


Fig. 6.1: **Screenshot of the video game.** The player wins the trial when the tray (top right side of the screen) –which in this example is partially filled– is completed with the remaining pieces (bottom of the screen) with the constrain of the cues showed (top left side of the screen). Placing the pink circle in the middle right spot and the brown square in the top right spot of the tray is the unique solution to this puzzle. The cue is accomplished in the bottom of the tray, where the circle cue is satisfied by the pink circle, and the pink cue with the pink square.

The present work includes data from 5 different trials with the characteristics mentioned before. Each child played all the 5 trials exactly one time, collecting 611 completed trials and 2947 events which consisted of: beginning of the trial, when a piece is placed or removed and when the player press check. There was no constrain of time per trial.

In the results section, the population was splitted according to the working hypothesis. The significance levels were reported taking to account corresponding Bonferroni corrections for multiple comparisons.

6.3 Results

When solving reasoning-integration tasks children may (or may not) change their plan when they reach a wrong partial solution. This ability of replaning the solution may be affected by different aspects. As mentioned in the introduction we pose 3 hypothesis: the age (H1), the one who taught (H2) and the speed (H3) influenced on the change of plan

ability.

In the current task the designed trials have only two missing pieces. Thus, a change of plan can be observed if after placing the first piece, instead of placing the remaining one to fill all the tray, the player changes the position (or just removes) the recently used piece.

The first movement could be done with any of the two pieces: the one that includes both features in the cue and the one that do not share any features with the cue. Since most of the trials first placed piece is the cued one (88.21% - 539 of 611), this work is focused on the possibility of changing the plan after this first move.

There are 2 possible positions where the cued piece can be placed: one is correct and the other is incorrect. Afterwards there are only two possible actions in order to fill the tray: keep going with the plan (and fill the remaining spot) or change the plan (removing or changing the recently placed piece). In Fig. 6.2 all the possible assortments of the tray are showed (the initial board in the top and complete trays in the bottom). Arrows represent change in the assortments product of placing or removing one piece. From the two possible boards that only include the cued piece (P_C), solid black arrows represent the change of plan, solid gray arrows represent keeping the plan under both conditions: first move correct and first move incorrect.

Statistical analysis confirmed that players whom played correctly the cued piece are less likely to change their plan (8.13%, 27 of 331) than the ones whom played it incorrectly (17.56%, 36 of 205) $p = 1.397E - 03$, two-sided Fisher's exact test - Fig. 6.3. Three trials were excluded because the participant checked the solution without placing the second piece.

Next, we explore the profiles of the participants to characterize which features influences their change of mind.

Does age determine the capacity of changing the plan? Data showed statistical difference among the group of older children: those who played correctly the first piece are less likely to change their plan than the ones whom played it incorrectly $p = 1.068E - 03$ two-sided Fisher's exact test. Meanwhile, younger children did not show any difference between the group who place the first piece correct against those who place it incorrectly, $p = 4.197E - 01$ two-sided Fisher's exact test (Fig. 6.4).

As explained in Materials and Methods, there were two group: a group who was taught by an adult, and a second group who children learned from a child of the first group. Does be taught by an adult promotes a better ability to change a plan? Data showed difference only in one of the two groups: those who learned from a peer $p = 2.044E - 03$ two-sided Fisher's exact test. Meanwhile, the group who learned from an adult $p = 1.492E - 01$ two-sided Fisher's exact test (Fig. 6.5).

Repeating the analysis but with the population split with two conditions: by grade groups and by the way each child learned, an statistical difference among conditions is only observed in the older children who learned from a peer: $p = 1.673E - 03$ two-sided Fisher's exact test (Table 6.1).

	Learn from adult	**Learn from peer**
1st and 2nd grade	7.258E-01	4.230E-01
3rd and 4th grade	1.407E-01	**1.673E-03**

Tab. 6.1: **Two-sided Fisher's exact test for subset of the population by grade group and by who taught them to play.**

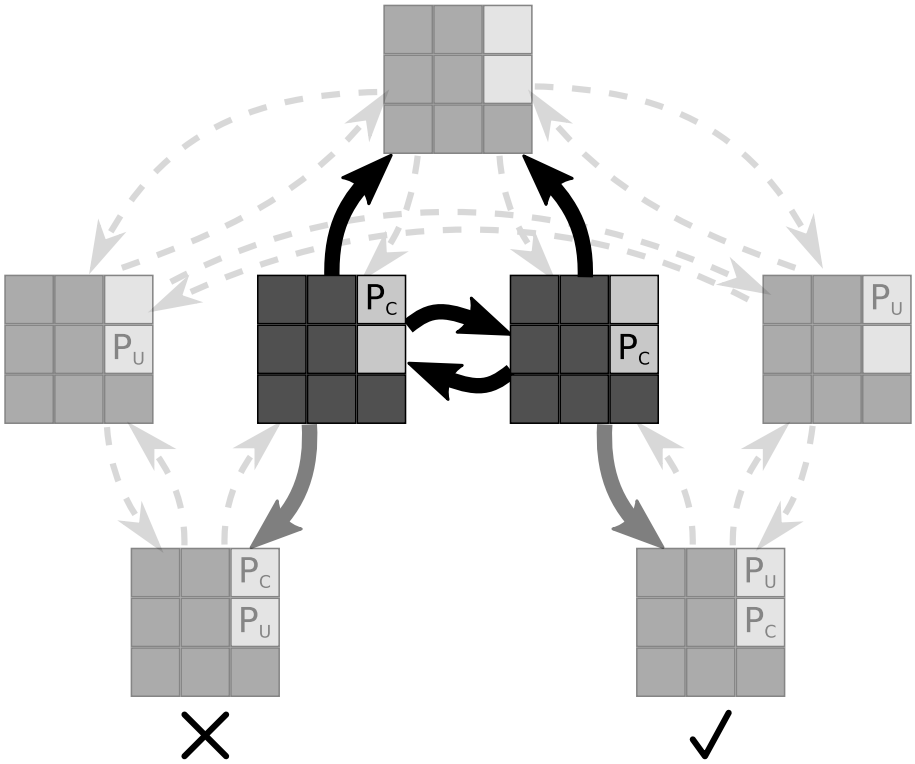


Fig. 6.2: **The space problem.** The top square represents the initial board with the empty spots in light gray, and filled spots in dark gray. The first row show boards where only one piece was added (P_c is the piece whit the attributes of the cue, P_u is the piece with no attributes in the cue). The third row has the two possible filled boards, the incorrect one (left side) and the correct one (right side). Each arrow represent a change in the board by adding, moving or removing a piece. The solid black arrows represent the situation of "change of plan", the solid gray the "continue with the plan" situation and the dashed gray arrows are movements not considered in this analysis. The present work focus on the movement after placing the cued piece (P_c), i.e. transition from the two boards of the middle of the figure.

Later on, this work presents results only of the group of children who learned from a peer.

Is there a characteristic of the players which influences this observed effect? Players move pieces at different speed. Can the speed of a player explain if there will be a change of plan or not? Speed can be related to fluid reasoning: the better the player the faster she plays; or to lack of inhibitory control: the player plays quickly without an exhaustive analysis of the current state.

First, we verified if response times were similar across grades. Data showed that older players play faster than younger ones. This was confirmed with a one-way ANOVA, which showed a significant effect of Grade on Response Time ($F(3, 139) = 6.674, p = 3.044E-04$). Since the speed distribution is not constant among participants over grades, the population was split in groups related to their mean move time within their grades: fast, regular and slow players (see Fig. 6.6). For each grade, the population was divided by their speed into fifths. The 2 quickest groups were tagged as *fast players* and the 2 slowest groups were tagged as *slow players*.

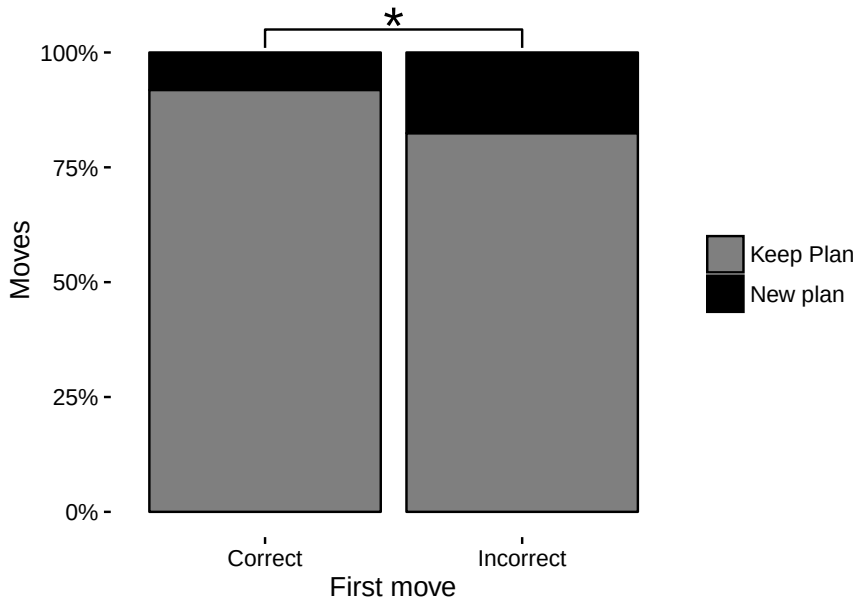


Fig. 6.3: **Change of plan.** We analyze what happened after the first move when P_c was placed, if the player kept filling the board (bold gray arrows, see Fig. 6.2) or if she changed her mind (black bold arrows), finding an statistical significant difference between the players whom placed it correctly against the players whom place it incorrectly ($p = 1.397E - 03$ two-sided Fisher's exact test). Players whom placed incorrectly the cued piece were more likely to change their original plan (i.e. fix the game).

Repeating the statistical analysis (players who played correctly the cued piece are less likely to change their plan) within these two groups a non statistically significant difference was observed in the *fast* group ($p = 4.976E - 01$, two-sided Fisher's exact test). Meanwhile, in the *slow* group the difference observed was significant ($p = 1.452E - 03$, Fisher's exact test - See Fig. 6.7). The *slow* players who place wrongly the cued piece in their first movement are more likely to change their plan –to the correct one– than the ones that place it correctly first. Opposite to the *fast* players who do not show this difference.

Finally, this works presented collected data from 1st to 4th graders students playing a reasoning-integration game. This behavioral data was explored in three dimensions of the population: the participant grade, the way she learn to play (a peer or an adult) and her speed. The data showed evidence to confirm the 3 hypothesis:

1. younger children were less likely to change their plan when they get to a wrong position,

2. those who received instruction from a peer were more likely to rearrange the plan and,

3. the slower participants were more likely to fix the wrong situation.

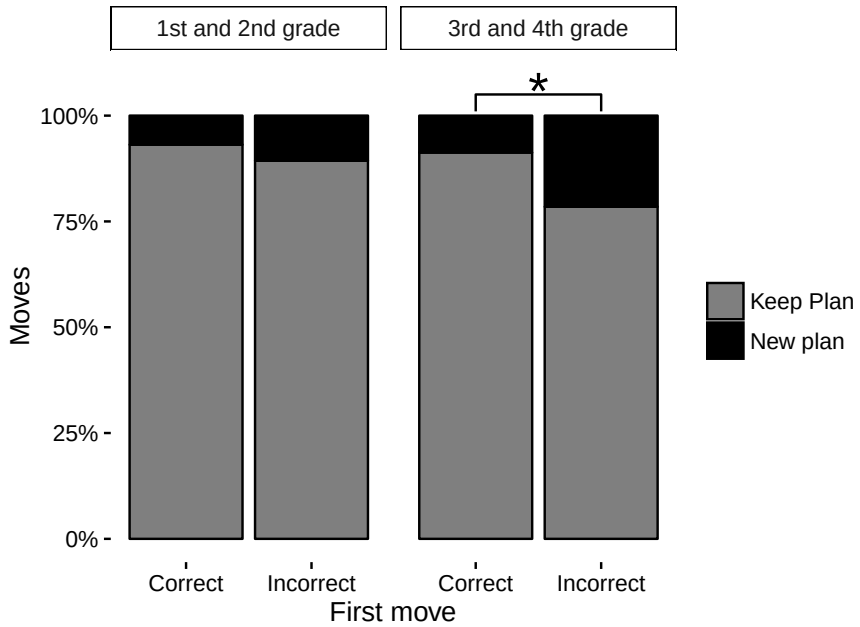


Fig. 6.4: **Difference of change of mind by age.** If we split the population by age in two groups: 1st and 2nd grade children and 3rd and 4th grade and we analyze if in each group whether the change of plan is possible after placing the first piece (cued piece). We find that only in the oldest children there was a difference ($p = 1.068E - 03$ two-sided Fisher's exact test meanwhile for the youngest group $p = 4.197E - 01$ two-sided Fisher's exact test).

6.4 Discussion and conclusion

Observing how children play a video game on their own is a unique opportunity to explore their minds. In the present study we explore children behavior playing a reasoning game. In this game, children are faced to complex trials with little tricks, which may not lead them to the correct answer.

The first exploration showed that players who played correctly the cued piece are less likely to change their plan than the ones who played it incorrectly. This first result suggests that players pay attention to their play, making more likely to fix a wrong position than displace a well-placed piece.

We explore if the ability of changing the plan is affected by the age of the child. Data showed that younger children were less likely to change. This confirmed H1. This result suggests that younger children have not developed the tools needed to self-asses and re-check initiated plan. Further studies should explore which cognitive abilities are related to this task: which are the competences involved? The study lacks of other cognitive test of the population. This does not allow us to explore possible relations among the ability of change of mind and the performance on cognitive tasks tests. Further studies like this, should measure working memory, planning, temperament and self regulation (as suggested in the literature [79, 17, 110]) to get a better understanding of the phenomenon of change of plan observed.

Later, we explored if the ability of changing the plan is affected by who taught the game rules. The findings showed that those who received instruction from a peer were more likely to rearrange the plan. This confirmed H2. This result arises various possible

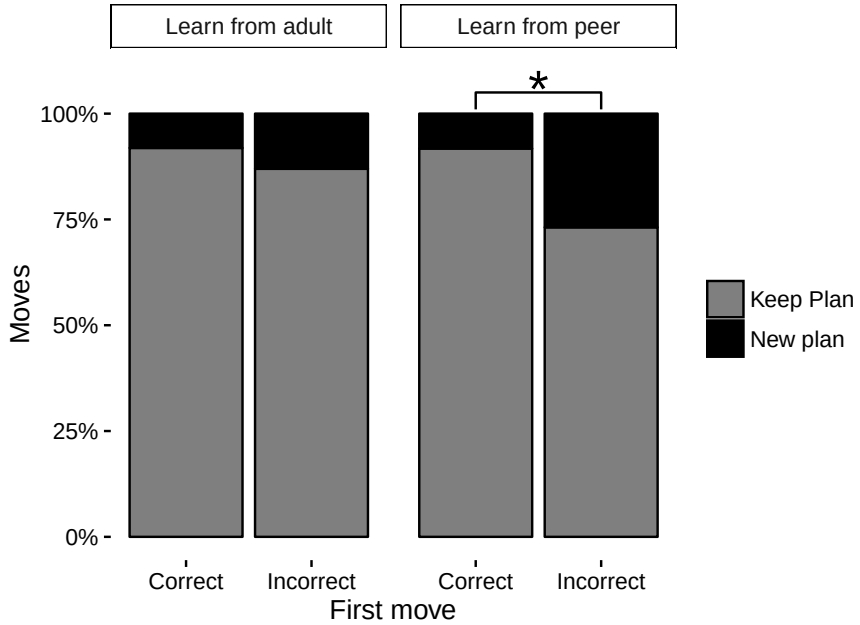


Fig. 6.5: **Difference of change of mind by who was the teacher.** If we split the population in two groups: those who learned from an adult and those who learned from a peer and we analyze if in each group whether the change of plan is possible after placing the first piece (cued piece). We find that only those who learned from a peer were a difference ($p = 2.044E-03$ two-sided Fisher's exact test) meanwhile for those who learned from an adult there was no significance difference ($p = 1.492E-01$ two-sided Fisher's exact test).

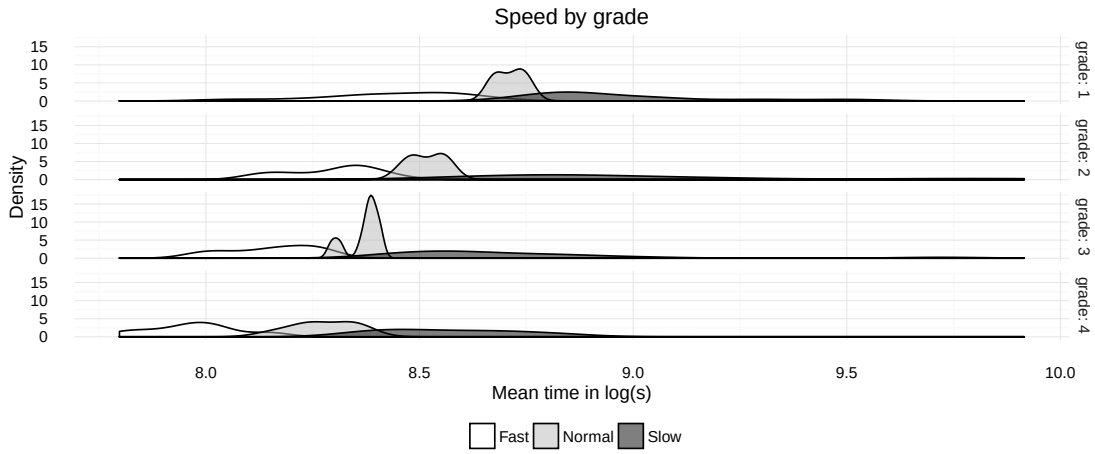


Fig. 6.6: **Speed labeling across grades.** Students where labeled in three groups according to their mean speed movements in relatively to students of the same grade.

explanations: a- children are better teachers than adults; b- since the peer-teacher had the experience playing, she explained in a personalized way with an special pragmatic view and; c- children are more attracted to learn from a peer; among others. All this possible explanations are out of bounds of the present study. Deep understanding of these explanation, if a- or c- are validated, would have great impact on how school learning is

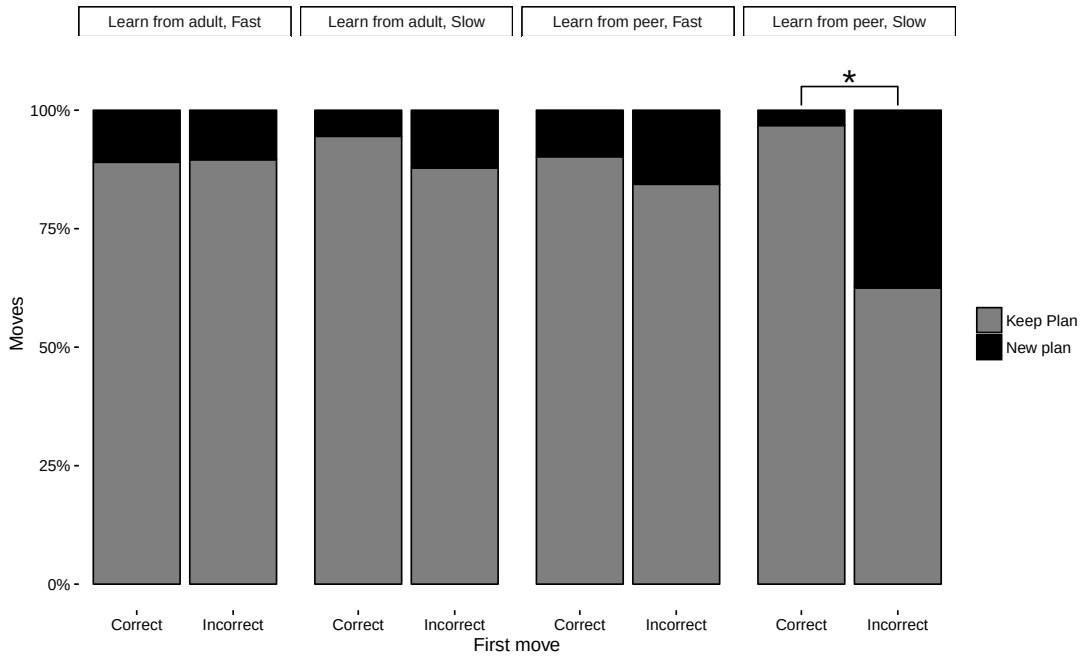


Fig. 6.7: **Speed influence in the subjects.** The Slow players which place incorrectly the cued piece in their first movement are more likely to change their plan than the ones that place it correctly first and learn from a peer ($p = 1.452E-03$ two-sided Fisher's exact test). There was no statistical difference among the Fast players who learn from a peer ($p = 4.976E-01$ two-sided Fisher's exact test). None of the speed conditions showed statistical difference among those who learn from an adult ($p = 1$ fast players and $p = 1.756E-01$ slow players in two-sided Fisher's exact test).

designed. Maybe an schema with peer tutoring should be incorporated.

