SCHOOL AGE CHILDREN'S COGNITION IDENTIFICATION BY MINING INTEGRATED COMPUTER GAMES DATA

Rob Whent[1], Dragana Martinovic[2], C. I. Ezeife[3*], Sabbir Ahmed[3], Yanal Alahmad[3] and Tamanna Mumu[3]

[1]*OTEP Inc, 13300 Tecumseh Road East, Suite 366, ON N8N 4R8, Tecumseh, Ontario, Canada*

[2]*Faculty of Education, University of Windsor, Windsor, Ontario, N9B 3P4, Canada*

[3]*School of Computer Science, University of Windsor, Windsor, N9B 3P4, Ontario, Canada*

Keywords: Cognitive Skill Mapping, Computer Games, Database Integration, Database Mining.

Abstract: National statistics confirm that nowadays about 20% of school children have some type of mental health issue, about 70% of adult mental health disorders originated in adolescence, while about 40% have unidentified learning differences that affect their learning abilities. Starting early enough with proper screening of a child's cognitive skills is critical for improved learning and mental wellbeing. Unfortunately, assessments and treatment can be costly, elusive or conflicting. This paper describes an approach to identifying child's cognitive skill level that is adopted by an online product called "Think2Learn" developed by OTEP Inc. (Online Training & Education Portal). OTEP uses ubiquitousness of the Internet and attractive features of online computer games to give parents automated opportunity to screen and follow their children's cognitive development. OTEP presently uses a collection of approximately 100 video games for children to play and while doing so, it records their score to continuously assess and monitor their cognitive strengths and weaknesses. The Web-based tool for identifying cognitive skill level is developed as an integration or data warehouse of a number of relevant data sources such as the cognitive skills categories data, games data, player inventory data and so on. The integrated data are continuously mined, analyzed and queried for proper and quick assessment or recommendations.

1 INTRODUCTION

Currently available digital technologies provide for multiple representations (e.g., visual, text, and sound) that draw youth as attractive, artistic and fast-paced. There is a body of research that points to unique learning habits of young people. Among else, they prefer short visual explanations, to receive information quickly and process it rapidly, prefer multi-tasking and non-linear access to information, have a low tolerance for lectures, prefer active rather than passive learning, and are kinaesthetic, experiential, hands-on learners who must be engaged with first-person learning, games, simulations, and role-playing ((Junco and Mastrodicasa, 2007); (Oblinger and Oblinger, 2005); (Tapscott, 2009)). The youth nowadays also rely heavily on communication technologies to access information and to carry out social and professional interactions ((Veen and Vrakking, 2006); (Pletka, 2007)). At the same time, playing computer and video games are lately being recognized as valid cognitive activities as they affect a player's capability to self-regulate, make right decisions, and problem-solve. Such environments are engaging but put a strain on the cognitive load and attention span of the user. It is notable that some parents deliberately avoid having computers at home, while some restrict access for their children out of fear that they will use computers to play games rather than for educational purposes (Dance, 2003). How should educators respond? Is there a way to convince parents that their children may learn while they are playing computer games?, that children may enhance their cognitive capabilities while having fun on a computer?. Cognition refers to the mental processes involved in gaining knowledge and comprehension, including thinking, knowing, re-

*This research was supported by the Natural Science and Engineering Research Council (NSERC) of Canada under an Operating grant (OGP-0194134) and Engage program.

Whent R., Martinovic D., Ezeife C., Ahmed S., Alahmad Y. and Mumu T..
SCHOOL AGE CHILDREN'S COGNITION IDENTIFICATION BY MINING INTEGRATED COMPUTER GAMES DATA.
DOI: 10.5220/0003917204950505
In *Proceedings of the 4th International Conference on Computer Supported Education* (SGoCSL-2012), pages 495-505
ISBN: 978-989-8565-07-5

membering, judging and problem-solving. These are higher-level functions of the brain that encompass language, imagination, perception and planning skills.
According to the U.S. Surgeon General, national statistics confirm that nowadays about 20% of school children have some type of mental health issue; the Mental Health Commission of Canada (Canada, 2011) states that 70% of adult mental health disorders originated in adolescence, while about 40% have unidentified learning differences that affect their learning abilities (e.g., the acquisition, retention, understanding, organization of information). Starting early enough with proper screening of a child's cognitive skills is critical for improved learning and mental wellbeing. Unfortunately, assessments and treatment can be costly, elusive or conflicting. This paper describes an approach to identifying a child's cognitive skill level that is adopted by an online product called "Think-2-Learn" (Whent, 2012) created by OTEP Inc. (Online Training & Education Portal). Cognitive abilities of individuals such as auditory, visual, sequential, conceptual, speed, and executive determine their learning abilities in the main cognitive categories such as basic reading, reading comprehension, math calculation, math reasoning, writing mechanics, writing content, oral expression, and listening comprehension. According to (Crouse, 2010), student needs to achieve, for example, high score in auditory domain, high score in conceptual and moderate score on speed of processing in order to have ability for reading comprehension. Determining a child's cognitive abilities in various categories of cognition and later identifying any improvement in these abilities in a natural environment such as their performances when playing relevant computer games is an approach being adopted in this paper. This paper proposes integrating into a data warehouse a number of relevant data sources such as video games data, cognitive data, student game play inventory and others, in other to quickly mine and compare their performances with the norms for cognitive achievements for different children from various age, gender, mental, physical, ethnic, social and other background they may belong to. While this approach does not claim to be a panacea for alleviating developmental difficulties in children, it should be taken as one attempt that, in conjunction with other preventive and treatment measures, may help children to improve their cognitive skills and strategies. More research is needed to establish to what extent is this approach helpful and in which cases; however our preliminary case studies and data mining method provide the ground work for future studies.

