

The Measurement of Intelligence in the XXI Century using Video Games

M. A. Quiroga¹, F. J. Román², J. De La Fuente¹, J. Privado¹ and R. Colom³

¹ Universidad Complutense (Spain)

² University of Illinois at Urbana-Champaign (USA)

³ Universidad Autónoma de Madrid (Spain)

Abstract. This paper reviews the use of video games for measuring intelligence differences and reports two studies analyzing the relationship between intelligence and performance on a leisure video game. In the first study, the main focus was to design an Intelligence Test using puzzles from the video game. Forty-seven young participants played “Professor Layton and the curious village”[®] for a maximum of 15 hours and completed a set of intelligence standardized tests. Results show that the time required for completing the game interacts with intelligence differences: the higher the intelligence, the lower the time ($d = .91$). Furthermore, a set of 41 puzzles showed excellent psychometric properties. The second study, done seven years later, confirmed the previous findings. We finally discuss the pros and cons of video games as tools for measuring cognitive abilities with commercial video games, underscoring that psychologists must develop their own intelligence video games and delineate their key features for the measurement devices of next generation.

Received 22 April 2016; Revised 11 October 2016; Accepted 20 October 2016

Keywords: abilities, commercial video games, intelligence, video games.

“Overreliance on conventional testing has greatly limited modern research on intelligence”. (Hunt, 2011, p. 24)

Decades ago, scientists discussed about intelligence assessment in the XXI century (Detterman, 1979; Horn, 1979; Hunt & Pellegrino, 1985; Resnick, 1979) and the main conclusions can be summarized around three points: use of computers, adaptive testing, and simulation of everyday problem solving situations.

Computers have been used to develop tasks for testing cognitive processes such as working memory, attention, processing speed, or visual search (Lavie, 2005; Miyake, Friedman, Quiroga et al., 2011; Rettinger, Shah, & Hegarty, 2001; Santacreu, Shih, & Quiroga, 2011; Wilding, Munir, & Cornish, 2001), but intelligence is still mainly measured by paper and pencil tests. This happens even when these printed-type intelligence tests can be administered using computers without modifying their main psychometric properties (Arce-Ferrer & Martínez-Guzmán, 2009). In this regard, Rubio and Santacreu (2003) elaborated an adaptive computerized intelligence test that has scarcely been sold and even retired from the publisher’s catalog. Thus, for intelligence, computerized assessment has not been generalized.

Adaptive testing enables reduced testing time, adjustment between difficulty levels and examinees’ ability, and improvement in the obtained ability estimates (ETS, Pearson & CollegeBoard, 2010; Weiss & Kingsbury, 1984).

When adaptive testing supports computerized tasks, the evaluation allows individualized administration, response time registering, immediate feedback, or Speed Accuracy Trade Off –SATO– corrections.

Simulation of everyday problem solving situations has been used in different occupational areas (Gray, 2002) such as flight training (Hays, Jacobs, Prince, & Salas, 1992), teamwork, leadership and planning (Szumal, 2000), education (Lunce, 2006) or medicine (Byrne, Hilton, & Lunn, 2007). Some of these simulations take advantage of video game environments. Sitzmann (2011) has shown that studies using active job training procedures based on video games produce better indexes of declarative knowledge, procedural learning, and retention rate.

However, neither of the above did devise the revolution that video consoles and video games have involved for the general population. Nobody anticipated that these changes would be implemented into games, and even that a new domain would appear: ‘gamification’. This topic refers to the application of game mechanics and design techniques to engage and motivate people to attain their goals at work, in the school or everyday life settings (Burke, 2014).

Video games

The first video game created in the 80s for analyzing cognitive processes was *Space Fortress* (Mané & Donchin, 1989). From the eyes of a video gamer from the XXI century it looks as an arcade game (http://hyunkyulee.github.io/research_sf.html), very simple in terms of

Correspondence concerning this article should be addressed to M. A. Quiroga. Facultad de Psicología. Campus de Somosaguas. UCM. 28223. Madrid (Spain).
E-mail: maquirog@ucm.es

graphic design and in 2D. Using this video game, Rabbitt, Banerji, and Szymanski (1989) analyzed the correlations between intelligence and video game performance through practice (five successive days). For measuring intelligence the AH4 was administered (Heim, 1967). The obtained correlations from 56 participants ranging from 18 to 36 years old were: .283 for the first session, .422 with the slope learning function and .680 with maximum scores. Rabbitt et al. (1989) concluded: *"a relatively unsophisticated video-game, on which performance may reasonably be expected to be independent of native language or acquired literacy, and which is greatly enjoyed by young people who play it, rank orders individual differences in 'intelligence' nearly as well as pencil and paper psychometric tests which have been specially developed for this purpose over the last 80 years"* (p. 13).