The last exploration was related to play speed. This can be related to fluid reasoning and inhibitory control. The data showed that slower the participant played, the more likely to fix the wrong situation she was. This confirms H3. The change of mind ability seems to be possible when students play slower. This also suggests further studies in which the time variable can be explored. For instance, the *let's see who finish it first* game may not contribute to a better possibility of fixing wrong situation. It could be needed to promote good results instead of quick solves, not only in schools, but also in games.

Resumen en español

Cambio de plan

¿Cómo aprenden los niños a resolver una tarea específica? ¿Y cómo solucionan las dificultades que pueden tener en el proceso? Avanzar en respuestas a estas preguntas contribuye al esfuerzo de la comunidad científica por realizar la ingeniería reversa de la mente. En este marco, el capítulo se centra en la comprensión del aprendizaje humano y la inferencia.

Para ello, se analizan las interacciones de niños, de entre 6 y 10 años, con un videojuego que emula una tarea de planificación. Concretamente se utilizó la versión digital de Chocolate Fix.

Asimismo, el estudio buscó explorar y caracterizar bajo qué condiciones, los niños pueden identificar sus propios errores y arreglarlos (o no). Los niños se encontraban separados en dos grupos: aquellos que aprendieron a jugar al videojuego con la explicación de un adulto y los que aprendieron de un compañero al que ya le había explicado un adulto con anterioridad.

Este estudio fue realizado en la sala de computación de una escuela primaria argentina para poder analizar el comportamiento de los niños en el ambiente más ecológico posible. Estas condiciones favorecerían la captación de sus habilidades para resolver los problemas. El videojuego fue utilizado con la plataforma Mate Marote detallada en el Capítulo 5, con la cual se registraron los datos.

De los datos recolectados, pudimos observar y medir la capacidad de los niños para arribar a una solución, cuando hicieron una mala jugada. Los resultados del estudio sugieren que esta capacidad de rehacer un plan para corregir los propios errores, no es innata, sino que se va desarrollando con la edad. También se demuestra que la forma en que cada niño/a aprendió a jugar, incide en la capacidad de volver a planificar: aquellos que aprendieron de un par son más eficaces en el desarrollo de estrategias alternativas que quienes lo hicieron con la explicación de un adulto.

7. CONCLUSION

Le das la copa al fin al vencedor, tarea fina perdida en mi soledad.
–Patricio Rey y sus Redonditos de Ricota

The world is going through one of the most important technological advances of history: the Information and Communication Technologies (ICT) are becoming part of the everyday life, in education, work and leisure. Access to ICT has become massive in a world wide scale. In Argentina institutional programs –Conectar Igualdad and Joaquín V. González (La Rioja), Plan Sarmiento (CABA)– provided computers to primary and secondary school students. Nowadays the population with access to internet is over 70% [7].

In this digitalized context many efforts emerged to promote education through ICT, such as computerized tutors, MOOCs (Massively Open On-line Courses), Games for Learning, etc. All of them have one challenge in common: gain students' attention and engage them with the activities. One approach to tackle this challenge is Gamification – the application of typical elements of game playing (e.g. point scoring, competition with others, rules of play) to other areas of activity –, which has been proven to motivate students and produce engagement in educational activities [76].

The inclusion of ICT in everyday life and, in particular, in schools generates lots of human-computer interactions which can be precisely registered and analized. These interactions can be used to explore *indirectly* how humans' mind work.

To improve education strategies, understanding how students learn is needed. There are many efforts from different perspectives in reverse engineering of the mind: analyzing log files or click streams of software usage; the analysis of gazed elements of a visual scene; the study cell-phone calls dynamics as large-scale networks; among others. With the knowledge built with this techniques new and more accurate learning methods can be designed.

The new methods for learning can be implemented and tested through interventions. These studies can handle two objectives. One of them is oriented to test the effectiveness of the proposed methodology. The other is collecting new data to answer empirical questions about the cognitive processes. This new knowledge will be used to improve the original method.

This thesis was a combined effort of Computer Science, Cognitive Science and Neuroscience fields which aimed to contribute to the design of software for cognitive training, looking for incorporate adaptive algorithms and tools for analysis to study cognitive processes and its impact.

From a vast amount of possible educational intervention tasks, we focus in planning, i.e. the ability to think a course of actions before applying the decisions made. We analyzed two tasks: Tower of London and Chocolate Fix. The characterization provided about this two tasks is important for further uses. The better the knowledge of the dimensions involved in the tasks, the better the understanding and results this task will produce. We expect the analysis exposed here to be useful when using these two tasks. In addition, it should be used as an example on how important is to understand the task used for research. The *black box* use of tasks may block the opportunity of a correct understanding

of the phenomenon observed of human behavior because used trials can hide particularities used to solve those task.

We presented a study where we measure how the usage of cell-phones interferes and affects perception and quality of face-to-face communication. After a brief (4 minutes) communication, data reported by the participants showed that total amount of attention is the major factor driving subjective beliefs about the message (the story being told) and of the conversational partner. The effect is observed on both the emitter and the receiver and is mostly independent on how attention is distributed in time. Interruptions during day to day communication between peers and also with children are extremely frequent. Our data should provide a note of caution, by signaling the consequences of these windows of neglect on the teller and receiver of a story.

From another perspective, we explored the transmission of concepts using a cell-phone app, showing the effectiveness of the different strategies used by the participants. The study was also a crowdsourcing experience, where we use *the crowd*, not only to solve the tasks, but also the assess in an impartial manner the process of communication involved. We found that written text is remarkably ambiguous when transmitting geometric figures. Textual style affects the effectiveness of the transmission: algorithmic description is the most effective method compared with others, such as metaphors.

The study of human computation requires an indirect way to observe the reflections of the mind. In this thesis we presented the first stage of the development of a free tool for educational interventions, called Mate Marote, which collected users' behavior playing video games. This stage was an articulation University-Industry experience. The result was the development of a robust and scalable tool which combines research needs with industry knowledge on heavy load systems.

The first test of the developed tool was an intervention in a primary school in including Chocolate Fix –the game exposed in this thesis–. We used the planning game with certain tricky puzzles to challenge the students. We could observe and measure the ability of children to fix a partial solution when they face a wrong position. Results suggest that this ability of re-arranging a plan is not universal and it is developed with the age. Data also suggests that it is modulated by the way each child learned to play –from an adult or from a peer–, showing that peer tutoring is a more effective strategy to develop this ability than been taught by an adult.

Summarizing, this thesis proposed, designed and implemented –with professional support of the Industry– Mate Marote, a robust and scalable tool for educational interventions using ICT. This work also includes analysis of the structure of two planning games that may benefit the proper use in educational interventions. Mate Marote was tested successfully in an intervention with the game Chocolate Fix in a primary school. In addition, two studies explore the impact the usage and inclusion of ICT in the everyday life. This thesis contributes from an interdisciplinary perspective to the introduction of techonology in educational interventions towards the better understanding of learning and cognitive processes in a quantitative evidence-based research.

7.1 Limitations

This thesis faced great challenges from theoretic to logistical problems. For example: establishing a common language among different disciplines, interacting with a software factory, dealing with times of shows like TEDx talks.

But it also faced limitations and problems. Reporting these limitations it is as important as reporting the results. The only way to avoid repeating mistakes is being aware of them.

The first problem we faced was a logging problem of our platform. For unknown reasons some sessions of some children just disappeared, some of them partially, other completely. This made us lose a lot of data. Part of this lost was not even possible to track and understand because the logs were messy and we were not able to reconstruct the session. This experience fed the knowledge of our team, and it leaded to better practices in logging human-computer interaction.

Other limitation was the difficulty of acquiring quality data which was not as easy as initially imagined. Collecting data from schools require a lot of paperwork and massive on-line systems require a lot of efforts to keep alive a web page with games. If the page does not offer novel games and challenges the audience reduces and it is harder to make them back.

In 2013, when we explored the crowdsourcing experiment, we did not have the time nor the experience to have a multi-platform app. This incompatibility made us lose many participants. iphone's are popular among the participants of TEDx talks. This also was useful to understand the importance of developing in a technology which worked cross-platform.

Unfortunately, we did not have enough time to explore tutoring in our games further than the scaffolding provided by the sequence of levels and the dynamics within this sequence.

7.2 Further work

The development of this thesis raised many questions. Although these question were not able to be addressed, these are relevant questions that deserved to be explored in the future.

A brief list of the future directions.

- There is still un-analyzed data. From Chocolate Fix we have two data sets remaining: Berkeley 2014, which includes Eye Tracking of around 60 undergrad students and; Primary School 2016, around 80 of 6- to 8-years-old students. These studies will also arise new questions.

- After the experiment with Chocolate Fix, many question are still without answer: a- Are children better explaining than adults?; b- What are the difference between an adult and a student teaching? Is it language? Is it the age? Are different messages?; c- Are there other possible puzzles to explore the change of mind ability?

- In the path of defining complexity and difficulty, a massive experiment which explores all possible Tower of London puzzles would be a great project to validate the hypothesis of the structure of the game explored in this thesis.

- We still need to analyze data from the new testing protocol designed with Tower of London. This protocol order was designed based on the chances of winning if playing as a random agent. Is this sequence of trials sorted by difficulty?

- Some projects developed were not included in the thesis, one of these projects is a massive problem solving (participants had to find a minimum path in a squared road-map). The experiment had 3 stages: the first one was in person; the second one was solved using Google Forms; and the third stage was done using the platform Mate Marote.
- Finally, after the success of using crowdsourcing to assess productions of participants. The design and implementation of a system to use this strategy for further studies would be extremely useful.

Resumen en español

Conclusiones

Esta tesis fue un esfuerzo combinado de los campos de Ciencias de la Computación, y Neurociencia para contribuir al diseño de software para el entrenamiento cognitivo, mediante la incorporación de algoritmos adaptativos y herramientas de análisis para estudiar los procesos de aprendizaje y su impacto.

De una gran cantidad de posibles tareas educativas, nos centramos en el estudio de la planificación, es decir, la capacidad de pensar un curso de acción antes de ejecutar las decisiones tomadas. Lo hicimos mediante el análisis de dos pruebas: Torre de Londres y Chocolate Fix. Su caracterización y la acabada comprensión de la estructura interna de las mismas es importante para su mejor utilización en intervenciones e investigaciones futuras. Cuanto mayor sea el conocimiento de las dimensiones implicadas en la planificación, será más efectivo su uso para el entrenamiento cognitivo. En el caso de fines de investigación, los datos recolectados mediante estos juegos, podrán tener una mejor comprensión. El uso a ciegas de tareas puede producir la pérdida de una oportunidad para una comprensión más acertada del fenómeno observado de la conducta humana, ya que pueden ocultar particularidades utilizadas por los participantes (de manera consciente o inconsciente) que les sirvieron para resolver las tareas pero que resultan desconocidos para el investigador.

Asimismo, presentamos los resultados de un estudio sobre cómo el uso de dispositivos móviles interfiere y afecta la percepción y la calidad de la comunicación cara a cara, y mostraron que la cantidad total de atención al relato del otro, es el factor principal que impulsa creencias subjetivas sobre el mensaje (la historia que se cuenta) y sobre el compañero de charla. El efecto se observa tanto en el emisor como en el receptor y es independiente de cómo se distribuye la atención en el tiempo. Las interrupciones durante la comunicación diaria entre compañeros y también con los niños son frecuentes. Lo cual es algo a tener en cuenta en la vida cotidiana y especialmente en los procesos de enseñanza, en donde la afectividad y los vínculos interpersonales son fundamentales.

Desde otra perspectiva, se exploró la transmisión de conceptos utilizando una aplicación de teléfono celular, mostrando las diferencias en la efectividad de distintas estrategias escritas elaboradas por los participantes. El estudio fue también una experiencia de *crowdsourcing*, donde usamos a la *multitud*, no solo para resolver las tareas, sino también para evaluar de manera imparcial y ciega el proceso de comunicación en estudio. Encontramos que el texto escrito es notablemente ambiguo cuando se transmiten figuras geométricas. El estilo utilizado en la redacción, afecta fuertemente la efectividad de la transmisión: la descripción algorítmica es el método más eficaz comparado con otros, como las metáforas.

El estudio de como los humanos realizan cómputos, requiere una forma indirecta de observar las reflexiones de la mente. En esta tesis presentamos la primera etapa de desarrollo de una herramienta de código abierto y de uso gratuito para intervenciones educativas, Mate Marote, que captura el comportamiento de los usuarios mientras juegan a videojuegos. Esta etapa fue una experiencia de arti-

culación Universidad-Industria que permitió la construcción de una plataforma robusta y escalable que combina las necesidades de investigación con el conocimiento del sector privado de software en sistemas preparados para alto régimen de carga.

La primera prueba de la herramienta Mate Marote fue una intervención en una escuela primaria utilizando el juego Chocolate Fix. Utilizamos este juego de planificación con puzzles especialmente diseñados para desafiar a los estudiantes. Allí observamos y medimos la capacidad de los niños para llegar a modificar su estrategia de resolución ante una posición incorrecta. Los resultados sugieren que esta capacidad de rehacer un plan no es innata, sino que se desarrolla con la edad. Además se demuestra que la forma en que aprendieron cómo se juega, modula dicha capacidad: los que aprendieron de sus pares son más eficaces para redefinir su estrategia luego de una equivocación que aquellos que aprendieron de un adulto.

Esta tesis contribuyó al desarrollo y análisis del uso de las TICs en intervenciones educativas y el modo en que éstas afectan la vida cotidiana, la afectividad, los vínculos interpersonales y la comprensión de un mensaje. En este sentido, aportó a la construcción de Mate Marote, una herramienta robusta y escalable para intervenciones educativas basadas en las TICs. En esta dirección se realizó un análisis exhaustivo de la estructura de dos juegos de planificación para un uso adecuado y eficaz de los mismo en intervenciones educativas. En síntesis, esta investigación aporta evidencia empírica que permite realizar recomendaciones sobre los diversos elementos a considerar en la implementación de intervenciones educativas basadas en las TICs y presenta buenas prácticas de incorporación de tecnologías en el aprendizaje y para el logro de mejores procesos cognitivos.

APPENDIX

A. APPENDIX

A.1 Tower of London protocol proposal

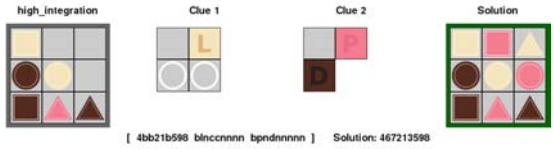
Order	Source	Target	#Path	Distance	P(win)
1			1	1	0.5
2			1	1	0.5
3			1	1	0.33
4			1	1	0.33
5			1	1	0.25
6			1	2	0.2
7			1	2	0.2
8			1	2	0.14
9			1	2	0.14
10			1	2	0.12
11			1	3	0.09
12			1	3	0.09
13			1	3	0.09
14			1	3	0.08
15			1	3	0.08
16			1	4	0.05
17			1	4	0.05
18			1	4	0.05
19			1	4	0.04
20			1	4	0.04

Order	Source	Target	#Path	Distance	P(win)
21			1	5	0.03
22			1	5	0.03
23			1	5	0.03
24			1	5	0.02
25			1	5	0.02
26			1	6	0.02
27			1	6	0.02
28			1	6	0.01
29			1	6	0.01
30			1	6	0.01
31			1	7	0.01
32			1	7	0.01
33			1	7	0.01
34			1	7	0.01
35			1	7	0.01
36			2	8	0.01
37			2	8	0.01
38			3	8	0.02
39			3	8	0.02
40			3	8	0.02

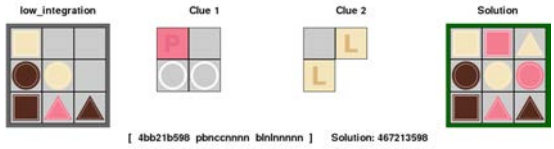
Tab. A.1: Tower of London protocol proposal according the definition of 2.1

A.2 Chocolate Fix

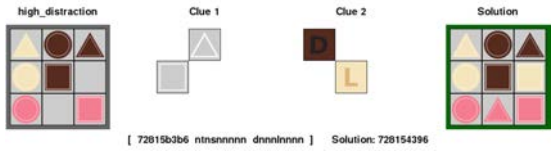
In the Fig. A.1 the puzzles explained in Section 2.1 are shown.



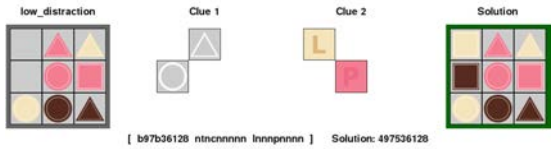
(a) High integration



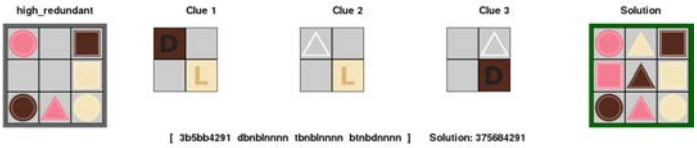
(b) Low integration



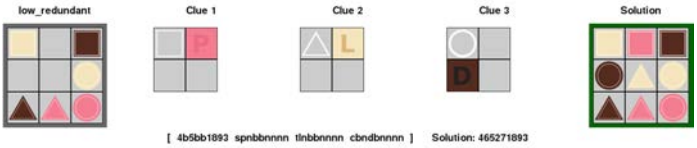
(c) High distraction



(d) Low distraction



(e) High redundant



(f) Low redundant

Fig. A.1: Manipulations of the puzzles

In the Fig. A.2 the puzzles analyzed in experiment of Chapter 6 are shown.