Section 2 presents the related work including those in computer gaming and cognition, data warehousing and mining. Section 3 presents the OTEP solution approach including the data warehouse integration algorithm, schema and example application. Section 4 discusses performance analysis of how games data can be effective in identifying cognition while section 5 presents conclusions and future work.

2 RELATED WORK

2.1 Youth and Gaming

Cognitively, young generations are known to have a shorter attention span and to need immediate answers (Pedro, 2006). This may be a consequence of the extensive propagation of video games among youth. In (Rideout et al., 2005), it is noted that in the U.S., 8-to10-year-olds spend more than an hour a day playing video games. Since the main features of many video games are quick reactions and immediate feedback, the games may reinforce the children's inclination towards fast, focused, and repetitive actions that result in direct (task-oriented and instant-feedback-based) and limited (fast, focused, repetitive) learning in a short time. In (Pedro, 2006), it is concluded that "nothing seems further from this than the students' daily school experiences requiring longer attention spans, engaging students in reflective activities, focusing intensely on one activity at a time, and involving properly written text" (p.11). The learning habits of youngsters not only differ from traditional learning habits, but they also demonstrate certain unique weaknesses. For example, nowadays students are strong visual learners (Berk, 2010), but usually weak in learning from text (Vaidhyanathan, 2008). In addition, while such students prefer fast-paced environments, their performance sometimes lacks depth and critical thinking ((Oblinger and Hawkins, 2006); (Rockman and Associates, 2004)).

2.2 Computer Games and Cognition

There exist conflicting reports on the social and cognitive consequences of playing computer and video games (Martinovic et al., 2011). In the interest of space, we will focus on studies that relate playing games to positive effects. There are recorded benefits emerging from playing computer games, such as improving visual intelligence skills (which are especially relevant for subject areas in which manipulating images on a screen may be particularly useful, such as science and technology) (Subrahmanyam et al., 2000). Computer-based games may enhance hand-eye coordination, visual scanning, auditory discrimination, and spatial skills (DeLisi and Wolford, 2002).

Moreover, emphasis on visual information processing may be connected to a significant increase in average non-verbal scores in various psychological tests across all groups tested (see (Greenfield et al., 1994), (Subrahmanyam et al., 2000)). A comparative study of children aged 10 to 11 who played two different computer games, one with strong visual content and the other text-based, showed that playing the first game improved the children's spatial performance, while playing the second did not (Subrahmanyam and Greenfield, 1994). Repetitive game playing may increase the young children's working memory (Thorell et al., 2009); mental rotation accuracy (DeLisi and Wolford, 2002); and spatial rotation, iconic skills, and visual attention (Subrahmanyam et al., 2001).

2.3 Computer Games and Learning

Playing the carefully and purposefully designed computer games may positively affect learning among children of wide range of ages. Early studies on high school students' use of educational software at home showed significantly higher scores on computer literacy tests among students who played these games (Subrahmanyam et al., 2000). Describing the study done with pre-school children, (Freeman and Somerindyke, 2001), concluded that the children benefited and thrived when given the opportunity to develop skills through the use of computers. For these authors, computers "are an ideal vehicle for learning in a social setting" (p. 206), and as such, have a role in an active-learning, play-based curriculum. Because playing computer games involves integration of "touch, voice, music, video, still images, graphics, and text" (IBM, 1991), children are faced with the environment geared to a variety of intelligences (e.g., linguistic, logical, spatial, kinaesthetic, musical), this may particularly influence development of literacy skills and ability to problem-solve. In (Gee, 2007), it was argued that as "problem-solving spaces", video games can be excellent tools for learning, as they provide an environment for recursive practices. It is through these practices that learners can reach the point of automaticity in a "non-boring" way. Players see computer games as systems, rather than as collections of discrete skills. In addition, exploration and successive challenges of increasing difficulty are primary principles of computer games. Beck and Wade (as cited in (Bogost, 2007), p. 240), maintain that the "videogame generation" (those born after 1970) is uniquely positioned to use meta-cognitive skills obtained through video game playing (e.g., reflecting on immediate situation, analyzing choices and comparing odds, finding the right strategy). In (Bogost, 2007), Gee's notion of computer games was taken as "situated … and … embodied learning" further, stating that "videogames … offer meaning and experiences of particular worlds and particular relationships" (p. 241, emphasis in original). In other words, "the higher-order thinking skills still matter, but so does the ninja" (p. 243). These statements are supported by the findings from the neurological research that identify emotion as one important element of learning (Zull, 2004). Another element of learning is practice, and both are present in playing computer games.