Since this pioneer study, none was devoted to this topic until the XXI century. Recently, researchers have analyzed custom games to measure abilities (McPherson & Burns, 2007; 2008; Ventura, Shute, Wright, & Zhao, 2013) or casual video games to test the cognitive abilities they might be tapping (Baniqued et al., 2013). Twenty years after, scientists are recovering the topic, probably because people have incorporated video games into their lives and, also, developing video games is now affordable.

In Europe data from the quarter 3 2015 by GameTrack (Interactive Software Federation of Europe/IpsosMediaCT, 2015) show an average percentage of population playing any type of game from 40 in UK to 62 in France (42% for Spain), for 6 to 8 hours per week. There are differences by age, but not interaction age by sex. Among the youngest (6 to 15 years old) 85% play, on average, any type of game. These figures decrease to 55% for 15 to 34 years old, and for the oldest group (45 to 64 yrs.) it is an 18%. Furthermore, people prefer consoles, computers and smartphones to handhelds or tablets. For USA citizens during 2014, data from the ESA (Entertainment Software Association, 2014) show that 59% of the population plays video games. People of all ages play video games: 29% under 18, 32% from 18 to 35 and 39% over 36 years old. No differences by sex were observed: 52% males and 48% females. 51% of USA households have a dedicated console. And more noticeable, figures increase yearly, in both Europe and USA.

One relevant question is: are video games something more than a way for investing free time? Granic, Lobel, and Engels (2014) have summarized four groups of benefits of playing video games: cognitive, motivational, emotional and social. The mentioned cognitive benefits are: training spatial abilities, changing neural processes and efficiency, as well as developing problem-solving skills and creativity enhancement. Importantly, if challenging enough, video games cannot be automated

and, therefore, they could be used for testing purposes (Quiroga et al., 2011). The same is true for the repeated administration of an intelligence test: if you are not able to deduce the more complicated rules the test includes, your performance will not improve.

Video games for measuring intelligence

Quiroga, Colom et al. (2009), Quiroga, Herranz et al. (2009) study was similar to Rabbitt et al.'s (1989) study. The focus was to analyze the effects of repeated playing to elucidate if practice leads to learning how to play and, ultimately, automatization. Three games from Big Brain Academy by Nintendo® were considered. Participants played five times each game in each session (1 session per week; two weeks between sessions). Obtained results showed that all participants improved their skill (d 's from .72 to 1.64) but more importantly, correlations between g (factor score obtained from a battery of tests) and video game performance through practice varied for the three games. One of the games showed increased correlations with intelligence (from .49 to .71), leading to the conclusion that some video games are not automated (consistent with Rabbitt et al., 1989).

However, perhaps some games require a higher amount of practice to be automated. For checking this issue, Quiroga et al. (2011) increased substantially the number of practice sessions from 2 to 5 weeks, from 10 blocks of ten items to 25 blocks of ten items. Results showed that even with this very intensive practice, some video games couldn't be automated (correlations between intelligence and performance remained stable across playing sessions: from .61 to .67) while others clearly show a different pattern (initial correlations were .61, but they decreased through practice to a value of .32). The fit between the expected and obtained patterns of correlations was very high: .65 and .81, respectively. As noted by Rabbitt et al. (1989) *"high correlations between test scores and game performance may occur because people who can master most of the wide range of problems included in IQ tests can also more rapidly learn to master complex systems of rules, to attend selectively to the critical portions of complex scenarios, to make rapid and correct predictions of imminent events and to prioritize and update information in working memory"* (p. 12).

Therefore, some video games seem to require intelligence each time they are played, but which abilities tap the different video games? Even more important, is intelligence measured by paper and pencil tests the same, at the latent level, than intelligence measured by video games?

Using 20 casual games and 14 tests measuring fluid intelligence, spatial reasoning, perceptual speed, episodic memory and vocabulary, Baniqued et al. (2013)

found that games categorized as measures of working memory and reasoning were highly correlated with fluid intelligence tests ($r = .65$). Quiroga et al. (2015) administered 10 tests measuring five abilities (Gf, Gc, Gv, Gy, and Gs) and twelve console games (representing four abilities' domains: memory, visualization, analysis, and computation). Results showed extremely high correlations between the latent factor for video games and for the g factor ($r = .93$).