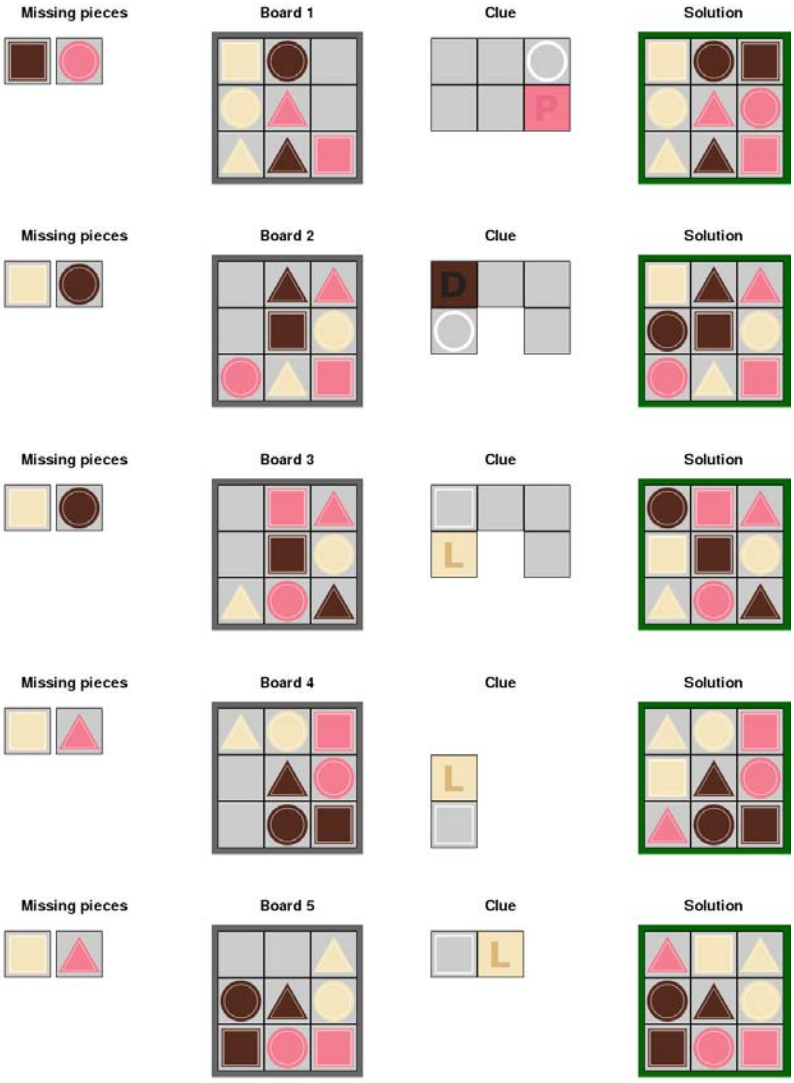


Fig. A.2: Puzzles of the study

A.3 TEDxperiments - Dialogs experiment materials

For the experiment of Chapter 3 the participants interact in pairs (the Speaker and the Listener) and they receive an envelop with two papers with instructions, one for each. The instructions where specially folded in 3 (see Fig. A.3): when it is closed no-one can see any instructions; when the first fold is unfolded, the instruction for the game can be read; when the second fold is unfolded the questionnaire is available to be answered. To prevent the second fold is unfolded before the *performance*, an staple was placed in the border. So, the second unfold also has to remove the staple.

A copy of a pair of instructions is included in Fig. A.4.

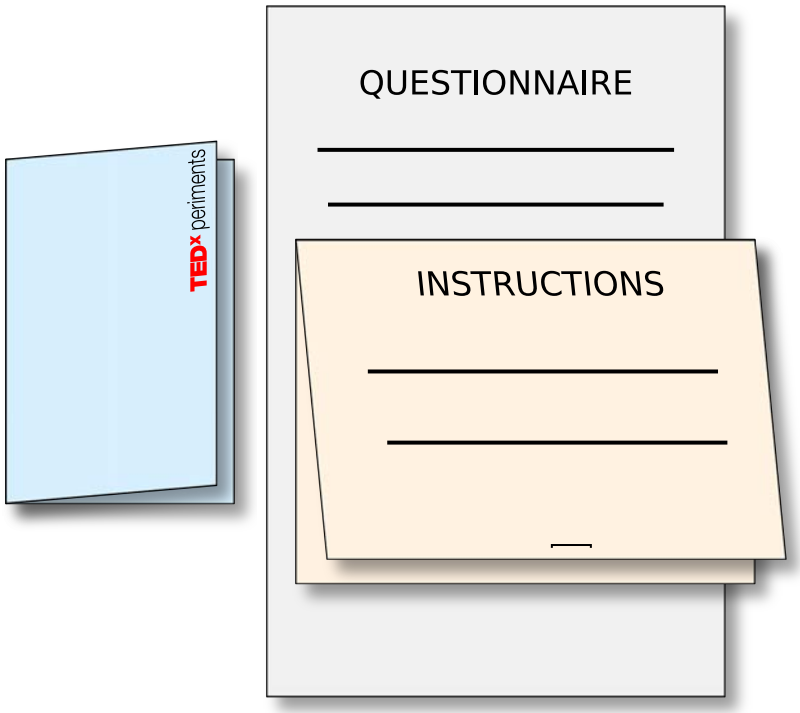


Fig. A.3: Experiment instructions folding in three stages: completely folded (left); partially folded, with instructions for the performance (right on-top); unfolded with the questionnaire (right bottom). In the bottom of the instructions the staple can be seen.

A.4 TEDxperiments - Draws experiment

A.4.1 Paper-version

In the Fig. A.5 an example of the experiment of Chapter 4 is shown.

B0001

TEDxperiments

Elegí una historia. La historia que quieras, puede ser personal o no.

El único objetivo es que sea lo más interesante posible para tu compañer@, a quien se le vas a contar. Pensá que estamos en TED y tenemos que elegir historias que nos partan la cabeza, que dejen una traza, que emocionen.

Tenés exactamente 4 minutos para contarla, ni un segundo más.
Te vamos a avisar cuando haya pasado un, dos y tres minutos como para que puedas ir redondeando.

Antes de empezar a contar la historia indica si la historia que elegiste es:
1) Entretenida (1-10)
2) Apasionante (1-10)
3) Emotiva (1-10)

Sobre la historia
En retrospectiva, la historia que elegiste es:
Entretenida (1-10)
Apasionante (1-10)
Emotiva (1-10)

Sobre el otro
Tu compañero de juego te parece una persona:
Interesante (1-10)
Atractiva (1-10)
Agradable (1-10)

La manera que contaste la historia fue
Efectiva (1-10)
Fluida, con buen ritmo (1-10)
Recibida con interés (1-10)

B0001

TEDxperiments

Tu compañer@ va a contarte una historia durante cuatro minutos.

Al principio tu rol es ignorar al que te está contando la historia. Podes chequear tu teléfono (twittear, mensajear, leer email...) buscar algo en el bolso, atarte los cordones, interrumpirlo con algo irrelevante (por ejemplo ofrecerle un caramelo). Es importante que lo ignores claramente pero que no sea tan exagerado como para que sea obvio y grotesco. Simplemente que parezca que realmente no te importa como en tantas situaciones de la vida donde estas compenetrado con tu teléfono mas que con el que te está hablando.

Cuando avisemos que faltan dos minutos tenes que cambiar de actitud. Como antes, esto tiene que ser sin ser demasiado obvio ni tremendamente exagerado. Progresivamente mostra mas y mas interés en la historia, si queres asintiendo o marcando algún comentario que indica que la seguís con pasión. Al final de la historia deberías estar fuertemente compenetrado.

Este cuestionario no lo va a ver tu compañero.

Sobre la historia
La historia que eligió tu compañer@ es:
Entretenida (1-10)
Apasionante (1-10)
Emotiva (1-10)
Contada con buen ritmo (1-10)

Sobre el otro
Reacción de tu compañero cuando lo ignoraste
(si dividís tiempo total en dos mitades):

1 mitad Abandono - persistencia
(1 abandono por tu actitud - 10 se esforzó por ganar mi interes)
2 mitad Abandono - persistencia
(1 abandono por tu actitud - 10 se esforzó por ganar mi interes)

Tu compañero de juego te parece una persona:
Interesante (1-10)
Atractiva (1-10)
Agradable (1-10)

Sobre tu performance
Evalúa la performance de tu rol
(1 suave, 5 justa 10 exagerada)
Uso de teléfono o símil como herramienta para ignorar
(1 nada 10 mucho)
Te parece que afecto a tu compañero
(1 no para nada, 10 si lo afligió)

Fig. A.4: Instructions for the experiment: on top the Speaker role instructions (on the left the instructions, the outside, and the folding guides), on the bottom the Listener role instructions.

A.4.2 Images used in the experiment

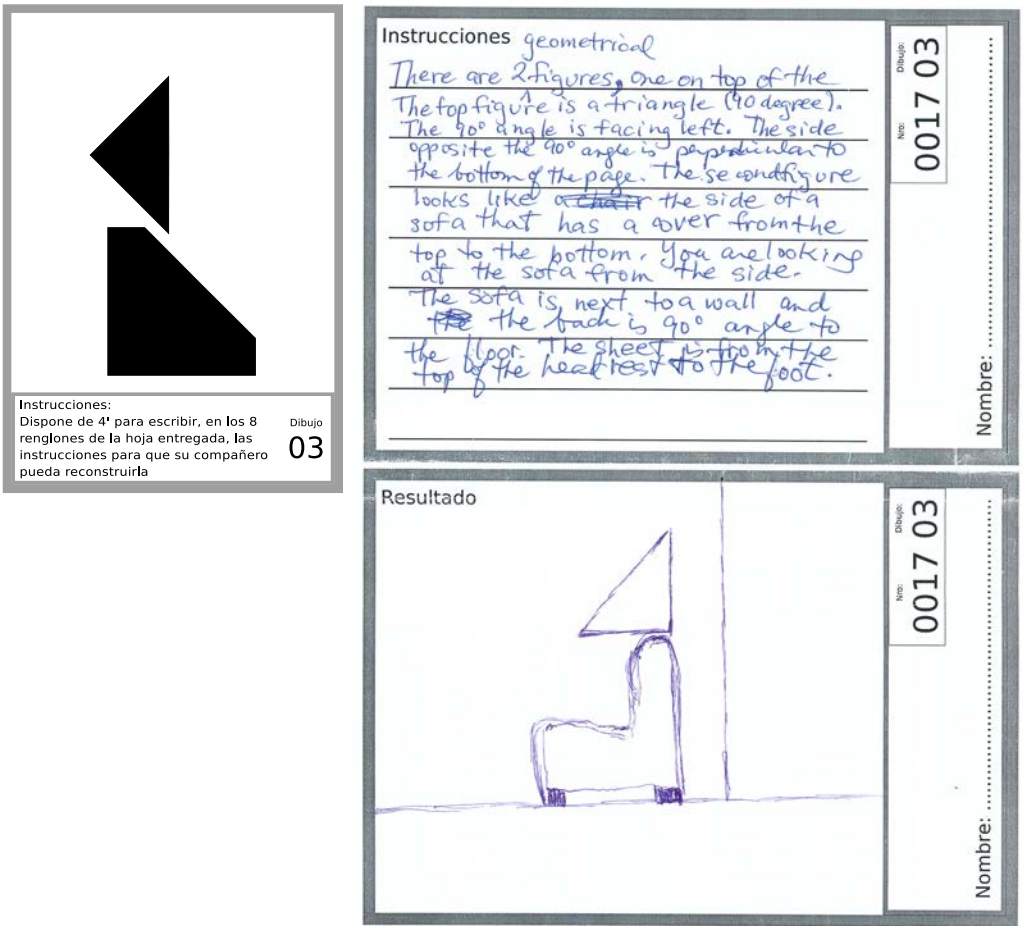


Fig. A.5: Experiment material. Top-left: original image received by the first participant who write down instructions (top-right). The second participant receive the instructions (top-right) and makes a draw according to what she understand from the text (bottom-right).

A.4.3 Mockups of the app

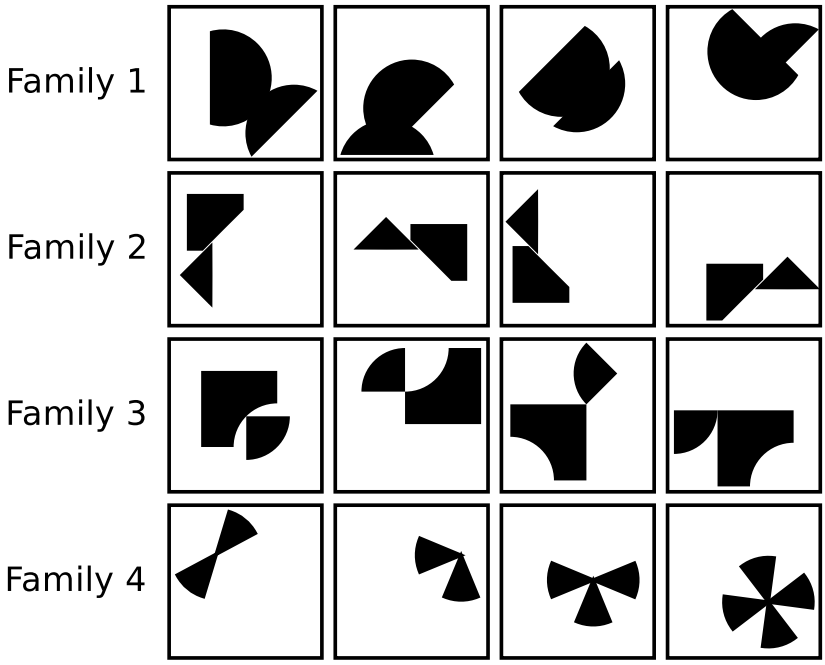


Fig. A.6: Source images of the experiment. Each row represent a family of similar images. The subject first received a randomly select image of one family in the writing stage, for the drawing part the description received was produced from an image of a different family of draws.

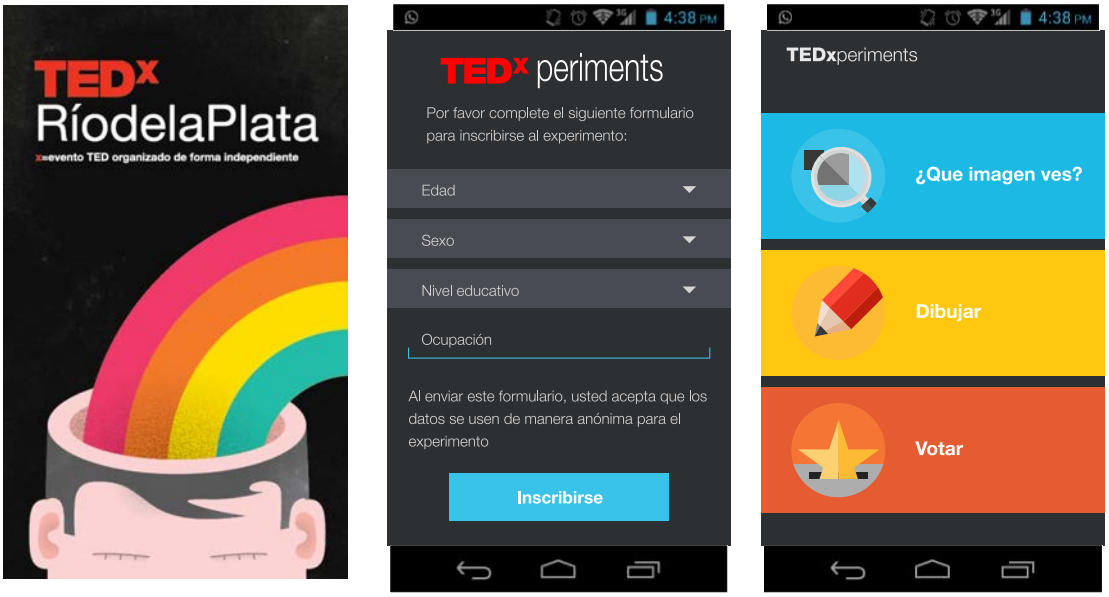


Fig. A.7: Mockups: Splash-screen (left); register and accept terms and conditions (middle) and; menu (right). The menu only has one of the three options enable at a time. It depends on the flow of the app.

A.4.4 Server database schema

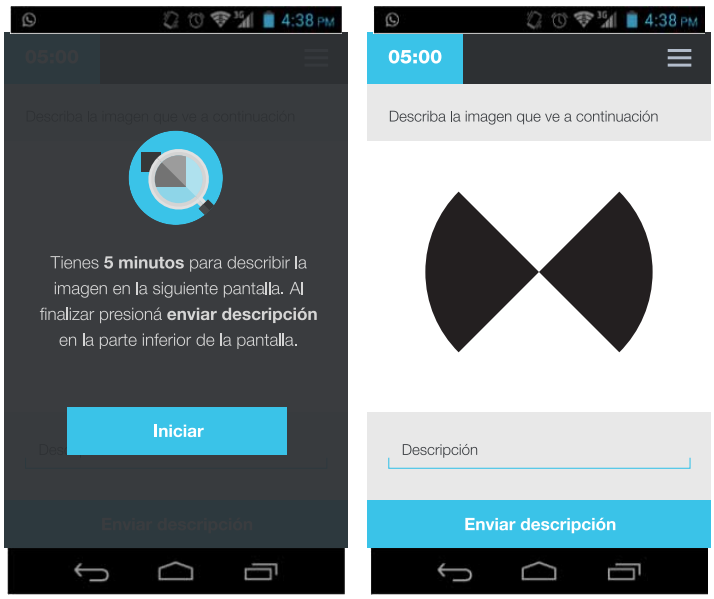


Fig. A.8: Mockups: Write stage instruction (left); write scenario with the image to be described (right).

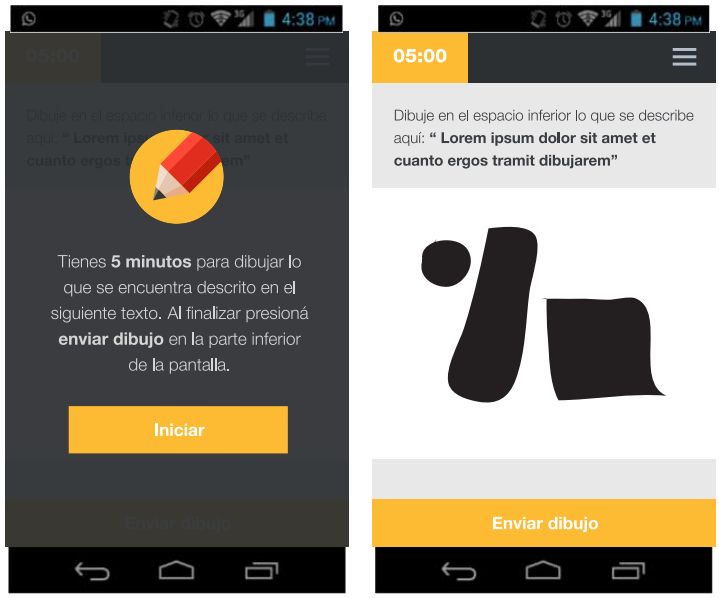


Fig. A.9: Mockups: Draw stage instruction (left); draw scenario with the description on top and the canvas to draw in the bottom (right).