2.4 Examples of Studies Involving Game Playing

There are several recommendations coming out of previous work that we took into account in this study. For example, Ko (2002) analyzed 7-10 year old children's behaviour and cognitive development while playing a game based on 'if-then' rules. The author used statistical methods to find if the child used the game rules or achieved the goal by chance only, and analyzed children's game performance in relation to their individual differences. Ko concluded that the success in playing depended on the child's age and planning skills; also, the children improved in using the game clues and developed their planning strategies through repeated game playing. Ko's recommendation is that educational software must be measured against how children think and learn (e.g., the purpose of the game, activities during the play, what and how the child performs, and what concepts are learned). Computer games of different types affect children cognitive processes differently. (Pillay, 2002) investigated the impact of two recreational computer games on 14-16 year old children's subsequent performance in a computer based instructional tasks, by employing both quantitative (i.e., speed and accuracy) and qualitative (i.e., cognitive strategies) measurements. The author found that playing recreational computer games may influence performance (e.g., time efficiency and correct problem solving) on subsequent computer-based educational tasks. Pillay concluded that although recreational and educational games may have different goals, commonality in their structure is what matters. By playing games of different types, children develop a repertoire of cognitive schemas that can help them later in performing learning tasks. To conclude, (Lieberman et al., 2009) suggested that more research is needed to understand cognitive and other development of children when they play computer games in order to support their healthy emotional, cognitive and learning development. We continue our report by introducing the OTEP approach and philosophy.

2.5 Data Warehousing and Mining Approaches

A data warehouse is a historical, integrated, subject-oriented database storing data from multiple data sources in the one data warehouse schema (Han et al., 2011). Construction of a data warehouse is done through processes of schema and data integration of different data sources which involve data cleaning (Ezeife and Ohanekwu, 2005), data transformation and loading with periodic refreshing. A popular data warehouse schema approach is the star schema where there is a central fact table having foreign key attributes that include the main subjects of interest, the integration attribute, the historical time attribute and some non-foreign key aggregate measures of interest. Other descriptive tables in the data warehouse design using the star schema are dimension attributes for describing the foreign key attributes in the fact table. A measure such as score achieved during a game by a child can be calculated from a multidimensional model version of the data warehouse called the data cube (Ezeife, 2001). Other online analytical processing (OLAP) queries such as drill-down and roll-up analysis can be performed on the data warehouse fact table or cube. Also, data mining querying for analyzing the correlation between groups of data in the warehouse such as association rule mining and clustering can be done to answer required queries.

3 THE OTEP TECHNIQUE OVERVIEW

3.1 The OTEP Model and Problem Addressed

As shown in the OTEP model of Figure 1, one goal of the model is to measure a child's cognitive abilities through its performances in repetitive playing of a variety of games in different cognitive categories. The model proposes to accomplish this goal by comparing the child's performance in these games with the performances of dynamically changing normalized performances (termed norms) of other children in similar comparison groups such as age, ethnic background, social background, learning or physical level, etc. In order to gather the necessary historical, integrated and correlated game playing scores and parameters of each child, this approach proposes using the data warehouse approach to integrate structured data sources. The structured data sources to be integrated consist of game playing database, cognitive inventory database, and other data sources such as learning inventory database. The game playing database can also result from a continuous integration of various gaming sites. From the integrated data warehouse, online analytical processing and mining of data for comparative purposes can be performed. For example, some analysis can be accomplished using multidimensional data cube views (Ezeife, 2001).

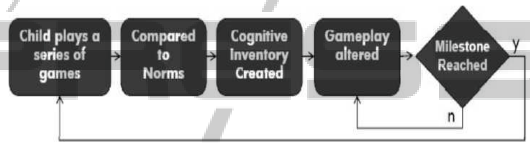

Figure 1: The OTEP Model.

Some Challenges to be Resolved
Some of the questions to be answered include:
1. The data warehouse techniques to be used for continuous integration of both database schemas and data of different game sites as different attribute names need to be merged as one name in the data warehouse. Also, data cleaning needs to be done to resolve differences in units of data (e.g., when one database stores a games parameter in inches and the other stores them in meters).
2. Establishing the approach for computing and revising the 'Norms' data. For example, the Norms for children who are 8 years old and who are playing a game involving reading can be computed using the average speed over a period of time. However, many other ways for computing the Norms are to be investigated.
3. Data structuring, integration and mapping of the cognitive inventory to various levels of Norms and performances as well as to learning inventory.
4. Determining types of analysis and reporting that are needed to indicate a child's cognitive abilities and improvement from the integrated data. Multidimensional online analytical data cube approach is one technique to be explored.
5. Establishing firmer correlations between playing visual games for example, to learning skills.
6. What does it mean when a first level of a game is played in comparison to when a second, third or higher level is played with regards to performance in cognitive abilities?