Therefore, two different types of video games, two different intelligence tests' batteries, but the same conclusion: intelligence can be measured with commercial video games. The hypothesis by Hunt and Pellegrino (1985) regarding the influence of the device used for presenting the items (paper and pencil or computer) fails to substantiate.

The study by Quiroga et al. (2015) included 12 games, and 10 of them came from *Big Brain Academy*® for Nintendo Wii®. Perhaps the high correlation obtained between the two latent factors (g and Video Game Performance) is due to the fact that the so-called "brain games" have been elaborated paralleling paper and pencil tests or lab tasks. Here we will test if intelligence can be measured with a commercial video game designed just for entertainment (leisure game), and if so, if we can create an Intelligence Test based on the game.

STUDY 1

Method

Participants

Participants were recruited from the Faculty of Psychology at the *Universidad Complutense de Madrid* (UCM) and from *Colegio Universitario Cardenal Cisneros* (CUCC) also located in Madrid, through flyers advertising to participate in a video game study. 55 people applied to participate, but 47 completed the whole experiment (38 women and 9 men). The mean age was 19.6 ($SD = 1.65$, range 18–25). Participants had no previous experience with the Nintendo DS or with the video game.

All participants signed an informed consent form, and accepted to avoid playing at home the video game they were playing in the Lab, during the 6 weeks required for completing the experiment. Upon study completion, participants received an exemplar of the video game played as gratification for their participation.

Procedure

Abilities' tests and *Video Games Habits Scale* (VGHS; Quiroga et al., 2011) were group administered in the first session and demographic data were collected. Afterwards, each participant agreed with the researchers the days per week (from Monday to Thursday) he/she

will come to the Lab for playing one hour with the video game. Participants should complete 15 hours of playing, during 6 weeks at a maximum rate of 4 hours per week.

Researchers did provide consoles and video games to each participant, and saved the game each time after completing an hour of playing, on a flash memory card (Secure Digital Card or SD). Consoles and video games remained in the lab during the 6 weeks.

Materials

Abilities tests administered were the Abstract Reasoning (AR), the Verbal Reasoning (VR) and the Spatial Reasoning (SR) subtests from the Differential Aptitudes Test Battery, level 2 (Bennett, Seashore, & Wesman, 1990; DAT-5 Spanish adaptation by TEA in 2000). Internal consistency values for AR and SR were excellent (.83 and .90) and adequate for VR (.78).

The selected video game was "*Professor Layton and the curious village*"® for Nintendo DS®, because it had been released only some months before starting our study (2009), and, therefore, it was unknown. This is a puzzle adventure game based on a series of puzzles and mysteries given by the citizens of towns that *Professor Hershel Layton*® and *Luke Triton*® visit. Some puzzles are mandatory, but it is not necessary to solve all the puzzles to progress through the game. However, at certain points in the story guiding the game, a minimum number of puzzles must be solved before the story can continue.

In 2015, *Professor Layton* has become a series consisting of six games and a film. By April 2015 the series has sold more than 15 million copies worldwide. From each participant, several outcomes from the video game performance were saved, from the many the video game provides: (1) number of puzzles found per hour of playing; (2) number of puzzles solved per hour of playing; (3) number of points (games' *picarats*) per hour and time needed to complete the video game if less than the maximum allowed of 15 hours.

The "*Video Games Habits Scale*", was developed by Quiroga et al. (2011). It consists of 16 questions covering the whole aspects of the playing experience: amount of time spent playing (hours per week), types of devices used, and type of video games played. For this study only answers to the first 6 questions were considered.

Results

First of all, ability scores were factor analyzed (principal axis factoring, PAF) to obtain a g score for each participant that was transformed into the IQ scale ($M = 100$, $SD = 15$). Factor loadings were .92 for AR, .78 for SR and .76 for VR. The percentage of explained common variance was 68%. Table 1 includes descriptive statistics

Table 1. Descriptive statistics for abilities tests and video game outcomes ($N = 47$)

	<i>M</i>	<i>SD</i>	$Z_{\text{asymmetry}}$	Z_{kurtosis}
DAT-AR	23.02	6.98	0.09	-1.41
DAT-VR	22.04	6.28	-0.16	-1.51
DAT-SR	25.89	9.42	-0.08	-1.35
IQ	100	13	-0.30	-1.45
Invested time	14.09	1.23	-3.23	0.67
Puzzles found	101.94	12.82	-4.39	11.62
Solved puzzles	91.28	15.80	-0.70	2.64
Found puzzles per hour ¹	7.78	1.40	-1.06	0.13
Solved puzzles per hour ¹	6.91	1.51	0.20	0.57

Note: ¹These variables refer only to the first 11 hours of playing.

for the ability measures, computed IQs and video game outcomes. All ability measures showed proper distributions (standard asymmetry and kurtosis values < 2.00) while some video game outcomes did not. Invested time and number of found puzzles showed a high negative asymmetry, meaning that a high percentage of participants needed almost the maximum time allowed to complete the game and also a high percentage of participants found a great number of puzzles.