A.4.5 Example of productions and assement

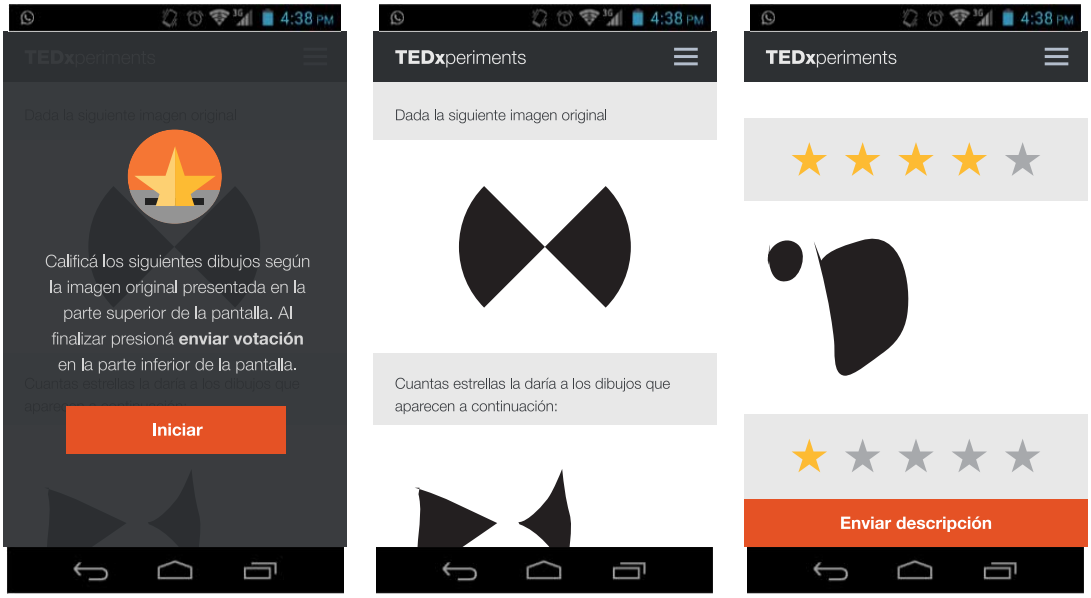


Fig. A.10: Mockups: Assess stage instruction (left); assess examples (middle and right)

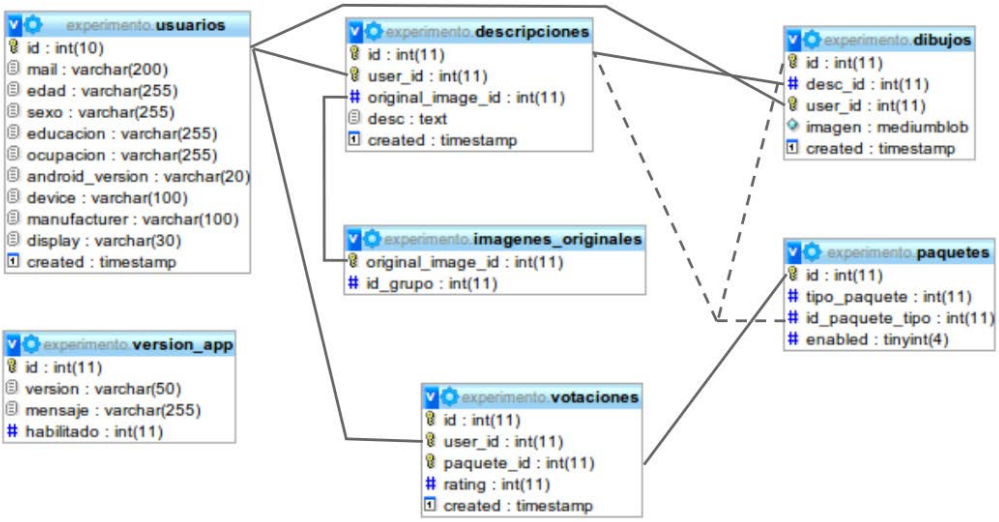


Fig. A.11: Server database schema.

A.5 Mate Marote - Games documentation

A.5.1 ANT

Fundamentos

El Test de las Redes Atencionales (Attentional Network Test, ANT) es una adaptación, realizada por Rueda y colaboradores [111], de una versión para adultos. Fue diseñada con el objeto de evaluar en niños de hasta alrededor de diez años tres aspectos fundamentales del procesamiento atencional: alerta, orientación y control ejecutivo. La tarea se presenta como un juego de computadora y requiere que el niño determine si un animal central

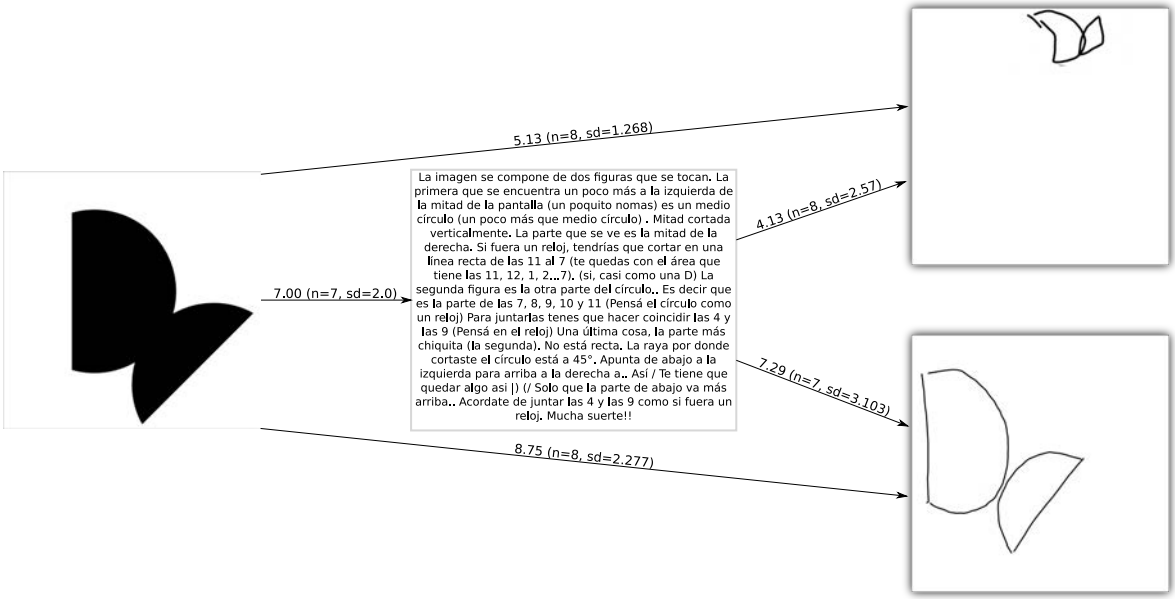


Fig. A.12: Example of a description with a procedural description.

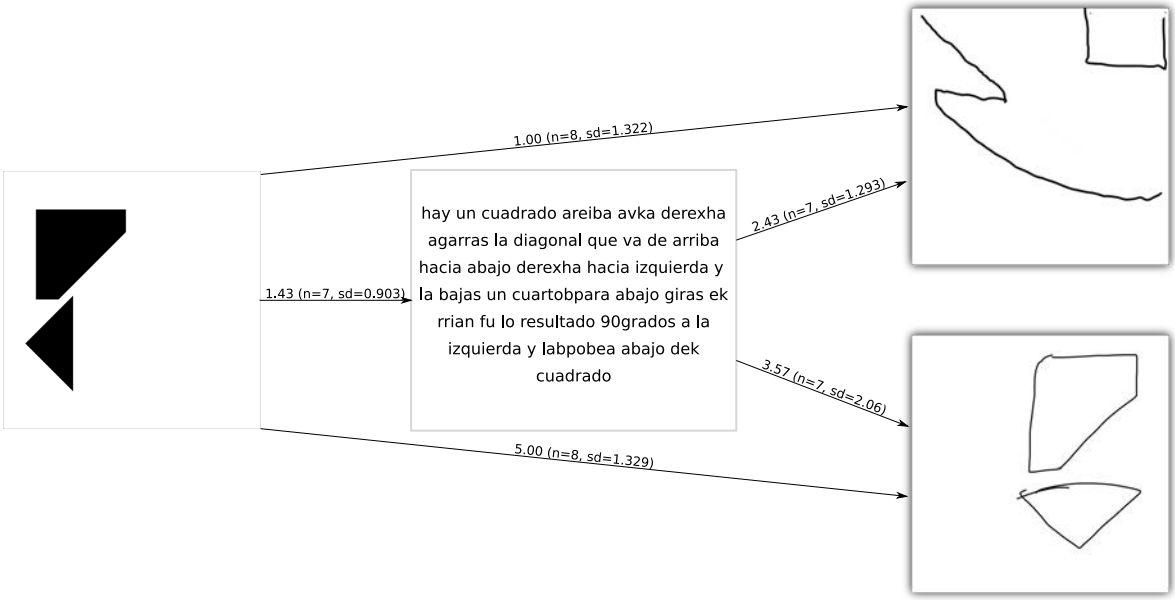


Fig. A.13: Example of a description with lots of mistakes but still effective.

apunta hacia la derecha o hacia la izquierda y que, lo más rápidamente posible, presione uno de los dos botones del mouse —que se encuentra exactamente frente a ellos— (el derecho para indicar derecha, el izquierdo para indicar izquierda).

La tarea consiste en una sucesión de trials separados en bloques. Cada bloque tiene un animal y color de fondo diferente (un pez con fondo celeste, un ratón con fondo rojizo o un pájaro con fondo verde, etc.; Fig. A.16). Entre bloques, los niños pueden descansar el

Fig. A.14: Example of an extremely metaphoric description and its difference of score among the different stages of the study.

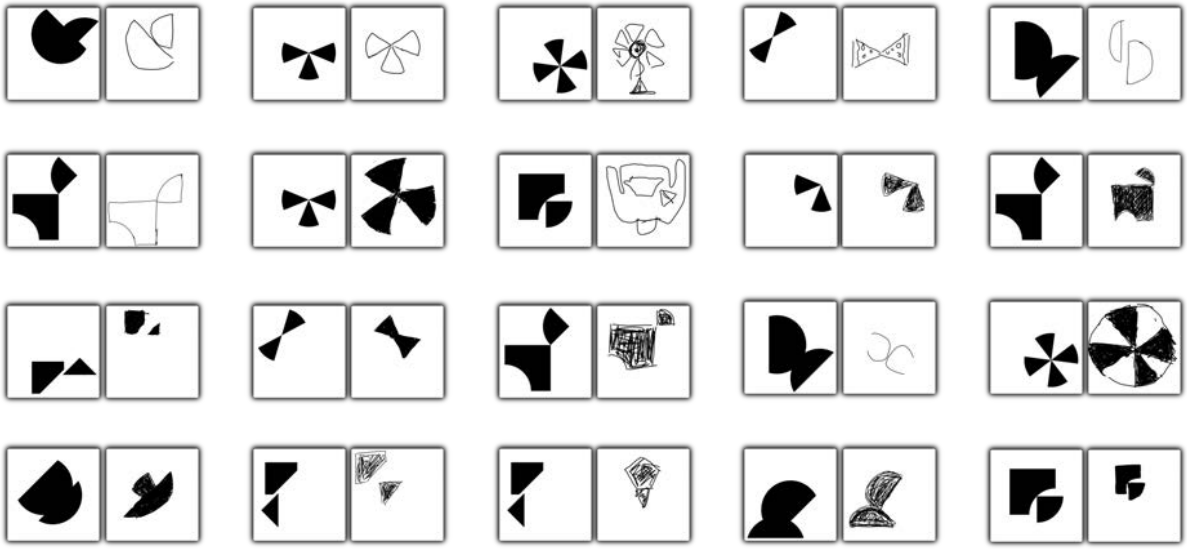


Fig. A.15: Example of pairs of images and the draw-result after transmission.

tiempo que precisen. Cada trial comienza con la presentación de una cruz de fijación en el centro de la pantalla. En cada trial, el animal central se presenta flanqueado por un animal de cada lado. Estos animales pueden mirar en la misma dirección que el central (caso trial congruente, ej. Fig. A.16 izq. arriba) o en la dirección opuesta (caso trial incongruente, ej. Fig. A.16 izq. abajo).

Además, en algunos trials aparece —por 150 ms— una de tres claves espaciales o no aparece ninguna clave (figura 2). En ningún momento se hace mención sobre las claves a los niños. Luego de la respuesta de los niños, se les puede dar feedback.

En cada bloque se definirá:

- la cantidad de trials,
- el animal que aparecerá en ese bloque,
- el tiempo entre trials (considerado como el tiempo desde que termina el feedback de

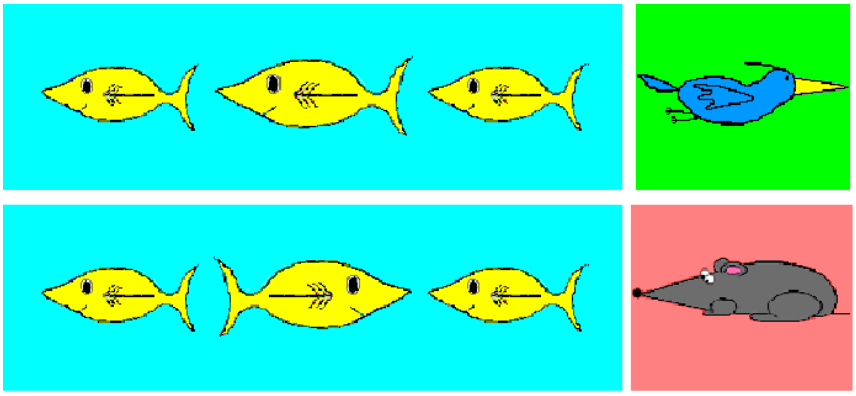


Fig. A.16: Dibujos que aparecen en la tarea ANT para niños. Peces, ratones, pájaros. Izquierda superior: ejemplo de trial congruente, izquierda inferior: ejemplo de trial incongruente.

un ensayo hasta que aparece la cruz de fijación del ensayo siguiente, llamado "ISI" en la Fig. A.17. Idealmente la cruz de fijación no debería aparecer allí),

- el tiempo en que aparece la cruz de fijación (panel izquierdo, figura 3),
- el tiempo de duración de las claves (panel "Clave", figura 3; ahora es de 150 ms. En caso de que la clave sea "Sin clave", la pantalla de cruz de fijación se mantendrá visible por 150 ms. El punto es que todos los trials tengan la misma duración.),
- el tiempo entre la desaparición de las claves y la aparición de los animales (tercer panel "Cruz", figura 3; ahora es 450 ms. En este tiempo debe seguir visible la cruz de fijación),
- y la composición de los trials:
  - proporción de congruentes e incongruentes,
  - proporción de los distintos tipos de clave (incluyendo "Sin clave"),
  - posición de los animales,
  - lado de la respuesta correcta,
  - si se da feedback o no (por default, si).

Un ejemplo de un trial (ver un esquema detallado de los tiempos en la Fig. A.17). La pantalla aparece del color característico del bloque y con una cruz de fijación en el centro. En ese momento puede o no aparecer algún óvalo ("Clave", ver Fig. A.18) y, luego, aparecen tres animales. El niño debe presionar lo antes posible la tecla que indica el lado al cual mira el animal del centro, independientemente del lado al cual miren los flancos. El ejemplo de la figura 1 (der.) sucede en el primer bloque (indicado por los peces y el fondo celeste) y la respuesta correcta es presionar la tecla izquierda. Un sonido indica el feedback y, además, si la respuesta es correcta, cambia la imagen y el animal se mueve. El estímulo queda presentado en pantalla un cierto tiempo (en el ejemplo, 1700 ms) y, si no hay respuesta en ese tiempo, el ensayo se considera incorrecto y se da feedback negativo. El intervalo entre trials (ISI) es de 1000 ms. Luego de que se suceden todos los trials definidos del bloque, aparece una pantalla de descanso y el niño puede tomar el tiempo que quiera hasta comenzar el siguiente bloque, que será igual excepto por el color de fondo

y el animal. Cabe aclarar que antes de comenzar el primer nivel, los niños tienen una serie de trials de prueba, que pueden ser definidos como un nivel diferente. Son exactamente iguales que el resto de los trials excepto que los niños no tienen tiempo máximo para responder (o el tiempo es mucho).

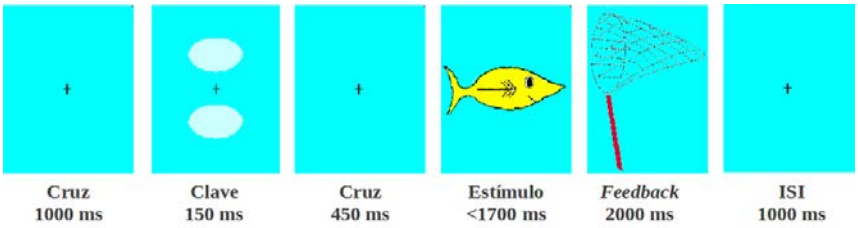


Fig. A.17: Desarrollo completo de un ensayo de la tarea ANT para niños. Sucede de izquierda a derecha. El tiempo de presentación del estímulo puede ser menor a 1700 porque, si el niño tarda menos, el estímulo desaparece y aparece el feedback. El tiempo entre ensayos (ISI) es de 1000 ms.

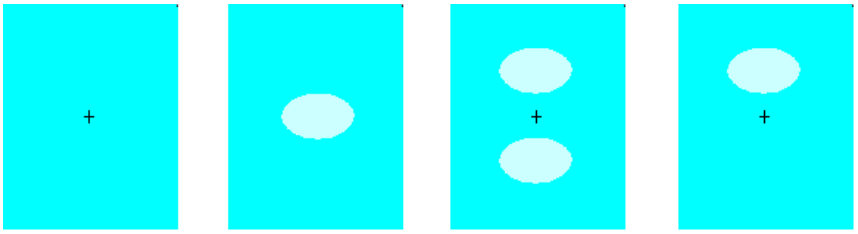


Fig. A.18: Claves que aparecen en la tarea ANT para niños. De izquierda a derecha: "Sin clave", "Central", "Doble" y "Espacial". La clave "Espacial" consiste en un único óvalo que puede aparecer arriba o abajo de la cruz de fijación (en este ejemplo se encuentra arriba).

Niveles

Hay un único tipo de bloque/nivel. Lo único que cambia es la configuración de los trials y la imagen/color de fondo que se muestra. Entre bloques pueden cambiar los tiempos de presentación (del estímulo, de la cruz, de la clave, del feedback, ISI).

Pantalla

No tenemos ejemplo de esta pantalla porque este es un juego nuevo. Ver pantalla ejemplo (por proporciones y posiciones) en la Fig. A.19.

Controles

Hasta el momento no hemos jugado a este juego de forma no supervisada y en el juego presencial los niños usaban los botones del mouse porque un adulto lo posiciona correctamente. Como no será posible controlar la posición del mouse para los niños que jueguen remoto, proponemos que en la versión en Web (PC con teclado) las teclas para indicar derecha o izquierda sean las flechitas del teclado correspondientes. Si esto no fuera posible, podrían ser dos teclas centrales y adyacentes del teclado, como v y b ó b y n. Para versiones táctiles hay que encontrar algún recurso como puede ser botones en el pantalla.