Table 1: General areas of cognitive processing and related abilities (based on Crouse, 2007).

Cognitive Category	Sub-categories
Visual processing	seeing differences between things, remembering visual details, filling in missing parts in pictures, remembering characteristics, visual-motor coordination, visualization and imagination, organization of their room, desk, artistic skills
Auditory Processing	hearing differences between sounds/voices, remembering specific words or numbers, remembering general sound patterns, understanding when they miss some sounds, blending parts of words together, music
Sequential Rational	Short-term memory for details, long-term retrieval of details, fine-motor coordination, finding the words you want to say or write, organization of your thoughts and materials, writing mechanics (spelling, punctuation), reading speed/sounding out new words, attention to details, putting words and thoughts in order
Conceptual Abstract	memory for general themes or ideas, reasoning spatial awareness, general knowledge, inferential thinking, estimation/approximation, conceptual understanding, creativity/inventiveness, reading comprehension, use of context rhythm, music, art
Processing Speed	short-term memory (with time pressure), long-term retrieval (with time pressure), talking speed, word-finding, writing speed, reading speed, attention reasoning (with time pressure), general response speed
Executive Functioning	ability to stay focused on tasks, plan and anticipate, organize thoughts and materials, follow-through and complete tasks, cope with unstructured situations, cope with changes in routine, regulate emotions

Table 2: Correlation between areas of cognitive processing and student achievement.

Cognitive skill	Auditory	Visual	Sequential	Concept	Speed	Executive
Basic Reading	high	moderate	high		high	
Reading Comprehension	high			high	moderate	
Math calculation		high	moderate		moderate	high
Math Reasoning		high		high	moderate	moderate
Writing Mechanics	high	high	high		high	high
Writing Content	high			high	moderate	
Oral expression	high		moderate		moderate	
Listening Comprehension	high		moderate			

7. How do we compare a single person's record and compare it to hundreds of people's game records.

3.2 The Cognitive Categories

The cognitive categories being addressed or targeted are based on six general areas of cognitive processing: Visual processing; Auditory processing; Sequential/rational processing; Conceptual/abstract processing; Processing speed; and Executive Functioning that involve abilities listed in Table 1.

3.3 The Video Games Categories and Mapping

The current collection of games has about one hundred games in ten categories encompassed by the six general areas of cognitive processing of Table 1. These ten cognitive categories are: 1. Visual Processing, 2. Processing Speed, 3. Reasoning, 4. Self Direction, 5. Self Regulation, 6. Auditory Processi ng, 7. Language Processing, 8. Motor Functioning, 9. Memory/Learning, and 10. Flexibility. Each of these main game categories has sub categories of games they contain. For example, the visual processing category has games that are classified as one of the four types of i) puzzle games (e.g., number balls, the missing jigsaw and find the pair), ii) action games (e.g., fruit collection, mouse and cat, christmas gifts), iii) sports games (e.g., free kick, freestyle soccer, asteroids) and iv) educational games (e.g., word search II, whack a difference). Some games may belong to more than one main game category. For example, motor functioning games consist of some types of action and sports games that may include some games of those types in the visual processing category. The games inventory continues to expand and each of these games has a web link where they can be played. The games results are classified as one of the following five outcomes: 1) extremely below average, 2) below average, 3) average, 4) above average, or 5) extremely above average.

3.4 Theoretical Correlation between Cognitive Processing and Student Achievement

According to (Crouse, 2010), a student needs to achieve, for example, high score in auditory domain, high in conceptual and moderate score on speed of processing in order to have the ability for reading comprehension. Similarly, one can determine student's strengths or weaknesses in the domains of mathematics calculation or oral expression, and similar (see Table 2). Table 2 presents the relationship between cognitive processing and student achievements. From Table 2, assessing a child's math calculation skills or abilities as being strong can be done with tests that would ascert the child's visual skills as high, conceptual skills as high, processing and executive skills as moderate. The OTEP approach is to use a natural testing environment consisting of computer games where these skills such as auditory, visual, sequential, conceptual, speed and executive are directly measured through the child's performance in repetitive playing of games in various categories.

3.5 The OTEP Data Warehouse Integration Approach

The goal of this system is to use the online games to screen or assess children's cognitive skills and later suggest a learning plan that would be most suitable for their learning success. While this paper describes gathering and integrating the relevant data from (1) video games data, (2) cognitive skills and mapping data and to obtain a data warehouse schema called OTEP_GamesDW, in the future, other data sources will be integrated including the learning achievement data and third party data. The current schemas of the games data source and the cognitive data source with the integrated data warehouse are provided in this section.