Table 2 shows descriptive statistics per hour for the video game outcomes. Video game outcomes provided by the console are aggregated. This is important because disaggregated measures are meaningless due to the fact that not all parts of the game include the same number of puzzles.

Table 2. Descriptive Statistics

Hour (N)	Puzzles found		Puzzles Solved		Points (Picarats)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	8.32	2.28	7.74	2.27	181.34	70.77
2	15.68	4.63	14.36	4.77	352.15	146.12
3	23.34	5.93	20.94	6.18	540.79	187.72
4	31.38	7.67	27.89	7.52	729.23	224.32
5	39.60	9.32	34.87	9.27	911.83	267.43
6	48.40	11.71	42.60	11.24	1111.02	325.49
7	55.94	13.86	49.26	13.18	1300.28	403.24
8	63.96	14.56	56.40	14.48	1506.36	459.98
9	72.32	15.29	63.66	15.77	1729.21	533.01
10	78.89	15.45	69.45	16.48	1914.55	582.59
11	85.57	15.35	75.98	16.59	2142.02	612.11
12 (46)	91.53	14.85	81.49	17.11	2348.51	660.00
13 (42)	96.55	14.77	86.49	17.80	2546.06	719.02
14 (31)	100.11	14.23	90.11	17.77	2687.11	715.41
15 (26)	104.02	14.22	94.62	18.16	2861.28	742.63

Note: Sample size decreases from the 11th hour, due to participants that had already completed the video game.

Figure 1 includes the correlations between IQ scores and video game performance (puzzles found and puzzles solved). The pattern of correlations shows an increasing value from the first hour to the 11th, both for found and for solved puzzles. The difference between correlation values is statistically significant in both cases ($Z_{\text{found puzz.}} = 3.51 > Z_c = 1.96$; $Z_{\text{solved puzz.}} = 3.65 > Z_c = 1.96$) from hour 1 ($r_{g\text{-found puzz.}} = .073$; $r_{g\text{-solved puzz.}} = .198$) to hour 11 ($r_{g\text{-found puzz.}} = .657$; $r_{g\text{-solved puzz.}} = .586$). After the 11th hour, correlations decrease, probably due to the sample size reduction caused by the fact that some participants have completed the video game. The correlation between *g* scores and invested time to complete the video game was $r = -.522$ ($p < .001$). Interestingly, participants with high and low IQ scores (over 100 and below or equal 100) clearly differed in the time required for completing the video game ($M_{\text{High IQ}} = 13.58$, $SD = 1.24$; $M_{\text{Low IQ}} = 14.61$, $SD = .99$; $F(1, 45) = 9.78$, $p = .003$; $d = -.92$). Participants with high IQ scores required, on average, one hour less to complete the game. Importantly, the 95% Confidence Interval showed no overlap between groups (High IQ: 13.06 to 14.11; Low IQ: 14.18 to 15.04).

Because of the high relationship observed between intelligence and video game performance through practice, difficulty (*P*) and discrimination (*D*) indices were computed for each puzzle to select the best ones to elaborate a test. *P* index represents the proportion of examinees correctly answering the item. *D* index represents the fraction of examinees from the upper group correctly answering the item minus the fraction of examinees from the lower group correctly

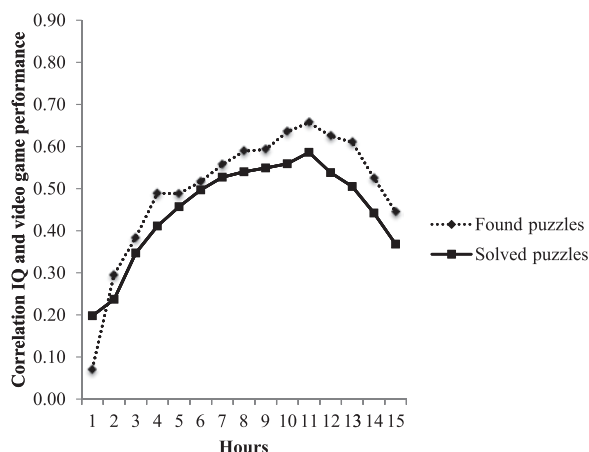


Figure 1. Correlations between g and video game performance through practice.