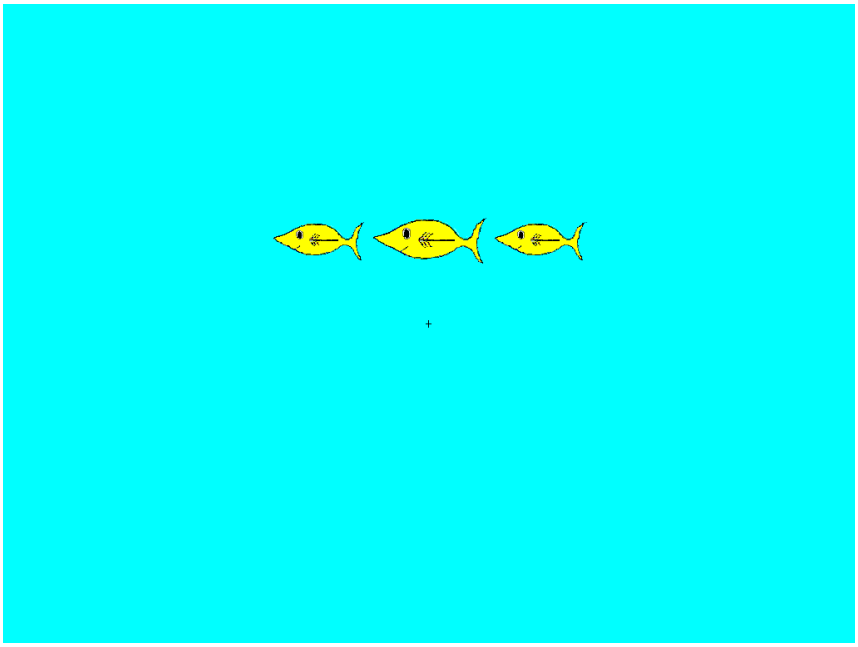


Fig. A.19: Captura de pantalla en un trial congruente del primer nivel donde se observa la posición respecto de la cruz de fijación y las proporciones de los animales en pantalla.

A diferencia del avioncito, en este juego precisamos que los botones estén juntos y en el centro de la pantalla.

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc. Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio del juego
- Inicio de bloque
- Construcción del bloque
- Inicio de trial
- Detalles de la construcción del trial
- Respuesta del usuario
- Fin del trial/Resultado
- Fin de bloque
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)

Configuraciones

Este juego debe poder configurarse de manera tal de definir la cantidad de niveles que tendrá.

Debe poder definirse para cada nivel/bloque:

- Animal (cuál y qué dibujo -de qué color-)
- Color de fondo
- Tiempos de presentación (del estímulo, de la cruz, de la clave, del feedback, ISI).
- Cantidad de estímulos
- Proporción de estímulos (congruentes/incongruentes)
- Proporción de claves (Sin clave/Central/Doble/Espacial)
- Si el nivel es de prueba, idealmente debería poder definirse un porcentaje/cantidad de ensayos incorrectos que al ser superado repita el nivel de prueba una única vez (ej.: si un niño contesta incorrectamente 6/8, es muy posible que no esté entendiendo las reglas y debería poder repetir una única vez esos 8 ensayos -no tienen necesariamente que ser los mismos-).

Estos niveles deben poder ser determinados por medio de porcentajes o de manera explícita dando la secuencia de estímulos a mostrar (ver ejemplo en Tabla A.2).

Nivel	Descripción
Nivel 00 (prueba)	Pez Fondo celeste 8 trials 50 % congruentes - 50 % incongruentes Claves: Sin clave: 100 %; Central: 0 %; Doble: 0 %; Espacial: 0 %
Nivel 01	Pez Fondo celeste 32 trials 70 % congruentes - 30 % incongruentes Claves: Sin clave: 70 %; Central: 10 %; Doble: 10 %; Espacial: 10 %
Nivel 02	Ratón Fondo rojo 40 trials 30 % congruentes - 70 % incongruentes Claves: Sin clave: 20 %; Central: 40 %; Doble: 10 %; Espacial: 30 %
Nivel 03	Pájaro Fondo verde 28 trials 50 % congruentes - 50 % incongruentes Claves: Sin clave: 10 %; Central: 30 %; Doble: 30 %; Espacial: 30 %

Tab. A.2: Ejemplo de configuración de 4 niveles

A.5.2 Avioncito

Fundamentos

Este juego está basado en tareas de flexibilidad cognitiva y control inhibitorio estilo stroop [88] y busca entrenar la capacidad de inhibir una respuesta dominante para conseguir una respuesta subdominante. Los participantes tienen dos opciones de dirección (derecha o izquierda) y deben decidir, lo más rápidamente posible, cuál es la correcta. En algún lugar visible de la pantalla aparece un avión de papel y el niño debe indicar una dirección. En la versión web, si la dirección elegida es la derecha, debe presionar la tecla L; si es la izquierda, la tecla A. Estas dos teclas fueron elegidas buscando maximizar la distancia espacial real entre las dos elecciones (creemos que utilizar las flechas puede ser más intuitivo en un principio, pero el hecho de que estén tan cerca puede también llevar a una interferencia espacial no deseada, sobre todo en situaciones de incongruencia entre la posición del avión en la pantalla y la dirección a ser seleccionada).

Los aviones pueden ser amarillos o rojos y estar mirando a la derecha o a la izquierda. Si son amarillos, la dirección que hay que elegir es la misma a la que apunta el avión; si el avión es rojo, en cambio, hay que elegir la dirección opuesta. Por ejemplo: si el avión es amarillo y mira hacia la derecha, hay que presionar la tecla L; pero, si el avión es rojo y mira hacia la derecha, hay que presionar la tecla A.

La respuesta del niño debe ser lo más rápida posible para lo cual se muestra en el centro de la pantalla un reloj que indica el tiempo máximo para resolver el trial. Si se supera ese tiempo, el trial se considera perdido y comienza uno nuevo.

Además debe haber otros recursos gráficos como un indicador de avance en el juego para ver cuánto falta para concluirlo, cuántos niveles son, por qué nivel va el niño. Cuanto más avanzado es el nivel más rápido corre el reloj del centro.

También debe haber un indicador el tiempo máximo de juego total.

El juego se termina cuando:

1. se comete un cantidad determinada de errores dentro de un mismo nivel
2. cuando el indicador del tiempo máximo de juego se agota

Niveles

El juego tendrá que poder aceptar distintas estructuras de niveles, pero que siguen una estructura fija. Lo único que varía es la composición de los niveles.

Cada vez que se inicia el juego el niño comienza desde el principio. Es decir, no se guarda avance del mismo.

Para pasar de un nivel al siguiente hay dos maneras, cada nivel tiene una única manera de ser pasado:

- se debe pasar una cantidad fija de aviones a los que no importa cuál es la respuesta (nivel de prueba)
- se debe contestar una cantidad fija de aviones de manera correcta (nivel común)

El niño no pierde en los niveles que son nivel de prueba. Sólo pierde si responde de manera errónea (o omite respuesta) una cantidad determinada veces en el mismo nivel (nivel común).

Un nivel es una sucesión de trials definida. Cada trial se puede ganar o perder, y esto debe ser fácil de visualizar para el niño que está jugando. Se gana se si presiona la tecla correcta, se pierde si se presiona la tecla incorrecta o si se acaba el tiempo disponible para realizar el trial.

Cada nivel consta de:

- Una cantidad de avioncitos fija, en proporciones previamente definidas
  - Color: Amarillo o Rojo
  - Invertido: Sí o No
- Dirección a la que apunta cada avión: Izquierda o Derecha
- Posiciones posibles en las cuales pueden aparecer los avioncitos del nivel
- Distractores: son otras imagenes que pueden o no estar que sólo molestan la vista del niño. Ej. nubes, pájaros, aviones verdes, etc. Se utilizan en los niveles más avanzados. No deben tapar el avioncito objetivo, sino aparecer más como fondo.
- Flancos: imágenes que aparecen cerca del avioncito.

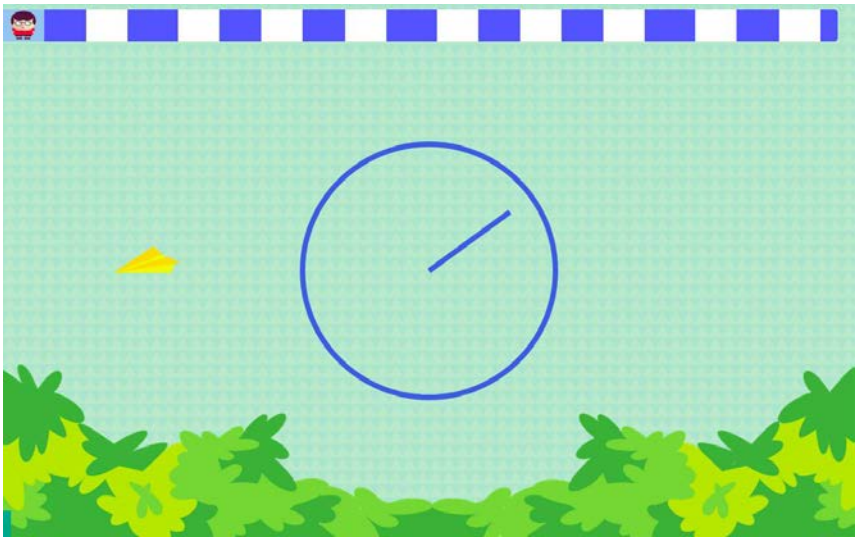


Fig. A.20: Pantalla de avioncito.

Pantalla

La Fig. A.20 muestra un ejemplo de pantalla.

Controles

En la versión en Web (PC con teclado) se utilizaron las teclas A y L. Para versiones táctiles hay que encontrar algún recurso como puede ser botones en el pantalla.

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc.

Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio de trial
- Detalles de la construcción del trial
- Respuesta del usuario
- Fin del trial/Resultado
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)
- Inicio del juego

Configuraciones

Este juego debe poder configurarse de manera tal de definir la cantidad de niveles que tendrá. Definir si tiene o no tiempo máximo de juego, y en particular expresar este tiempo en función de cantidad de avioncitos mostrados.

Se debe poder definir para cada nivel:

- Cantidad de aviones, tipo y posiciones posibles
- Cantidad de distractores, tipo y posiciones posibles (puede o no aparecer)
- Cantidad de flancos, tipo y posiciones posibles (puede o no aparecer)
- Tiempo disponible para responder todos los trials del nivel

Se debe poder definir algunos niveles, y para los demás tener un default de cómo debe ser.

El tiempo de base por trial se debe establecer y luego se pueden usar fracciones de ese tiempo para los niveles.

Estos niveles deben poder ser determinados por medio de porcentajes o de manera explícita dando la secuencia de avioncitos a mostrar (ver ejemplo en Tabla A.3).

Nivel	Descripción
Nivel 01	10 aviones nivel de prueba 100 % amarillos no invertidos (50 % izq - 50 % der)
Nivel 02	30 aviones nivel común 60 % amarillos no invertidos (50 % izq - 50 % der) 40 % rojos no invertidos (90 % izq - 10 % der)
Nivel 03	30 aviones nivel común 30 % amarillos no invertidos (50 % izq - 50 % der) 50 % rojos no invertidos (50 % izq - 50 % der) 20 % amarillos invertidos (0 % izq - 100 % der)
Nivel por defecto	30 aviones nivel común 30 % amarillos no invertidos (50 % izq - 50 % der) 50 % rojos no invertidos (50 % izq - 50 % der) 20 % amarillos invertidos (50 % izq - 50 % der) 0.7 x Tmax
Global	10 niveles, 3 con configuración especial, 7 con configuración por defecto 3 avioncitos mal máximo por nivel Se pierde por tiempo a los 279 aviones Tiempo máximo (Tmax) de trial es 2 segundos

Tab. A.3: Ejemplo de configuración del juego Avioncito

A.5.3 Memomarote

Fundamentos

La memoria de trabajo refiere a la capacidad de almacenar y manipular información por períodos muy cortos de tiempo [137]. Su capacidad es limitada a unos pocos ítems, y aún se discute el número, aunque hay consenso en que se encuentra entre cuatro y siete [22, 35]. El juego de memoria de trabajo se basa en un paradigma que requiere memoria de reconocimiento para patrones visuales pero no para ubicaciones espaciales [86, 102]. Cada ensayo implica varias jugadas y consiste en un tablero y en un cierto número de fichas distribuidas aleatoriamente en él (ver Fig. A.21). Cada ficha es una imagen definida por una serie de atributos. Los atributos pueden estar presentes o ausentes (por ejemplo, una ficha puede, o no, tener baldes) y, cuando presentes, pueden asumir un amplio rango de valores. Los atributos incluyen color de fondo (puede ser un color liso o la imagen de una playa, como en la Fig. A.21), forma de la ficha, sombrilla (en diferentes colores o ausente), cantidad de estrellas, cantidad de baldes, personaje (pueden ser tres en diferentes posiciones, o ninguno), etc. En el ejemplo de la Fig. A.21, el ensayo es de seis fichas, todas con el mismo fondo y forma; todas tienen diferente color de sombrilla y diferente número de estrellas; ninguna tiene baldes. Sólo la mitad tiene al personaje Ana sentado.

Cada ensayo, constituido por varias jugadas, tiene por objetivo seleccionar (cliquear) todas las fichas que aparecen en el tablero. Luego de que el niño selecciona una ficha, las fichas desaparecen y reaparecen, después de un tiempo, configurable pero cercano a los 2 segundos, en una ubicación diferente (siguen siendo las mismas fichas pero están mezcladas). En la siguiente jugada, el niño tiene que seleccionar una ficha no elegida previamente. La dinámica continúa hasta que el niño selecciona, sin repetir, todas las fichas o hasta que comete un error. Por ejemplo, completar correctamente el ensayo de la Fig. A.21 requerirá seis jugadas (seis clics). Cada vez que el niño selecciona una ficha anteriormente elegida, se computa un error y pierde el ensayo. Cada vez que el niño hace una selección completa correcta, gana el ensayo.

Es importante notar que, para números pequeños de fichas, los niños pueden, simplemente, intentar recordar cuáles seleccionaron (o cuáles falta seleccionar). Sin embargo, cuando el número de fichas excede el límite de ítems procesables por la memoria de trabajo, los niños deben adoptar otra estrategia (agrupar, ordenar) para poder resolver el problema [92]. Los niños nos han demostrado que generan muchas más estrategias que las que a nosotros se nos ocurrieron (hay algunos que, por ejemplo, se arman canciones para recordar). En cada ensayo, cualquier atributo puede ser definido como presente o ausente, y dentro de los presentes elegir si varía entre todos los posibles o sólo sobre algunos. Por ejemplo, en cada ficha de la Fig. A.21, el personaje toma el valor “Ana” o “ninguno”. Así, esta variable binaria puede ser potencialmente utilizada por los niños para agrupar y categorizar los ítems. Las estrategias de búsqueda de los niños son registradas por el juego, lo que significa que, en un futuro, el juego podrá identificar automáticamente errores para cada nivel de dificultad, así como estrategias utilizadas por los niños, y actuar en consecuencia. En los resultados aquí presentados, aquellos niños que parecían no generar una estrategia que les permitiera ganar, luego de tres repeticiones del mismo nivel de dificultad sin lograr el objetivo, recibieron orientación de un experimentador para promover la generación de estrategias.

Fig. A.21: Versión esquemática del juego.

Niveles

El juego tendrá que poder aceptar distintas estructuras de niveles, pero que siguen una estructura fija. Lo único que varía es la composición de los niveles.

Se debe poder configurar si cada vez que se inicia el juego el niño debe comenzar desde el principio o desde un punto más avanzado que fuera salvado en una sesión anterior (checkpoint).

Para pasar de un nivel al siguiente se debe pasar una cantidad fija de trials de manera correcta ó de manera consecutivas.

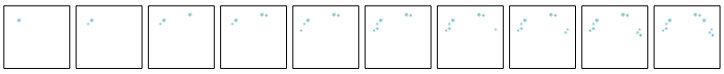
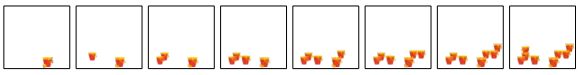
Un nivel es una sucesión de trials definida. Cada trial se puede ganar o perder, y esto debe ser fácil de visualizar para el niño que está jugando. Se gana si el niño presiona con el mouse/dedo una única vez en cada una de las fichas, se pierde si se presiona dos veces la misma ficha. Todos los trials de un nivel tienen la misma cantidad de fichas, las cuales son todas distintas entre sí. Las fichas están construidas como combinación de atributos. Actualmente los atributos que se están utilizando son:

Para cada una de las categorías, ese objeto puede o no estar, excepto las 2 primeras que siempre están. (esta es una guía tentativa de objetos, se agregarán y modificarán las cantidades de dichas categorías).

Cada categoría además tiene definidos clusters que sirven para usar esa característica para separar a los grupos (como la presencia o no del personaje Ana en el ejemplo previo).

Las fichas no son elegidas al azar, sino que a partir de cierto número de fichas se garantiza que se puedan separar en subgrupos, como para resolverlos como problemas más chicos (en el ejemplo anterior, la presencia de Ana en la mitad de las fichas garantiza esta posibilidad).

Actualmente cada nivel tiene 5 etapas que se suceden en orden al resolver correctamente un trial (podrían ser más de uno)

Propiedad	Valor
Posición	es un par (X, Y) con, X∈[133, 263, 393, 523, 653, 783] Y∈[230, 360, 490, 620]
Forma (*shape*)	
Fondo (*back*)	
Sombrillas	
Estrellas	
Baldes	
Personajes	

Tab. A.4: Cararcterísticas originales de las fichas.

1. tiene únicamente fondo y forma (uno de los dos separa las fichas y el otro es al azar)
2. idem 1. con el agregado de una característica constante. Ej. todos con Ana Sentada.
3. idem 1. con el agregado de una característica variable. Ej. algunos con Ana Sentada, otros con Pancho.
4. idem 1. con el agregado de los ítems 2. y 3., es decir, una característica variable y una constante.
5. idem 4. salvo que ambas características son variables.

Actualmente estamos elaborando reglas más complejas sobre como pasar de etapas. y cómo están compuestas las mismas.

Para separar en grupos, se asegura que dada una cantidad se dividan en grupos del mismo tamaño:

- Menos de 3 fichas: 1 grupo
- Entre 4 y 10 fichas: 2 grupos
- Entre 11 y 15 fichas: 3 grupos
- Entre 16 y 24 fichas: 4 grupos
- Más de 25 fichas: 5 grupos

Ejemplos de cluster para la característica fondo:

- Grupo 1: marino, celeste, azul

- Grupo 2: fucsia, violeta, rosa
- Grupo 3: marron, gris
- Grupo 4: amarillo, naranja
- Grupo 5: playa
- Grupo 6: verde

Cada trial tiene 2 x #fichas + 2 cajas (o algún número similar). Estas cajas se ubican de forma aleatoria en un posiciones de una grilla en la que se divide la pantalla. Propuestas previas le daban formas a las posiciones de las cajas, formando figuras, por ejemplo, un puente, una persona, etc. Es una alternativa a analizar si es realizable.

Ya en el juego, las fichas se ubican de manera aleatoria en las cajas (ubicadas al inicio del trial también de manera aleatoria). Cada vez que el niño haga clic en una de las fichas, si es una ficha que ya había presionado, entonces termina el trial dando el feedback correspondiente; si es una ficha que no había seleccionado previamente, entonces se ocultan todas las fichas y vuelven a repartirse al azar en las cajas que se mantienen fijas. En caso de ser el clic con el que se completa la tarea, es decir le hizo clic a todas, se gana el trial y se pasa a un nuevo trial.

Pantalla

La Fig. A.22 muestra un ejemplo de pantalla.