The Games Data Source
The database schema for the games data source to be integrated holds data descriptors for each of the 100 or more games that a child can play. Information on the games include the type of game it is. For example, a game called "Building Blocks" belongs to the game category type of "Puzzle" with a description of the type of activity involved as "Scanning". For this game, the maximum allowed number of replays is 3 and the web link where this game can be found is included as well as the override options that this game can be set on. Other tables that are presented, contain games instance, and records of each user's play of each game. The games data source schema is represented through the following eight table schema.:

1. Usertable(userid, useridc, userlogon, usermail, timecreated);
2. Gamestable(gameid, gametag, gamepath, swfpath, confpath, gametype, title, typedesc, maxplay, link, options);
3. GamesCategorytable(catid, cattag);
4. Gamesidentification(catid, gameid);
5. GamesConfiguration(confid, gameid, configmatch, timestamp_config, configuration);
6. GameInstance(usergamesid, userid, catid, insttag, timecreated_inst, timecompleted_inst, gameslist);
7. UserGameInstance(usergamesid, gameid, timeplayed, timecompleted, gamecompleted, numfinished, numtries); and
8. UserGamePlayLog(usergamesid, gameid, timestamp, confid, status, completed, playresult);

The Cognitive Data Source
The current cognitive categorization has each game belonging to one of the following ten main categories: 1. Visual Processing, 2. Processing Speed,

3. Reasoning, 4. Self Direction, 5. Self Regulation, 6. Auditory Processing, 7. Language Processing, 8. Motor Functioning, 9. Memory/Learning, and 10. Flexibility. Each of these main cognitive game categories (i.e., Visual Processing, Processing Speed, Reasoning, Self Direction, Self Regulation, Auditory Processing, Language Processing, Motor Functioning, Memory/Learning, and Flexibility) is then mapped to one of the currently defined eight categories. The sub categories are defined based on the main category. For the ten main categories, the corresponding two to eight subcategories are provided next to them below:

1. Visual Processing has Saccadic tests, visual array, feature detection, scanning, symbol processing, visual perception, spatial perception, and non-verbal solving.
2. Processing Speed has speed-efficiency, verbal output, and written output.
3. Reasoning has verbal reasoning, complex problem solving, non-verbal reasoning, and social component.
4. Self-Direction has planning, strategy generation, organization, prioritizing, initiation/activation, and task monitoring.
5. Self-Regulation has emotional, motor output (writing), self-monitoring, attention, behavior, and cognitive inhibition.
6. Auditory Processing has conduction, perception, processing, direction following, and comprehension.
7. Language Processing has language concepts, verbal logic test, reading, verbal expression, word finding, and spelling.
8. Motor Functioning has speed, dexterity, static steadiness, motor planning, and visual motor processing.
9. Memory/Learning has verbal, non-verbal, spatial, and knowledge aspect.
10. Flexibility has working memory, and set shifting.

The following database schema presents the cognitive data source, which describes the cognition levels and the connection to game play instances.

1. UserCogMeasure(cogid, cogsubid, usergamesid, userid, usercatid, normcogid, measure);
2. GamesCogMap(gameid, cogid, cogsubid);
3. CognitionCat(cogid, cogdescr);
4. CognitionTable(cogid, cogsubid, cogsubdesc);
5. GamesCogNorm(gameid, age, sex, highscore, mediumscore, lowscore);

The Data Warehouse Schema

Using the Star schema which consists of a central main fact table called FactTable, the data warehouse contains an integration of the relevant data from the Games data source and the Cognitive data source that will enable continuous answering of needed queries. One such needed query is “Given the game play data of a child, what is the cognition language processing skill of this child (an in particular, her word finding abilities) in comparison with the normal cognition level for this game and for this child’s age, ethnic, and other group the child may belong to?”. In addition to the FactTable, the data warehouse star schema also contains dimension tables (e.g. Users-Dim) , which are descriptor tables for the foreign key attributes (e.g., userid) in the FactTable. The following schema is similar to the current data warehouse.

1. FactTable(userid, gamid, gameseq, gameDB, gamelevelid, catid, normcogid, cogid, cogsubid, time-m, coglevel, gamescore, duration, tries);
2. GameCategory-Dim(catid, catname);
3. Users-Dim(userid, gender, age, userlogon, user-mail);
4. GamesDB-Dim(gameDB, weblink);
5. GamesLevel-Dim(gamelevelid, gameid, levelid);
6. Games-Dim(gameid, title, maxplay, gametag, gamepath, gametype, link, cogid, cogsubid);
7. Cognition-Dim(cogid, cogsubid, cogdesc);
8. CogCategory-Dim(cogid,cogdescr);
9. CogGameMap-Dim(gameid,cogid,cogsubid);
10. GameCogNorm(gameid,age,sex, highScore, mediumScore, lowScore);

3.6 The Data Integration Algorithm

The data integration algorithm used to integrate the various relevant data sources into a single historical, subject-oriented, data warehouse which is refreshed periodically and mined to obtain changing normal data and user data is provided as algorithm 1. The current implemenation details of the system are as follows: the data warehouse and source databases are stored in mySQL database management system, the data integration algrorithms are implemented using mySQL stored procedures and triggers. The interface for all the databases including the data warehouse is implemented using PHP, JavaScript and jQuery tools. The OLAP interface for multi-dimensional querying and data cube views is implemented using JAVA, JSP, Mondrian OLAP engine and MDX language for data cube querying.