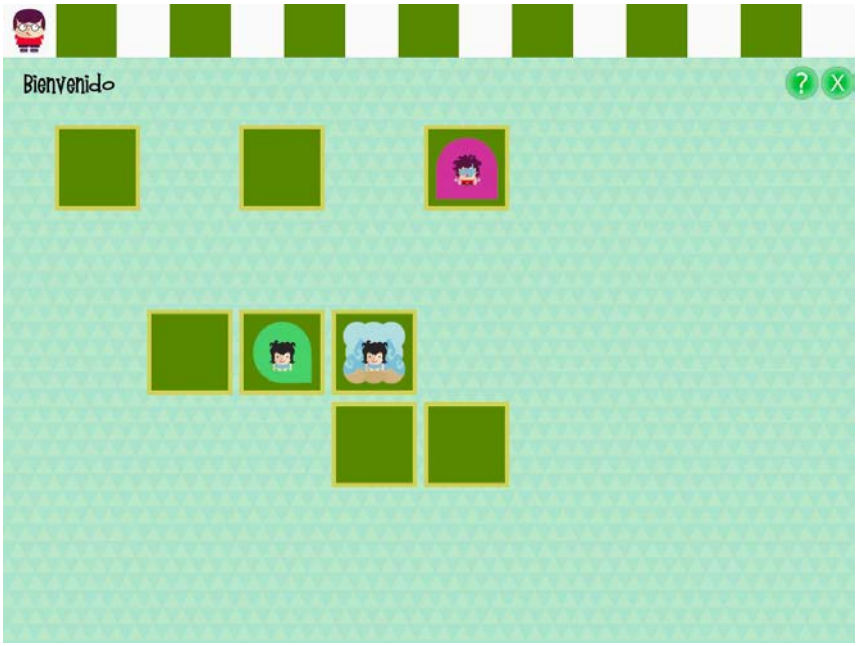


Fig. A.22: Pantalla de Memomarote.

Controles

Para jugar únicamente se utiliza el mouse (o dedo en caso de dispositivo movil).

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc.

Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio del juego
- Inicio de trial
- Detalles de la construcción del trial
- Respuesta del usuario
- Trayectoria del mouse (cuando exista)
- Tiempos de respuesta
- Fin del trial/Resultado
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)

Configuraciones

De momento no hay planeadas configuraciones en cuanto a la jugabilidad y flujo del juego. Sí existirán más categorías de atributos, y más ejemplares dentro de las creadas.

A.5.4 Number Catcher - El atrapanúmeros

Fundamentos

Este juego está diseñado por el Inserm (Institut national de la santé et de la recherche médicale). Actualmente se encuentra implementado en Flash y puede jugarse en: `http://thenumbercatcher.com`.

El juego consiste en llenar distintos vehículos con objetos. La dificultad radica que cada vehículo tiene cierta capacidad y hay que llenarlo hasta completarlo. A medida que avanzan los niveles, hay que cargar cada vez más objetos, y los mismos se presentan de distintas maneras. Además existe una sierra con la cual se pueden cortar un bloque para lograr una cantidad menor.

Elementos del juego

Bloques todos los bloques tiene un tamaño entre 1 y 10. Además se traducen en objetos que varían según el nivel (frutas, flores, peces), los bloques pueden ser de la siguiente forma:

- bloques de ancho N, que muestran N objetos
- bloques de ancho N, que muestran el número N y al ser clickeados se transforman en fruta
- bloques de ancho fijo con el número N o con una operación (Ej. 6+3, 10-4)

los bloques aparecen desde arriba cuando:

- se aprieta el botón de "more boxes"
- se elige un bloque, automáticamente aparece uno nuevo con cierta probabilidad a definir en cada nivel (independienetme del éxito del click, si no entraba también aparece uno nuevo)
- pasa el tiempo, luego de que cierta cantidad de tiempo desde la última aparición

el tamaño y formato de los bloques estará definido para cada nivel. Una opción a contemplar es que una vez iniciado un trial se "emita" un bloque óptimo para dicho ejercicio con cierta probabilidad.

Transporte son los medios de transporte que cargan los objetos, tienen las siguientes características:

- Cantidad de objetos a cargar, estos están organizados en grupos de decenas más un extra que pueden ser menos. Cada vez que se desee completar una decena, por más que no se llene el transporte hay que cortar el bloque para que entre justo, y se carga en 2 partes. El vehículo puede tener una parte o dos (como si fuera un camión con acoplado). Posibles configuraciones:
  - 3
  - 7
  - 2
  - 10 + 2 (un acoplado de 10, y adelante 2)
  - 10 + 9

  - 3x10 + 4 (un acoplado de 10 que tiene 3 niveles que llenar)
- Tiempo de espera máximo, opcional para niveles avanzados

Sierra la misma se utiliza para cortar un bloque cuando es demasiado grande para un lugar, o cuando hace falta dividirlo porque no entra todo entero, ya sea porque va en 2 filas distintas (Ej, hay 8 en el transporte que es de 2x10 + 4 y elijo un 5, entonces debo cortarlo en 2 y 3, para completar la primer fila)

Bomba si se apretan varias veces rápido aparece una bomba que cae como los bloques (un bloque de 1x1 con forma de bomba) que al apretarlo explota y explotan los bloques cercanos a ella impidiendo que se llene la caja y se pierda.

Pájaro El pájaro va a hablar dándole pistas al usuario, presentándole elementos nuevos (sierra, bomba), formas nuevas (camión de 2 partes, ó de más de una fila) y a su vez dando un feedback positivo sobre las elecciones óptimas, presentando alternativas sobre las elecciones sub-óptimas y avisando cuando hay un error.

Tipos de respuesta de al clickear un bloque

Correcto entra perfectamente en el vehículo

Error es demasiado grande para entrar

Sub-óptimo es demasiado pequeño para llenar el vehículo y hay a la vista un bloque que es del tamaño justo.

Puntos

Se dan puntos que se reflejan en una barra de avance, la misma se modifica:

- subiendo por:
  - velocidad al clickear los bloques (bien rápido y correcto)
  - completar una fila del mismo color
  - por precisión
  - por solución óptima
- bajando continua y gradualmente

Fin de un nivel

Un nivel termina cuando:

- La barra de avance se llena, pasando de nivel
- No hay más lugar en la caja de bloques y estos se atascan, perdiendo el nivel
- Se baja la barra de avance hasta que queda vacía (esta barra se va decrementando

Niveles

Hay dos tipos de configuración en principio. En las tablas A.5 y A.6 detallan los niveles con su temas y comentario de qué sucede en ese nivel.

Nivel	Paisaje	Tema	Descripción
1	ciudad	día	Introducción - Nivel muy fácil
2	ciudad	día	
3	ciudad	día	
4	ciudad	noche	Empezar a usar digitos
5	ciudad	noche	
6	campo	día	Presentar sierra
7	campo	día	El vehículo excede el tamaño 10 por primera vez
8	campo	día	Color
9	campo	día	
10	campo	día	Bomba
11	mar	día	
12	mar	día	
13	mar	noche	
14	mar	noche	Ejercicios adicionales
15	mar	noche	Ejercicios adicionales + Tamaño fijo

Tab. A.5: Configuración modo fácil

Nivel	Paisaje	Tema	Descripción
1	ciudad	día	Introducción - Nivel muy fácil
2	ciudad	día	
3	ciudad	día	Empezar a usar digitos
4	ciudad	noche	Sierra
5	ciudad	noche	El vehículo excede el tamaño 10 por primera vez
6	campo	día	Un color de bonus
7	campo	día	Incorporación del límite de tiempo
8	campo	día	El tamaño del vehículo llega a 20
9	campo	día	
10	campo	día	Bomba
11	mar	día	Barco que se hunde; tiempo límite difícil
12	mar	día	Ejercicios adicionales
13	mar	noche	Ejercicios de resta
14	mar	noche	Sumas y restas
15	mar	noche	

Tab. A.6: Configuración modo difícil

Pantalla

La Fig. A.23 muestra un ejemplo de pantalla.

Controles

Para jugar únicamente se utiliza el mouse (o dedo en caso de dispositivo movil).

Fig. A.23: Pantalla de Number Catcher.

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc.

Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio del juego
- Inicio de trial
- Detalles de la construcción del trial
- Respuesta del usuario
- Trayectoria del mouse (cuando exista)
- Tiempos de respuesta
- Tiempo entre eventos (ej. apretó la ayuda, la sacó, apretó un bloque, apretó otro, etc)
- Fin del trial/Resultado
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)
- Estado al salir del juego (éxito, muchos errores, muy lento, se apilaron las cajas y perdió

Configuraciones

De momento no hay planeadas configuraciones en cuanto a la jugabilidad y flujo del juego. Sí existirán más categorías de atributos, y más ejemplares dentro de las creadas.

A.5.5 Stroop

Fundamentos

Esta es una prueba desarrollada por Davidson y colaboradores [25] para evaluar control inhibitorio y flexibilidad cognitiva en niños pequeños. La tarea consiste en pulsar uno de dos botones de acuerdo con la figura que aparezca en pantalla y su ubicación. Hay dos figuras estímulo posibles: un corazón o una flor (Fig. A.24 a y b) y dos posibles posiciones: derecha o izquierda de la pantalla (Fig. A.24 c). Si la figura que aparece es el corazón, hay que apretar el botón del mismo lado de la figura; si, en cambio, aparece la flor, hay que apretar el botón del lado contrario, generando ensayos congruentes (compatibles, en este ejemplo los del corazón) o incongruentes (incompatibles, en este ejemplo los de la flor). La tarea proviene de dos paradigmas clásicos de la psicología cognitiva: la tarea Simon [84, 120] y la alternancia de tareas (el task switching [88]). En el paradigma de Simon, un aspecto no espacial del estímulo (como su identidad; en este caso, flor o corazón) es relevante, mientras que su ubicación espacial es menos relevante. El llamado "efecto Simon", o de incompatibilidad espacial, refiere a que las respuestas son más rápidas y certeras en ensayos espaciales compatibles que en incompatibles —es decir, más éxito cuando el estímulo y la respuesta se encuentran en el mismo lado que cuando están en lados opuestos [84]—. Entre otras cosas, esto demuestra una tendencia fuerte a responder por un atributo casi irrelevante (en este caso, la posición espacial) que se inhibe cuando la ubicación del estímulo y la respuesta correcta son incompatibles, permitiendo así estudiar algunos aspectos del control inhibitorio. Esta tarea, además, aumenta los requerimientos de memoria de trabajo y flexibilidad, porque la respuesta correcta requiere cierta manipulación mental: en lugar de la regla "para A, presionar la derecha", la regla es "para A, presionar el lado opuesto/mismo que A". Por eso, además de utilizar ciertas reglas asociadas con los dos estímulos (como es el requerimiento mnésico clásico para las tareas de Simon), en este test los participantes deben elegir y utilizar en cada uno de los ensayos la regla apropiada para el estímulo.

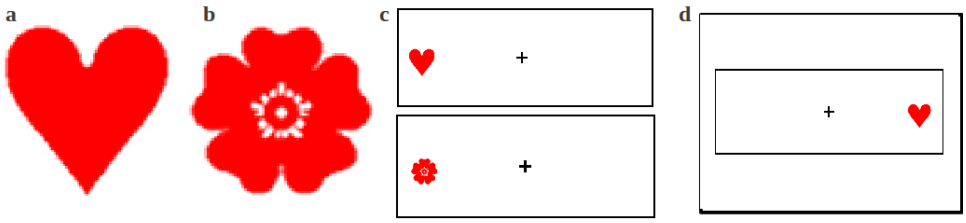


Fig. A.24: Estímulos que conforman la tarea de Stroop Corazón-Flor. a) Corazón, estímulo congruente. b) Flor, estímulo incongruente. c) Posición de los estímulos con respecto a la cruz de fijación. d) Captura de pantalla en un ensayo congruente, donde se observan las proporciones del estímulo completo en pantalla.

Niveles

Hay tres tipos de niveles a través de las cuales cambian las condiciones de presentación del tipo de estímulos, de manera de ir aumentando progresivamente la dificultad de las demandas de control inhibitorio y flexibilidad. Entre niveles los niños pueden descansar todo el tiempo que precisen/quieran.

Tipo 1 - congruente una cantidad de trials a definir en los que sólo aparece la figura congruente (en este ejemplo, corazón, Fig. A.24, a y c superior) y la consigna es presionar el botón del mismo lado en que aparece esta figura.

Tipo 2: incongruente una cantidad de trials a definir en los que sólo aparece la figura incongruente (en este ejemplo, la flor, Fig. A.24, b y c inferior) y la consigna es presionar el botón del lado opuesto al que aparece esta figura.

Tipo 3: Mixta Es la condición en la que se combinan estímulos congruentes e incongruentes. Consiste en una cantidad de trials a definir en los que puede aparecer la figura de la flor o la del corazón y la consigna es presionar el botón del lado opuesto —si aparece la flor— o el del mismo lado —si aparece el corazón—.

Además, antes de comenzar el primer nivel los niños tienen una serie de trials de prueba que pueden ser definidos como un nivel diferente. Son exactamente iguales que el resto de los trials de ese nivel (congruentes, incongruentes o mixtos) excepto que los niños no tienen tiempo máximo para responder (o el tiempo es mucho). En los trials de prueba se da feedback al niño. En los niveles propiamente dichos, no se da feedback.

Para cada nivel se definirá: cantidad de trials y distribución de estos (ej, 10 flores y 5 corazones, y las posiciones), además se definirá si el estímulo se muestra hasta que exista una respuesta (trial de prueba) o si se muestra hasta un tiempo máximo (ej. 2500 ms). Por otro lado, es necesario poder definir cuánto tiempo transcurrirá entre trial y trial (ej. 1000 ms) y cuánto tiempo transcurrirá entre que comienza un trial (aparece la cruz de fijación) y que aparece el estímulo (ej. 250 ms).

Pantalla

No tenemos ejemplo de esta pantalla porque este es un juego nuevo. Ver pantalla ejemplo (por proporciones y posiciones) en la Fig. A.24d. La idea es usar un fondo fijo sobre el que aparezca la cruz central de fijación y los estímulos (corazón y flor en el ejemplo). El recuadro que se observa en la Fig. A.24 no hace falta, nosotros no lo tendremos.

Controles

Hasta el momento no hemos jugado a este juego de forma no supervisada y en el juego presencial los niños usaban los botones del mouse porque un adulto lo posiciona correctamente. Como no será posible controlar la posición del mouse para los niños que jueguen remoto, proponemos que en la versión en Web (PC con teclado) las teclas para indicar derecha o izquierda sean las flechitas del teclado correspondientes o dos teclas centrales y adyacentes del teclado, como v y b ó b y n. Idealmente no deberían ser las mismas que las utilizadas en ANT.

Para versiones táctiles hay que encontrar algún recurso como puede ser botones en el pantalla. A diferencia del avioncito, en este juego precisamos que los botones estén juntos y en el centro de la pantalla.

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc.

Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio del juego
- Inicio de trial
- Inicio del nivel
- Detalles de la construcción del trial
- Respuesta del usuario
- Fin del trial/Resultado
- Fin del nivel
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)

Configuraciones

Este juego debe poder configurarse de manera tal de definir la cantidad y tipo de niveles que tendrá. Se debe poder definir para un juego completo cuál será el par de estímulos congruente-incongruente a utilizar (en el ejemplo visto es un corazón o una flor; nosotros propondremos otros pares de estímulos).

Se debe poder definir para cada nivel:

- Cantidad y tipo de estímulos y sus posiciones,
- Tiempos para esperar respuesta (fijo o sin límite -en las pruebas-),
- Tiempo entre estímulos (TISI),
- Tiempo hasta que aparezca el estímulo.

Estos niveles deben poder ser determinados por medio de porcentajes o de manera explícita dando la secuencia de estímulos a mostrar (ver ejemplo en Tabla A.7).

Nivel	Descripción
Nivel 01	4 estimulos 100 % corazones 50 % izq - 50 % der Sin tiempo límite para respuesta TISI 500ms
Nivel 02	30 estimulos 100 % corazones 40 % izq - 60 % der Ttrial 2500ms TISI 500ms
Nivel 03	4 estimulos 100 % flores 50 % izq - 50 % der Sin tiempo límite para respuesta TISI 500ms
Nivel 04	30 estimulos 100 % flores 40 % izquierda - 60 % derecha Ttrial 2500ms TISI 500ms
Nivel 05	8 estimulos 50 % corazones - 50 % flores 50 % izquierda - 50 % derecha Ttrial 2000ms TISI 1000ms
Nivel 06	30 estimulos 60 % corazones - 40 % flores 50 % izquierda - 50 % derecha Ttrial 2000ms TISI 1000ms

Tab. A.7: Ejemplo de configuración

A.5.6 Torres de Londres

Fundamentos

La prueba Torre de Londres (ToL, por Tower of London) fue desarrollada por Shallice [117] como una variante facilitada de la prueba Torre de Hanói. Evalúa procesos de planificación de acciones y representaciones, y de memoria de trabajo espacial, que contribuyen a la conformación de una estrategia para el logro de un plan. Este es el único test no computarizado que hemos tomado. Se utilizan dos aparatos iguales, semejantes a los propuestos por Shallice consistentes en una base de madera con tres varillas de alturas crecientes (ver esquema ejemplo en la Fig. A.25). Cada aparato contiene tres bolitas, cada una de un color (rojo, azul o amarillo). Uno de los aparatos se usa para disponer los diseños correspondientes al modelo final (que queda siempre en poder del experimentador) y el otro, para el inicial, sobre el cual trabaja el niño.

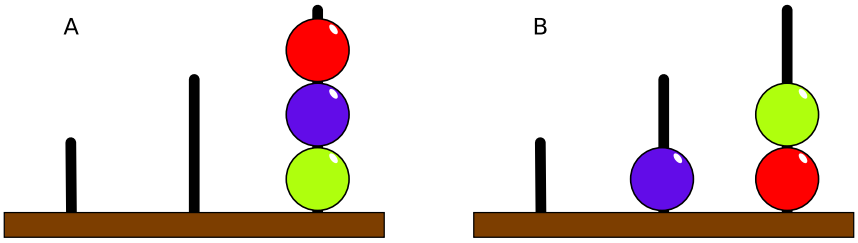


Fig. A.25: Ejemplos de un ensayo de ToL. A: configuración inicial (la que recibe el niño), B: configuración final (la que muestra el experimentador). Este es un ensayo que puede resolverse en un mínimo de cinco movimientos.

El objetivo de la tarea es alcanzar una configuración final de las bolitas en las varillas a partir de otra inicial, para lo cual el niño deberá mover de a una bolita por vez, y hacerlo en una cantidad mínima de movimientos explicitada al comienzo de cada ensayo ("distancia al objetivo"). En la varilla más alta entran tres bolitas, en la del medio, dos y en la más pequeña, sólo una.

Un ensayo correcto es aquel en el que el niño logra alcanzar la configuración del modelo final realizando la cantidad ideal de movimientos mínimos. La versión que implementaremos es una modificación sobre la original. Puede haber más de tres bolitas y, en lugar de existir 3 varillas de largo fijo, estas pueden tener distinto largo (dependiendo de la cantidad de bolitas)- Dado que es un juego que no es sencillo de entender, es particularmente deseable un tutorial interactivo que muestre cómo jugar.

Niveles

Los niveles estarán definidos como una sucesión de trials donde cada uno tendrá alguna cantidad de movimientos mínima para resolverlo.

Se definen 2 modos de juego:

- exploración libre: donde el niño puede realizar tantos movimientos como desee. En este modo, puede configurarse si se da feedback o no y cómo se consideran los ensayos cuando la cantidad de movimientos es mayor a la mínima (puede ser que se le permita jugar hasta conseguir el objetivo pero que se considere incorrecto el trial al momento de computo).

- movimientos restringidos: donde se le informa cuál es la cantidad mínima de movimientos y si se pasa, pierde. En este modo no se da feedback.

Importante: el juego no termina cuando los dos tableros quedan iguales sino cuando el niño lo decide. Debería haber algo (por ejemplo un botón) donde el niño pueda indicar que ya terminó ese trial. Independientemente de cómo le haya ido, cuando el niño indique que terminó un trial, se continúa el juego según la configuración del admin y se computa el trial correcto/incorrecto según como se haya definido.

Cada nivel tendrá una estructura básica determinada por:

- El modo de juego,

- La altura de cada una de las 3 varillas,

- La cantidad de bolitas (cada una tendrá un color distinto),

- La cantidad de trials que tendrá ese nivel.
- La forma de pasar de nivel (ej.: x trials correctos consecutivos, o z trials correctos cada w trials de esa cantidad de bolitas)
- La forma de perder (ej.: y trials incorrectos consecutivos)

Cada nivel, a su vez estará determinado por dos tableros, uno con la configuración inicial y otro con la final. Debe estar calculada la distancia mínima para poder resolverlo. Si una bolita es sacada completamente de su palito, se cuenta un movimiento independientemente de si va a otro palito o vuelve al mismo. Las únicas posiciones posibles son dentro de los palitos y arriba de la última bolita o del piso. Si se intenta depositar una bolita fuera de los palitos o sobre otra bolita, esta vuelve a su posición previa y no se computa movimiento.

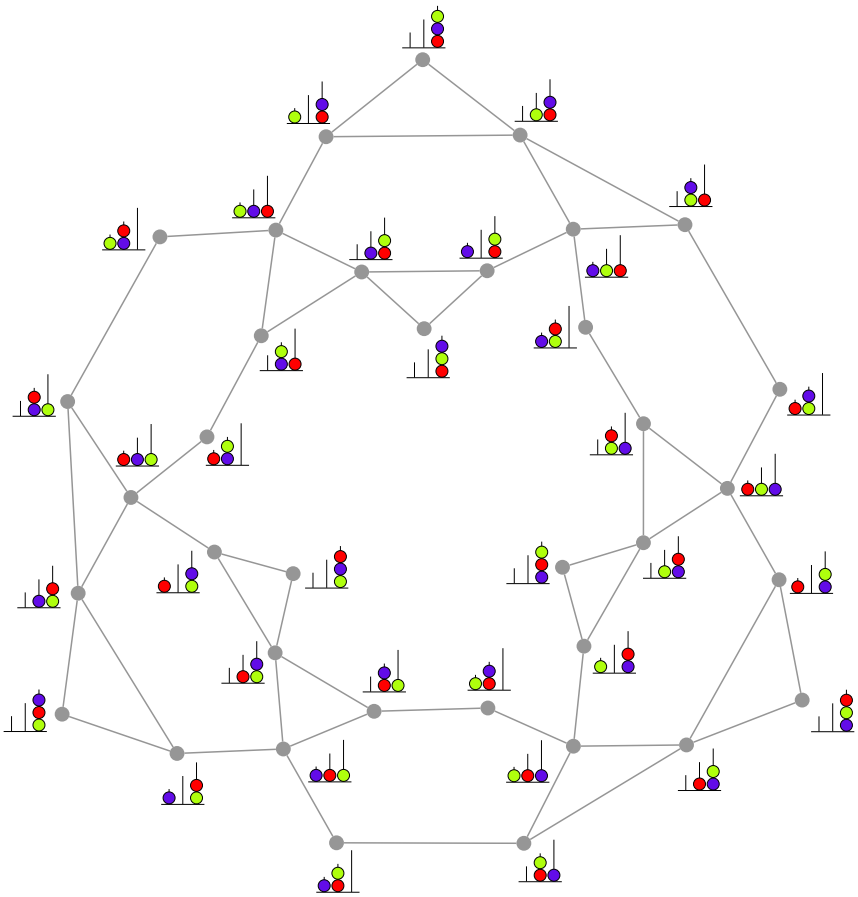


Fig. A.26: Espacio de problema de ToL clásico (3 bolitas). Cada trial se corresponde con un par de nodos de este gráfico. La dificultad de un trial se define como la cantidad de ejes mínima que se debe transitar para ir de uno a otro.

Pantalla

La pantalla deberá ser estructuralmente parecida a la Fig. A.27. Cambiando el fondo por el de otro juego hasta que se definan bien los recursos gráficos de este juego. Cabe aclarar que el tablero superior es una imagen con la que no se puede interactuar.

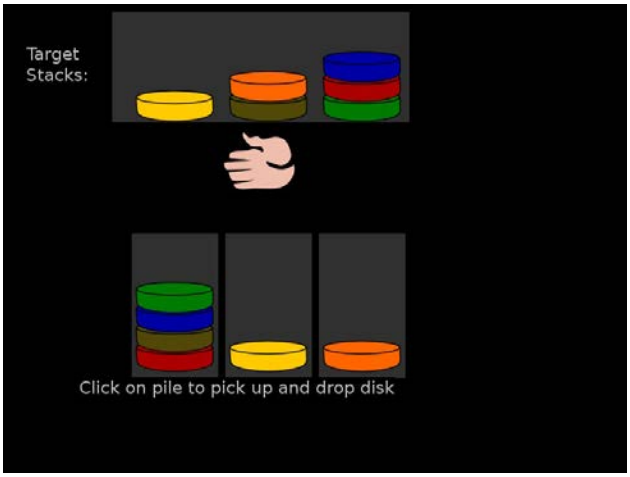


Fig. A.27: Pantalla de un juego de TOL con varios discos (equivalentes a bolitas) y sin palitos. En este caso, el tablero superior es la configuración final y el inferior, la del usuario.

Controles

Para jugar, únicamente se utiliza el mouse (o dedo en caso de dispositivo móvil). Hay que mover bolitas (drag and drop) de un palito a otro (o al mismo).

Datos a registrar

Para todos los trials se deben registrar todos los datos que permitan reconstruirlo. Es decir, los elementos mostrados, la posición de estos eventos, respuesta del usuario, tiempos del trial, etc.

Cada dato registro debe estar acompañado por un timestamp lo más preciso posible, un formato posible es: "YYYY-MM-DD_HH:mm:ss.ms".

Eventos a registrar:

- Inicio del juego
- Inicio de trial
- Detalles de la construcción del trial
- Respuesta/movimientos del usuario
- Fin del trial/Resultado
- Avance del juego (nivel, cantidad de correctos, cantidad de incorrectos)

Configuraciones

Este juego debe poder configurarse de manera tal de definir la cantidad de niveles que tendrá y la forma de avance/finalización.

Se debe poder definir, para cada nivel:

- Modo de juego
- Tamaño de las varillas

- Imagen de las varillas
- Colores de las bolitas
- Cantidad de bolitas
- Cantidad de trials y cantidad de movimientos mínima

En la Tabla A.8 se puede observar un ejemplo de configuración de 3 niveles.

Nivel	Descripción
Nivel 01	modo exploración libre tamaño varillas: [1,2,3] cantidad de pelotas: 3 10 trials = 2 de distancia 2, 3 de distancia 4, 5 de distancia 7
Nivel 02	modo movimientos restringidos tamaño varillas: [3,3,3] cantidad de pelotas: 5 30 trials = 2 de distancia 4, 15 de distancia 4, 13 de distancia 7
Nivel 03	modo movimientos restringidos sin pistas tamaño varillas: [2,3,4] cantidad de pelotas: 6 30 trials = 2 de distancia 4, 15 de distancia 4, 13 de distancia 7

Tab. A.8: Ejemplo de configuración

Actualización 2016

Se agregan para la configuración de niveles de Mate Marote la configuración establecida por la Fig. A.28 y A.29. Ejemplos de su uso se pueden ver en la Fig. A.30. La presente notación y visualización se usará como modelo para la documentación de los demás juegos.

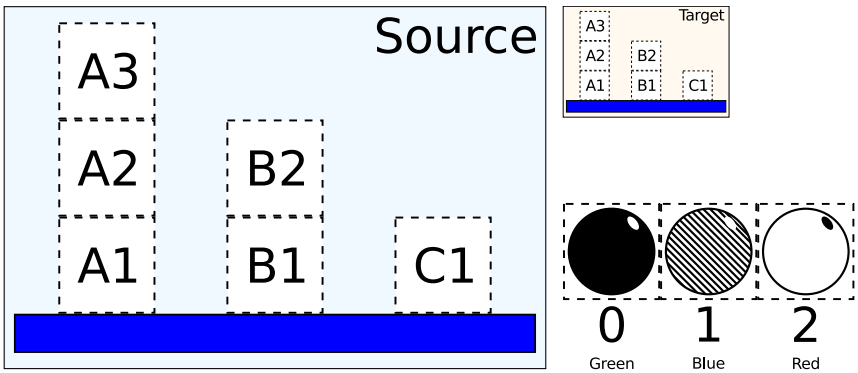


Fig. A.28: Identificación de los elementos que component un ensayo del juego Torres de Londres.

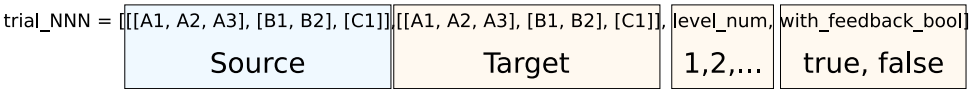
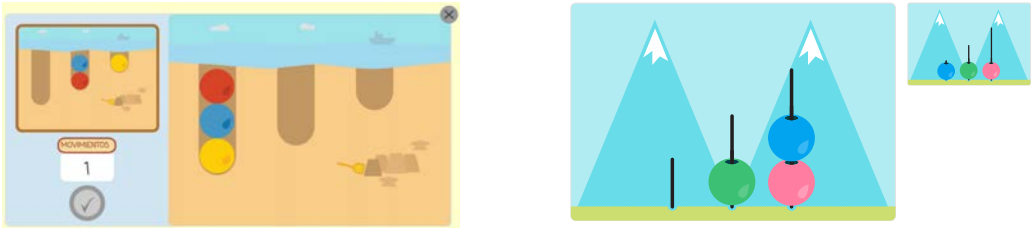


Fig. A.29: Escritura de la representación de un ensayo.



trial_001=[[[0,1,2],[],[]],[[],[2,1],[0]],1, true] trial_002=[[[2,1],[0],[]],[[2],[0],[1]],1, true]

Fig. A.30: Ejemplos de la interpretación de dos ensayos.

A.6 Mate Marote initial backlog

Las tareas serán divididas en grupos en una jerarquía de 3 niveles:

1. Áreas: división inicial del proyecto en las principales ramas de trabajo: servidor, API juegos, plataforma web, plataforma offline (Android/OLPC/iOS), juegos, otros.
2. Tema: descripción de alto nivel de las características a desarrollar/investigar incluidas en el Área
3. Tareas (Feature/Issues): elementos a implementar

Backlog Inicial: Se enumeran elementos del backlog de las primeros dos niveles de jerarquía

Servidor

- ABM Sitio
- ABM Idioma
- ABM Juegos para todas las plataformas
- ABM Gameflow
- ABM Otras entidades (por ej, usuarios)
- Generador de paquetes de instalación para plataformas offline.
- Generador de paquetes de updates para plataformas offline.
- Scripts para manejo masivo (ej. alta masiva de usuarios desde CSV)

API Juegos

- API Log
- API estado/avance de juegos
- API Idioma
- API Sonidos

- API Dibujos
- API Multiplayer
- API Reglas de juegos (parametrización de los juegos)
- API Highscores
- API Badges/Premios
- Diseño general de elementos para los juegos (avance de nivel, tiempo transcurrido, puntaje, etc)

Plataforma Web

- Multisitio
- Multiidioma
- Usuarios con perfil (avatar, redes sociales, premios/medallas, opciones default)
- Diseño gráfico de gameflow
- Páginas estáticas/institucionales
- Usuario invitado
- Migración de juegos
- Versión tablet, versión mobile

Plataforma Offline (Android/OLPC/iOS)

- Targets: OLPC-XO, OLPC-Android, Android, iOS, BlackBerry
- ABM Usuarios
- Gameflow
- Upload automático de logs
- Update automático de games
- Update de plataforma
- Sincronización de datos globales (ej. highscores)
- Multiplayer según plataforma (con/sin conexión internet)

Juegos

- Avioncito
- Memomarote
- ANT
- Stroop
- TOL
- Bloques de Corsi
- Numerosidad
- Linea numérica
- Chocolate Fix (colaboración con ThinkFun y Dra. Silvia Bunge)
- Numbercatcher (colaboración con Dr. Stan Dehaene)

- Lectoescritura (colaboración con Dra. Beatriz Diuk)
- Lenguaje (Colaboración con Dr. Juan Valle Lisboa, Dra. Kathy Hirsh-Pasek, Dr. Jacques Mehler)
- Spatial reasoning (colaboraación con Dra. Nora Newcombe)
- Test de memoria de trabajo no espacial
- KBit
- Geometria (colaboración con Dra. Elizabeth Spelke)
- Metacognición (colaboración con Dra. Cecilia Calero)
- Go-no go
- Number Race (colaboración con Dr. Stan Dehaene)
- Casitas
- Test de sumas/restas

Otros

- Jerarquía de gameflows
- Estrategias de gaming para adicción

BIBLIOGRAPHY

[1] V.A. Aleven and K.R. Koedinger. An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26(2):147–179, 2002.

[2] Vincent Aleven, Bruce M. McLaren, Jonathan Sewall, and Kenneth R. Koedinger. The Cognitive Tutor Authoring Tools (CTAT): Preliminary evaluation of efficiency gains. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4053 LNCS:61–70, 2006.

[3] Geoff Ward & Alan Allport. Planning and Problem solving Using the Five disc Tower of London Task. *The Quarterly Journal of Experimental Psychology Section A*, 50(1):49–78, 1997.

[4] John R Anderson, Frederick G Conrad, and Albert T Corbett. Skill acquisition and the lisp tutor. *Cognitive Science*, 13(4):467–505, 1989.

[5] J.R. Anderson, A.T. Corbett, K.R. Koedinger, and Ray. Pelletier. Cognitive tutors: Lessons learned. *The journal of the learning sciences*, 4(2):167–207, 1995.

[6] Mark A Attiah and Martha J Farah. Minds, motherboards, and money: futurism and realism in the neuroethics of bci technologies. *Frontiers in systems neuroscience*, 8, 2014.

[7] The World Bank. Internet users (per 100 people), 2017.

[8] Pieter J Beers, Henny PA Boshuizen, Paul A Kirschner, Wim Gijselaers, and Jochem Westendorp. Cognitive load measurements and stimulated recall interviews for studying the effects of information and communications technology. *Educational Technology Research and Development*, 56(3):309–328, 2008.

[9] Maryjane Blasi, Sally C Hurwitz, and Sally C Hurwitz. For parents particularly: To be successful—let them play! *Childhood Education*, 79(2):101–102, 2002.

[10] Rafal Bogacz, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, 113(4):700, 2006.

[11] E.M. Brannon. The development of ordinal numerical knowledge in infancy. *Cognition*, 83(3):223–240, 2002.

[12] Juan C Burguillo. Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education*, 55(2):566–575, 2010.

[13] Jerome R Busemeyer and James T Townsend. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, 100(3):432, 1993.

[14] Eric Butler, Erik Andersen, Adam M Smith, Sumit Gulwani, and Zoran Popovic. Automatic Game Progression Design through Analysis of Solution Features. *Chi'15*, pages 2407–2416, 2015.

[15] Alejandro Calderón, Mercedes Ruiz, and Elena Orta. Integrating serious games as learning resources in a software project management course: the case of prodec. In *Proceedings of the 1st International Workshop on Software Engineering Curricula for Millennials*, pages 21–27. IEEE Press, 2017.

[16] Julián Candia, Marta C González, Pu Wang, Timothy Schoenharl, Greg Madey, and Albert-László Barabási. Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical*, 41(22):224015, 2008.

[17] Stephanie M. Carlson, Louis J. Moses, and Laura J. Claxton. Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, 87(4):299–319, 2004.

[18] Hao Cen, Kenneth R Koedinger, and Brian Junker. Is over practice necessary? Improving learning efficiency with the Cognitive Tutor through educational data mining. *Proceedings of the 13th International Conference on Artificial Intelligence in Education AIED 2007*, 158:511–518, 2007.

[19] Gregory KWK Chung and Eva L Baker. An exploratory study to examine the feasibility of measuring problem-solving processes using a click-through interface. *The Journal of Technology, Learning and Assessment*, 2(2), 2003.

[20] Pamela M Cole, Jason J Bendezú, Nilam Ram, and Sy-Miin Chow. Dynamical systems modeling of early childhood self-regulation. *Emotion*, 17(4):684, 2017.

[21] Albert T Corbett and John R Anderson. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction*, 4(4):253–278, 1994.

[22] Nelson Cowan. The magical mystery four: How is working memory capacity limited, and why? *Current directions in psychological science*, 19(1):51–57, 2010.

[23] Eveline A Crone, Carter Wendelken, Linda Van Leijenhorst, Ryan D Honomichl, Kalina Christoff, and Silvia A Bunge. Neurocognitive development of relational reasoning. *Developmental science*, 12(1):55–66, 2009.

[24] Gergely Csibra and György Gergely. Natural pedagogy. *Trends in cognitive sciences*, 13(4):148–153, 2009.

[25] Matthew C Davidson, Dima Amso, Loren Cruess Anderson, and Adele Diamond. Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11):2037–2078, 2006.

[26] Rudolf Debelak, Johanna Egle, Lena Köstering, and Christoph P Kaller. Assessment of planning ability: Psychometric analyses on the unidimensionality and construct validity of the Tower of London task (TOL-F). *Neuropsychology*, 30(3):346–360, 2015.

[27] S. Dehaene. Précis of the number sense. *Mind & Language*, 16(1):16–36, 2001.

[28] S. Dehaene. *The number sense: How the mind creates mathematics*. Oxford Univ Pr, 2011.

[29] S. Dehaene and R. Akhavein. Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2):314, 1995.

[30] S. Dehaene, V. Izard, E. Spelke, and P. Pica. Log or linear? distinct intuitions of the number scale in western and amazonian indigene cultures. *science*, 320(5880):1217–1220, 2008.

[31] J Dunlosky and J Metcalfe. Metacognition: A textbook of cognition, educational, life span, and applied psychology, 2009.

[32] Nathan Eagle, Alex Sandy Pentland, and David Lazer. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences*, 106(36):15274–15278, 2009.

[33] Martin Ebner and Andreas Holzinger. Successful implementation of user-centered game based learning in higher education: An example from civil engineering. *Computers & education*, 49(3):873–890, 2007.

[34] Maria K Eckstein, Belén Guerra-Carrillo, Alison T Miller Singley, and Silvia A Bunge. Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental cognitive neuroscience*, 2016.

[35] Lisa Feigenson and Susan Carey. On the limits of infants' quantification of small object arrays. *Cognition*, 97(3):295–313, 2005.

[36] Lisa Feigenson, Stanislas Dehaene, and Elizabeth Spelke. Core systems of number. *Trends in cognitive sciences*, 8(7):307–314, 2004.