3.7 Using MultiDimensional Data Cube Views for Queries

From the historical data in the data warehouse FactTable, a data cube can be used to compute normal cognitive measures and other results. For example, the average score of a specific game such as ‘Number

Algorithm 1: The OTEP_GamesDW Game Play Integration Algorithm.

```
Algorithm OTEP_GamesDW()
Input:     New GamePlay DB (GDB), Cognitive DB (CDB)
           Data Warehouse (DWH)
Output:    Clean updated Data Warehouse (UDWH)
Variables:
BEGIN
UDWH = DWH
FOR each new record in GDB.UserGamePlayLog
BEGIN
EXTRACT and CLEAN
(userid, gameid, gameDB, gamelevelid,
tries, score, timereq, completed)
FROM GDB.UserGamePlayLog
IF userid NOT IN UDWH.Users-Dim
    THEN
    EXTRACT and CLEAN User
    information FROM DS.usertable
    INSERT into UDWH.Users-Dim
    END
IF gameid NOT IN UDWH.Games-Dim
    THEN
    EXTRACT and CLEAN Games
    information including
    FROM GDS.gametable
    INSERT INTO UDWH.Games-Dim
    END
IF gameid is not IN
UDWH.CognitiveMain-Dim
    THEN
    EXTRACT and CLEAN Cognitive
    Categories of gameid FROM CDB
    INSERT INTO UDWH.CognitiveMain-Dim
    END
isVALID = DataValidatonProcess
(GamePlayData)
IF (GDB.UserGamePlayLog.completed
= 'Y' AND isVALID = 'TRUE')
BEGIN
    INSERT
    (userid, gameid, gameDB,
    gamelevelid, tries, score,
    timereq, completed )
    INTO UDWH.FactTable
END
END
END
```

Balls' for different ages of children between 3 and 8 years, during several attempts of 1st, 2nd, 3rd and 4 tries. This is accomplished with an SQL query on the fact table data that computes the average gamescore group by user age and group by tries. The sample data cube view for this query is given as Figure 2.

3.8 An Example Application of our Olap Approach for Queries

The algorithm 1 provides a way to keep updating a data warehouse that is built as an integration of a number of relevant data sources such as games data and cognitive data. The data warehouse now has the central fact table and the adjoining dimension tables and multi-dimensional or data cube querying (also called online analytical processing) can be used to answer a lot of needed comparative queries. Thus, when a child plays a game, a relevant record from the game play session that is integrated into the data warehouse has the schema

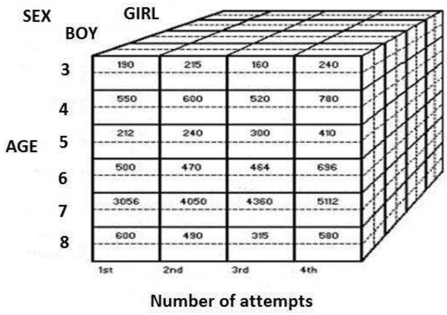

Figure 2: Example compute norm score cube.

GameplayRecord (usergamesid, gameid, timestamp, completed, value);

An example of such stored record is:

(26, 28, 2010-11-09 02:04:48, Y, {"totaltime":253.496, "levels":({"time":253.496, "success":0, "score":140, "level":1}), "totalscore":140, "status":3})

The sample data warehouse fact table schema after integration and a sample record from this schema are given below.

FactTable(userid, gamid, gameseq, gameDB, gamelevelid, catid, normcogid, cogid, cogsubid, time-m, coglevel, gamescore, duration, tries);

(33, 28, 20101109020448, 'think2learn', 1, 1, 5, 4, 20, '2010-11-09 02:09:48', 3, 140, 253.496, 4)

To answer a query similar to query 1 in section 3.1, given below: "What are the norms (cognitive norms and game achievement norms) for children who are 8 years old and who have difficulty reading for a reading game?". This query may be answered with an SQL instruction on the data warehouse fact table joined with its dimension tables given below:

```
SELECT
u.age,u.gender,g.cogdesc,avg(c.coglevel),
avg(c.score)
FROM cleanfact c,users_dim u,gamecog_dim g
WHERE c.user_id=u.user_id and
c.game_id=g.game_id and g.cogid=100 and u.age=8
GROUP by u.age,u.gender;
```

Sample results from the above query is given below.
(Age, Gender, Cognitive Type, Average Cognition, Average Score)
(8, F, Visual Processing, 2, 1980.00);

(8, M, Visual Processing, 3, 2155.71);

The above query has shown only a slice of the data cube for children of age = 8. To display the entire norms for children of all ages, we would just remove the restriction condition in the "where" clause for the specific age = 8.