[37] Emilio Ferrer, Elizabeth D O'Hare, and Silvia A Bunge. Fluid reasoning and the developing brain. *Frontiers in neuroscience*, 3:3, 2009.

[38] Leon Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1962.

[39] John H Flavell. Metacognitive aspects of problem solving. *The nature of intelligence*, pages 231–235, 1976.

[40] Nancy Garon, Susan E Bryson, and Isabel M Smith. Executive function in preschoolers: a review using an integrative framework. *Psychological bulletin*, 134(1):31, 2008.

[41] Hector Geffner and Blai Bonet. A concise introduction to models and methods for automated planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8(1):1–141, 2013.

[42] Joshua I Gold and Michael N Shadlen. The neural basis of decision making. *Annu. Rev. Neurosci.*, 30:535–574, 2007.

[43] Andrea Paula Goldin, María Julia Hermida, Diego E Shalom, Martín Elias Costa, Matías Lopez-Rosenfeld, María Soledad Segretin, Diego Fernández-Slezak, Sebastián J Lipina, and Mariano Sigman. Far transfer to language and math of a short software-based gaming intervention. *Proceedings of the National Academy of Sciences*, 111(17):6443–6448, 2014.

[44] Susan Goldin-Meadow and Susan M Wagner. How our hands help us learn. *Trends in cognitive sciences*, 9(5):234–241, 2005.

[45] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

[46] U. Goswami. Neuroscience and education: from research to practice? *Nature Reviews Neuroscience*, 7(5):406–413, 2006.

[47] C Shawn Green and Daphne Bavelier. Action video game modifies visual selective attention. *Nature*, 423(6939):534–537, 2003.

[48] Mark Griffiths. Violent video games and aggression: A review of the literature. *Aggression and violent behavior*, 4(2):203–212, 1999.

[49] Thomas Hainey, Thomas M Connolly, Elizabeth A Boyle, Amanda Wilson, and Aisya Razak. A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, 102:202–223, 2016.

[50] Sherif Halawa, Daniel Greene, and John Mitchell. Dropout prediction in moocs using learner activity features. *Experiences and best practices in and around MOOCs*, 7:3–12, 2014.

[51] Brenna Hassinger-Das, Tamara S. Toub, Jennifer M. Zosh, Jessica Michnick, Roberta Golinkoff, and Kathy Hirsh-Pasek. More than just fun: a place for games in playful learning / Más que diversión: el lugar de los juegos reglados en el aprendizaje lúdico. *Infancia y Aprendizaje*, 40(2):191–218, 2017.

[52] Kathy Hirsh-Pasek, Jennifer M Zosh, Roberta Michnick Golinkoff, James H Gray, Michael B Robb, and Jordy Kaufman. Putting education in "educational" apps: lessons from the science of learning. *Psychological Science in the Public Interest*, 16(1):3–34, 2015.

[53] T L Hodgson, a Bajwa, a M Owen, and C Kennard. The strategic control of gaze direction in the Tower-of-London task. *Journal of cognitive neuroscience*, 12(5):894–907, 2000.

[54] Kate S Hone and Ghada R El Said. Exploring the factors affecting mooc retention: A survey study. *Computers & Education*, 98:157–168, 2016.

[55] Celia Hoyles and Richard Noss. Synthesizing mathematical conceptions and their formalization through the construction of a logo-based school mathematics curriculum. *International Journal of Mathematical Education in Science and Technology*, 18(4):581–595, 1987.

[56] Ira E Hyman, S Matthew Boss, Breanne M Wise, Kira E McKenzie, and Jenna M Caggiano. Did you see the unicycling clown? inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, 24(5):597–607, 2010.

[57] World Telecommunication/ICT Development Report International Telecommunication Union and database. Mobile cellular subscriptions (per 100 people), 2017.

[58] International Telecommunication Union (ITU). *Measuring the information society report: 2014*, volume 8. -, 2014.

[59] Shafika Isaacs. *Transforming education: The power of ICT policies*. Unesco, 2011.

[60] V. Izard, P. Pica, E.S. Spelke, and S. Dehaene. Flexible intuitions of euclidean geometry in an amazonian indigene group. *Proceedings of the National Academy of Sciences*, 108(24):9782, 2011.

[61] Zhi-Qiang Jiang, Wen-Jie Xie, Ming-Xia Li, Boris Podobnik, Wei-Xing Zhou, and H Eugene Stanley. Calling patterns in human communication dynamics. *Proceedings of the National Academy of Sciences*, 110(5):1600–1605, 2013.

[62] Daniel Kahneman, Barbara L Fredrickson, Charles A Schreiber, and Donald A Redelmeier. When more pain is preferred to less: Adding a better end. *Psychological science*, 4(6):401–405, 1993.

[63] Christoph P. Kaller, Josef M. Unterrainer, Benjamin Rahm, and Ulrike Halsband. The impact of problem structure on planning: Insights from the Tower of London task. *Cognitive Brain Research*, 20(3):462–472, 2004.

[64] Márton Karsai, Kimmo Kaski, and János Kertész. Correlated dynamics in egocentric communication networks. *Plos one*, 7(7):e40612, 2012.

[65] W Keith Berg and Dana Byrd. The Tower of London spatial problem-solving task: enhancing clinical and research implementation. *Journal of clinical and experimental neuropsychology*, 24(5):586–604, 2002.

[66] Torkel Klingberg, E Fernell, and PJ Olesen. Computerized training of working memory in children with adhd. *Year Book of Psychiatry & Applied Mental Health*, 2006:42–43, 2006.

[67] K.R. Koedinger, J.R. Anderson, W.H. Hadley, and M.A. Mark. Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education (IJAIED)*, 8:30–43, 1997.

[68] Gabriela V Koppenol-Gonzalez, Samantha Bouwmeester, and a Marije Boonstra. Understanding planning ability measured by the Tower of London: an evaluation of its internal structure by latent variable modeling. *Psychological assessment*, 22(4):923–934, 2010.

[69] Celina Graciela Korzeniowski, Mirta Susana Ison, and Hilda Difabio. Cognitive stimulation in children at social risk: its transference to school performance. *International Journal of Psychological Research*, 10(2):34, 2017.

[70] Lauri Kovanen, Kimmo Kaski, János Kertész, and Jari Saramäki. Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences. *Proceedings of the National Academy of Sciences*, 110(45):18070–18075, 2013.

[71] Jussi M Kumpula, Jukka-Pekka Onnela, Jari Saramäki, Kimmo Kaski, and János Kertész. Emergence of communities in weighted networks. *Physical review letters*, 99(22):228701, 2007.

[72] Donald Richard John Laming. Information theory of choice-reaction times. *London*, 1968.

[73] Ming-Xia Li, Zhi-Qiang Jiang, Wen-Jie Xie, Salvatore Miccichè, Michele Tumminello, Wei-Xing Zhou, and Rosario N Mantegna. A comparative analysis of the statistical properties of large mobile phone calling networks. *Scientific reports*, 4, 2014.

[74] Steve W Link. The relative judgment theory of two choice response time. *Journal of Mathematical Psychology*, 12(1):114–135, 1975.

[75] Stephen W Littlejohn and Karen A Foss. *Theories of human communication*. Waveland Press, 2010.

[76] Tsung-Yu Liu and Yu-Ling Chu. Using ubiquitous games in an english listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, 55(2):630–643, 2010.

[77] Tharindu Rekha Liyanagunawardena, Andrew Alexandar Adams, and Shirley Ann Williams. Moocs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*, 14(3):202–227, 2013.

[78] Kathrin Lockl and Wolfgang Schneider. Precursors of metamemory in young children: The role of theory of mind and metacognitive vocabulary. *Metacognition and Learning*, 1(1):15–31, 2006.

[79] Gordon D Logan, Russell J Schachar, and Rosemary Tannock. Impulsivity and Inhibitory Control. *Psychological Science*, 8(1):60–64, 1997.

[80] Derek Lomas, Dixie Ching, Eliane Stampfer, Melanie Sandoval, and Ken Koedinger. "battleship numberline": A digital game for improving estimation accuracy on fraction number lines. *Society for Research on Educational Effectiveness*, 2011.

[81] Matías Lopez-Rosenfeld, F. Carrillo, G. Garbulsky, D.F. Slezak, and M. Sigman. Quantitative pedagogy: A digital two player game to examine communicative competence. *PLoS ONE*, 10(11), 2015.

[82] Matías Lopez-Rosenfeld, Andrea Paula Goldin, Sebastián Lipina, Mariano Sigman, and Diego Fernandez Slezak. Mate marote: A flexible automated framework for large-scale educational interventions. *Computers & Education*, 68:307–313, 2013.

[83] Matías Lopez-Rosenfeld, CeciliaI. Calero, Diego Fernandez Slezak, Gerry Garbulsky, Mariano Bergman, Marcos Trevisan, and Mariano Sigman. Neglect in human communication: quantifying the cost of cell-phone interruptions in face to face dialogs. *PloS one*, 2015 in press".

[84] Chen-Hui Lu and Robert W Proctor. The influence of irrelevant location information on performance: A review of the simon and spatial stroop effects. *Psychonomic bulletin & review*, 2(2):174–207, 1995.

[85] Xin Lu, Linus Bengtsson, and Petter Holme. Predictability of population displacement after the 2010 haiti earthquake. *Proceedings of the National Academy of Sciences*, 109(29):11576–11581, 2012.

[86] Monica Luciana and Charles A Nelson. The functional emergence of prefrontally-guided working memory systems in four-to eight-year-old children. *Neuropsychologia*, 36(3):273–293, 1998.

[87] Robert J MacFadden and Dick Schoech. Neuroscience, the unconscious and professional decision making: Implications for ict. *Journal of Technology in Human Services*, 28(4):282–294, 2010.

[88] Colin M MacLeod. Half a century of research on the stroop effect: an integrative review. *Psychological bulletin*, 109(2):163, 1991.

[89] Thomas W Malone. Toward a Theory of Intrinsically Instruction Motivating. *Cognitive Science*, 5(4):333–369, 1981.

[90] Anne T McCartt, Laurie A Hellinga, and Keli A Bratiman. Cell phones and driving: review of research. *Traffic injury prevention*, 7(2):89–106, 2006.

[91] Danielle S McNamara. *Reading comprehension strategies: Theories, interventions, and technologies*. Psychology Press, 2012.

[92] George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[93] Sherwyn P Morreale, Brian H Spitzberg, and J Kevin Barge. *Human communication: Motivation, knowledge, and skills*. Cengage Learning, 2007.

[94] Donald H Naftulin, John E Ware Jr, and Frank A Donnelly. The doctor fox lecture: A paradigm of educational seduction. *Academic Medicine*, 48(7):630–5, 1973.

[95] Sharlene D. Newman, Patricia A. Carpenter, Sashank Varma, and Marcel Adam Just. Frontal and parietal participation in problem solving in the Tower of London: fMRI and computational modeling of planning and high-level perception. *Neuropsychologia*, 41(12):1668–1682, 2003.

[96] Sharlene D Newman and Gregory Pittman. The Tower of London: a study of the effect of problem structure on planning. *Journal of clinical and experimental neuropsychology*, 29(3):333–342, 2007.

[97] J-P Onnela, Jari Saramäki, Jörkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336, 2007.

[98] Jukka-Pekka Onnela, Jari Saramäki, Jörkki Hyvönen, Gábor Szabó, M Argollo De Menezes, Kimmo Kaski, Albert-László Barabási, and János Kertész. Analysis of a large-scale weighted network of one-to-one human communication. *New Journal of Physics*, 9(6):179, 2007.

[99] Ilse Oosterlaken and Jeroen van den Hoven. Ict and the capability approach. *Ethics and Information Technology*, 13(2):65–67, 2011.

[100] Seymour Papert. *Mindstorms: Children, computers, and powerful ideas.* Basic Books, Inc., 1980.

[101] Bruce D Perry. Childhood experience and the expression of genetic potential: What childhood neglect tells us about nature and nurture. *Brain and mind*, 3(1):79–100, 2002.

[102] Michael Petrides and Brenda Milner. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 20(3):249–262, 1982.

[103] Michael I Posner and Mary K Rothbart. *Educating the human brain.* American Psychological Association, 2007.

[104] Roger Ratcliff. A theory of memory retrieval. *Psychological review*, 85(2):59, 1978.

[105] Roger Ratcliff and Jeffrey N Rouder. Modeling response times for two-choice decisions. *Psychological Science*, 9(5):347–356, 1998.

[106] Jessa Reed, Kathy Hirsh-Pasek, and Roberta Michnick Golinkoff. Learning on hold: Cell phones sidetrack parent-child interactions. *Developmental psychology*, 2017.

[107] Arbora Resulaj, Roozbeh Kiani, Daniel M Wolpert, and Michael N Shadlen. Changes of mind in decision-making. *Nature*, 461(7261):263–266, 2009.

[108] K. Richard Ridderinkhof and Maurits W. Van Der Molen. Mental resources, processing speed, and inhibitory control: A developmental perspective. *Biological Psychology*, 45(1-3):241–261, 1997.

[109] Ricardo Rosas, Miguel Nussbaum, Patricio Cumsille, Vladimir Marianov, Mónica Correa, Patricia Flores, Valeska Grau, Francisca Lagos, Ximena López, Verónica López, et al. Beyond nintendo: design and assessment of educational video games for first and second grade students. *Computers & Education*, 40(1):71–94, 2003.

[110] Mary K Rothbart, Lesa K Ellis, and Michael I Posner. Temperament and self-regulation. *Handbook of self-regulation: Research, theory, and applications*, 2:284–299, 2004.

[111] M Rosario Rueda, Jin Fan, Bruce D McCandliss, Jessica D Halparin, Dana B Gruber, Lisha Pappert Lercari, and Michael I Posner. Development of attentional networks in childhood. *Neuropsychologia*, 42(8):1029–1040, 2004.

[112] M Rosario Rueda, Mary K Rothbart, Bruce D McCandliss, Lisa Saccomanno, and Michael I Posner. Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the national Academy of Sciences of the United States of America*, 102(41):14931–14936, 2005.

[113] M Soledad Segretin, Sebastián J Lipina, M Julia Hermida, Tiffany D Sheffield, Jennifer M Nelson, Kimberly A Espy, and Jorge A Colombo. Predictors of cognitive enhancement after training in preschoolers from diverse socioeconomic backgrounds. *Frontiers in psychology*, 5, 2014.

[114] Charlott Sellberg and Tarja Susi. Technostress in the office: a distributed cognition perspective on human–technology interaction. *Cognition, Technology & Work*, 16(2):187–201, 2014.

[115] Patrick Shafto and Noah Goodman. Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In *Proceedings of the thirtieth annual conference of the cognitive science society*, pages 1632–1637. Cognitive Science Society Austin, TX, 2008.

[116] Patrick Shafto, Noah D Goodman, and Thomas L Griffiths. A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive psychology*, 71:55–89, 2014.

[117] Tim Shallice. Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 298(1089):199–209, 1982.

[118] Filippo Simini, Marta C González, Amos Maritan, and Albert-László Barabási. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, 2012.

[119] Herbert A Simon and Allen Newell. Human problem solving: The state of the theory in 1970. *American Psychologist*, 26(2):145, 1971.

[120] J Richard Simon and Kevin Berbaum. Effect of conflicting cues on information processing: the 'stroop effect'vs. the 'simon effect'. *Acta psychologica*, 73(2):159–170, 1990.

[121] Diego Fernandez Slezak, Andrea P Goldin, Matias Lopez-rosenfeld, Mariano Sigman, Patricia Forradellas, and Adrian Anacleto. Up-scaling Mate Marote : a university-industry interaction experience. In *43 JAIIO - 8 Jornadas de Vinculacion Universidad-Industria, JUI*, pages 19–24, 2014.

[122] Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.

[123] Elizabeth Spelke. Initial knowledge: Six suggestions. *Cognition*, 50(1):431–445, 1994.

[124] Herman H Spitz, Nancy A Webster, and Suzanne V Borys. Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. *Developmental Psychology*, 18(6):922–930, 1982.

[125] Manfred Spitzer. *Demencia digital: el peligro de las nuevas tecnología*. Ediciones B, 2013.

[126] Sidney Strauss, Cecilia I Calero, and Mariano Sigman. Teaching, naturally. *Trends in Neuroscience and Education*, 3(2):38–43, 2014.

[127] David L Strayer, Frank A Drews, and William A Johnston. Cell phone-induced failures of visual attention during simulated driving. *Journal of experimental psychology: Applied*, 9(1):23, 2003.

[128] Joshua B Tenenbaum, Charles Kemp, Thomas L Griffiths, and Noah D Goodman. How to grow a mind: statistics, structure, and abstraction. *Science (New York, N.Y.)*, 331(6022):1279–1285, 2011.

[129] M Usher and JL McClelland. On the time course of perceptual choice: A model based on principles of neural computation. *Psychological Review*, 108:550–592, 2001.

[130] MJvd Van der Molen, JEH Van Luit, MW Van der Molen, I Klugkist, and MJ Jongmans. Effectiveness of a computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research*, 54(5):433–447, 2010.

[131] Luis Von Ahn. *Human Computation*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2005. AAI3205378.

[132] Luis Von Ahn and Laura Dabbish. Designing games with a purpose. *Communications of the ACM*, 51(8):58–67, 2008.

[133] Chong-Wen Wang, Rainbow TH Ho, Cecilia LW Chan, and Samson Tse. Exploring personality characteristics of chinese adolescents with internet-related addictive behaviors: Trait differences for gaming addiction and social networking addiction. *Addictive behaviors*, 42:32–35, 2015.

[134] David Whitebread, M Basilio, M Kuvalja, and M Verma. The importance of play. *Brussels, Belgium: Toys Industries for Europe*, 2012.

[135] Wikipedia. List of countries by number of mobile phones in use. `http://en.wikipedia.org/wiki/List_of_countries_by_number_of_mobile_phones_in_use` (Retrieved Jun, 2014).

[136] Benjamin R. Williams, Jonathan S. Ponesse, Russell J. Schachar, Gordon D. Logan, and Rosemary Tannock. Development of inhibitory control across the life span. *Developmental Psychology*, 35(1):205–213, 1999.

[137] Richard Wilton. The psychology of learning and motivation. *Psyccritiques*, 17(11):612–613, 1972.

[138] ChienHsing Wu and Chia-Fang Liu. Acceptance of ict-mediated teaching/learning systems for elementary school teachers: Moderating effect of cognitive styles. *Education and Information Technologies*, 20(2):381–401, 2015.

[139] Ye Wu, Changsong Zhou, Jinghua Xiao, Jürgen Kurths, and Hans Joachim Schellnhuber. Evidence for a bimodal distribution in human communication. *Proceedings of the national academy of sciences*, 107(44):18803–18808, 2010.

[140] Diyi Yang, Tanmay Sinha, David Adamson, and Carolyn Penstein Rosé. Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Proceedings of the 2013 NIPS Data-driven education workshop*, volume 11, page 14, 2013.

[141] Bu Zhong, Marie Hardin, and Tao Sun. Less effortful thinking leads to more social networking? the associations between the use of social network sites and personality traits. *Computers in Human Behavior*, 27(3):1265–1271, 2011.

[142] Daniel John Zizzo. Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98, 2010.

[143] Jennifer M Zosh, Sarah Roseberry Lytle, Roberta Michnick Golinkoff, and Kathy Hirsh-Pasek. Putting the education back in educational apps: How content and context interact to promote learning. In *Media Exposure During Infancy and Early Childhood*, pages 259–282. Springer International Publishing, 2017.