4 PERFORMANCE STUDY OR ANALYSIS

In this section we provide readers with an example of user performance analysis based on a sample (N=106) of randomly selected test cases stored in our database. Our analysis takes into account the following variables: (a) demographics information, such as user's age and gender; (b) qualitative information, such as type of game as given by the production company, two major cognitive attributes of the game as determined by an OTEP psychologist, and two cognitive subcategories (4 in total); and (c) quantitative parameters of user performance, such as, score achieved in each game, number of trials for each game level, and time needed to complete each level. After a user plays a number of games from one cognitive cluster, his or her average score for this cognitive category is compared to the normative score of the comparative group of users (in terms of age and gender). The user is then ranked according to the following schema: if the score deviates up to one standard deviation (1SD) from the normative score (mean = M), the user demonstrates the 'average range of skill' for the cognitive category; if the score deviates more than 1SD-2SD from the normative score, the user demonstrates the 'moderate relative' strength or weakness (depending on whether the user's performance is above or below the average score); similarly, 'significant relative' strength or weakness is recorded if the user deviates 2SD-3SD from the average. The idea is to recommend action based on the child's current performance: if the child has scores such that point to his/her weakness in the area of Conceptual reasoning (see Table 2, this may affect Math Calculation, Math Reasoning and Writing Content). So, our hypothesis is that if the system suggests to the child more games in the area of Conceptual Reasoning, the child will eventually develop cognitive strengths relevant for his/her school achievement. Recall that standard deviation (SD) is a measure of how spread out numbers (e.g., n game scores $X_1 \ldots X_n$) are from the average or mean (M) and it is it is the square root of the Variance (V). The formulas for computing the mean M, variance V and standard deviation SD are respectively: $M = \sum_{i=1}^{n} \frac{X_i}{n}$
$V = \sum_{i=1}^{n} \frac{(X_i - M)^2}{n}$
$SD = 1/n\sqrt{\sum_{i=1}^{n}(X_i - M)^2}$.
Standard Deviation provides a way to know what is normal range of data (usually those within 1SD or within 68% from the mean), and what is moderate performance (usually those within 1SD-2SD or within 68.1% – 95% from the mean), and those that are extra large (strength) or extra small (weakness) (usually those within 2SD-3SD or within 95.1% – 99% from the mean).

4.1 Performance Study Example 1

Our query showed that in the sample of test data, there were 28 girls and 78 boys, with average age of 13.5 years. There was no statistical difference between boys' and girls' general performance (score or time needed to complete the game(s)). Say that an 8-year old girl (id = 95), played a game titled "Stationery", which was by the vendor described as Type = "spatial" game, and by the psychologist as Cognitive Category 1 = "visual processing", with Subcategory 1= "visual perception". The girl achieved Score = 24185 points on this game, for which it took her 42 minutes to complete. Overall, this child played 5 games in the visual processing category and obtained mean game score of M = 6937 points (SD = 994.4 points), and played for mean time of M = 80 min (SD = 50.3 minutes). The question is how this child compares with other girls or children of similar age who played games in the same category. After performing a one sample t-test, there was no statistically significant difference between the score of this child and other children who played similar games (t = .810, df = 58, $p > .4$). There was no statistically significant difference between the time it took this child to play the visual processing games and it took other children to do so (t = 1.09, df = 59, $p > .2$) However, when this child's average score on visual processing games was compared to scores of other 8-year olds, the statistically significant difference was established (t = -3.30, df = 10, $p < .01$). According to Table 1, the child who is lacking in the visual processing may be at risk in the mathematics calculation and reasoning, and in writing mechanics. Remediation could be recommended in form of playing more games in the related category and following the child's improvement accordingly.

5 CONCLUSIONS

In this paper, we reported on our current work on the the online product called "Think2Learn" developed by OTEP Inc. (Online Training & Education Portal).

OTEP presently contains about 100 video games for testing purposes, it records players' scores to continuously assess and monitor their cognitive strengths and weaknesses with regards to the main cognitive categories. The Web based tool for identifying cognitive skill level is developed as an integration or data warehouse of a number of relevant data sources such as the cognitive skills categories data, games data, player inventory data and so on. The integrated data are continuously mined, analyzed and queried for proper and quick assessment or recommendations.

We are presently working on: (a) increasing number of games; (b) increasing a reliability of categorization of the games by achieving an agreement between 2-3 scorers; (c) developing a formula that will incorporate the features of the game (including differentiating the impact of different cognitive sub categories), the number of trials, the scores achieved and the time spent playing. Once we have established collaboration with some schools, the system will associate with each child a unique identifier tracking the children's cognitive development, proposing remediation, increasing validity and reliability of our approach.

REFERENCES

Berk, R. (2010). How do you leverage the latest technologies, including Web 2.0 tools, in your classroom? . *International Journal of Technology in Teaching and Learning*, 6(1):1–13. .

Bogost, I. (2007). *Persuasive Games: The Expressive Power of Videogames*. The MIT Press.

Canada, M. (2011). Child and youth development. http://www.mentalhealthcommission.ca/English/Pages/ChildandYouth.aspx.

Crouse, S. (2010). *The Cognitive Processing Inventory (CPI)*. LDinfo Publishing.

Dance, F. (2003). The digital divide. In Strate, L., Jacobson, R. J., and Gibson, S., editors, *Communication and cyberspace: Social interaction in an electro nic environment*. Hampton Press, Inc., Cresskill, New Jersey.

DeLisi, R. and Wolford, J. (2002). Improving children's mental rotation accuracy with computer game playing. *Journal of Genetic Psychology*, 163(3):272–282.

Ezeife, C. (2001). Selecting and Materializing Horizontally Partitioned Warehouse Views. *Journal of Data & Knowledge Engineering*, 4(1). Elsevier Sciences.

Ezeife, C. and Ohanekwu, T. (2005). Use of smart tokens in cleaning integrated warehouse data. *International Journal of Data Warehousing and Mining (IJDWM)*, 1(2):1–22.

Freeman, N. and Somerindyke, J. (2001). Social play at the computer: Preschoolers scaffold and support peers' computer competence. *Information Technology in Childhood Education Annual*, 2001(1):203–213.

Gee, J. (2007). *Good Video Games and Good Learning: Collected Essays on Video Games,Learning, and Literacy*. New York: Peter Lang.

Greenfield, P., deWinstanley, P., Kilpatrick, H., and Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology*, 15(1):105–123.

Han, J., Kamber, M., and Pei, J. (August 2011). *Data Mining: Concepts and Techniques* . Morgan Kaufmann Publishers.

IBM (1991). Literacy development. (ibm's writing to read, and principles of adult literacy system multimedia/educational software). T H E Journal (Technologic al Horizons in Education).

Junco, R. and Mastrodicasa, J. (2007). *Connecting to the net.generation: What higher education professionals need to know about today's students*. Student Affairs Administrators in Higher Education (NASPA), Washington, DC.

Lieberman, D., Fisk, M., and Biely, E. (2009). Digital games for young children ages three to six: From research to design. *Computers in The Schools*, 26(4).

Martinovic, D., Freiman, V., and Karadag, Z. (2011). Child and youth development beyond age 6 - transitions to digitally literate adults. Unpublished Report. Ministry of Child and Youth Services.

Oblinger, D. and Hawkins, B. (2006). The myth about student competency. *EDUCAUSE Review*, 41(2). http://www.educause.edu/apps/er/erm06/erm0627.asp.

Oblinger, D. and Oblinger, J. (2005). Educating the net generation. educause. net.educause.edu/ir/ibrary/pdf/pub7101.pdf.

Pedro, F. (2006). The new millennium learners: Challenging our views on ict and learning. http://www.oecd.org/dataoecd/1/1/38358359.pdf.

Pillay, H. (2002). An investigation of cognitive processes engaged in by re creational computer game players: Implications for skills for the future. *Journal of Research on Technology in Education*, 34(3):336–350.

Pletka, B. (2007). *Educating the net generation*. Santa Monica Press, Santa Monica, CA.

Rideout, V., Roberts, D., and Foehr, U. (2005). *Generation M: Media in th e lives of 8-18 year-olds*. The Henry J. Kaiser Family Foundation.

Rockman, I. and Associates (2004). *Integrating information literacy into the higher education curriculum: Practical models for transformation*. Jossey-Bass, San Francisco, CA.

Subrahmanyam, K. and Greenfield, P. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology*, 15(1):13–32.

Subrahmanyam, K., Greenfield, P., Kraut, R., and E.Gross (2001). The impact of computer use on children's and adolescents' development. *Journal of Applied Developmental Psychology*, 22(1):7–30.

Subrahmanyam, K., Kraut, R., Greenfield, P., and Gross, E. (2000). The impact of home computer use and children's activities and development. *The Future of Chil-*

dren: Children and Computer Technology, 10(2):123–14.

Tapscott, D. (2009). *Grown up digital: How the net generation is changing your world.* McGraw Hill, New York, NY.

Thorell, L., Lindqvist, S., Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental Science*, 12(1):106–113. .

Vaidhyanathan, S. (2008). Generational myth: Not all young people are tech-savvy. *The Chronicle of Higher Education*, 55(4). B7.

Veen, W. and Vrakking, B. (2006). *Homo Zappiens, Growing up in a Digital Age*. Network Continuum Education, London.

Whent, R. (2012). Otep think-2-learn: A private online prescreening solution for children. http://www.think-2-learn.com/.

Zull, J. (2004). The art of changing the brain. *Teaching for Meaning*, 62(1). .