answering the item. To compute these indices, the formulae provided by Salkind (1999) were used. The maximum value of D is 1.0 and is obtained when all members from the upper group (27% with the highest scores) succeed and all members from the lower group (27% with the lowest scores) fail on an item. Table 3 includes these indices for the 41 puzzles with P and D values $> .40$. Only 36 participants had completed data for these puzzles. With these 41 puzzles, a test and three equivalent versions were elaborated (VG-T1k = 14 puzzles; VG-T2k = 14 puzzles and VG-T3k = 13 puzzles) similar in P and D indices ($F(2, 40) = 1.12, p = .337$ and $F(2, 40) = 1.794, p = .180$, respectively). The obtained reliabilities (internal consistency) for the whole puzzles test and the three equivalent versions were: .937, .823, .806 and .840.

Afterwards, a parallel analysis was computed with the three ability scores and the three video game tests scores to determine the number of factors that best represent the shared variance. Table 4 includes the obtained results showing a different solution whether the mean or the 95th percentile is used as criterion. With the mean, 2 factors could be a good solution because the eigenvalue is higher than the one simulated with the mean. With the 95th percentile, results should be grouped in one single factor. Thus two Exploratory Factor Analyses were run. Oblimin rotation was used for the two factors solution.

Table 5 includes both factorial solutions. Explained variance for the one factor solution is 54.17%, whereas explained variance for the two factors solution is 76.21% (57.30% for the first factor and 18.83 for the second); the correlation between the two factors is .478. Note that factorial outcomes from small samples are unstable, and, therefore, replication is mandatory (Thompson, 2004).

Discussion

Results obtained in this first study show that video game performance, measured either with Found Puzzles or Solved Puzzles, shows an increased correlation with intelligence through practice. This suggests that playing requires the systematic use of participants' cognitive resources, as observed in previous studies (Quiroga et al., 2011; Rabbitt et al., 1989). Moreover, the medium-high and negative correlation between g scores and invested time to complete the game indicates that intelligence predicts the time required for completing the video game.

On the other hand, Difficulty and Discrimination indices have been used to select the best puzzles. Puzzles with both indices higher than .40 were selected and forty-one puzzles (34% of those included in the video game) passed the criterion. The resulting test showed a high reliability (.94).

Finally, exploratory factor analysis results suggested that video games and paper and pencil can be described either as measures of the same intelligence factor (all factor loadings $> .50$), or as two correlated intelligence factors ($r = .48$).

STUDY 2

Method

Participants

Participants were recruited from the Faculty of Psychology at the UCM, through flyers advertising to participate in a video game study. 45 people applied to participate but 27 had free time to complete the whole experiment (21 women and 6 men). Mean age was 20.56 ($SD = 2.72$, range 18–28). Participants from this study were older than those of the first study ($d = -.43$). Selected participants had no previous experience with the video game "Professor Layton and the Curious village"®, but many were familiarized with the Nintendo DS and the Layton series video games. Nowadays it is almost impossible to find young adults without video consoles experience. The sample in this study does not differ from the sample in the first study regarding sex, but they are slightly older and with more experience playing video games.

All participants signed an informed consent form and accepted to avoid playing at home during the 6 weeks required for completing the experiment. To reward participation, upon study completion, participants took part in a raffle of 9 video games as the one they had played.

Procedure

Exactly the same as in study one.

Table 3. Selected puzzles from “Professor Layton and the curious village” with Difficulty (P) and Discrimination (D) indices $\geq .40$

Puzzle number	Name	Maximum Number of Points (Picarats)	P	D	Test to which the item was ascribed
10	Four digits	10	0.8	0.5	1
49	1000 Times	20	0.7	0.7	3
26	Bottle full of germs	20	0.7	0.4	1
55	The odd sandwich	20	0.8	0.5	3
80	Too many Queens 1	20	0.8	0.5	1
43	Three umbrellas	20	0.8	0.4	1
54	Monster!	20	0.8	0.4	2
110	The vanishing cube	20	0.8	0.4	2
104	A sweat treat	30	0.5	0.6	3
105	Rolling a three	30	0.5	0.6	1
19	Parking lot gridlock	30	0.7	0.5	3
117	Painting a cube	30	0.7	0.5	1
32	Candy jars	30	0.8	0.5	2
52	Find a star	30	0.8	0.5	1
84	Which boxes to move?	30	0.8	0.5	2
96	On the stairs	30	0.8	0.5	2
79	Apples to oranges	40	0.6	0.9	2
103	Wood cutouts	40	0.6	0.7	2
73	How many squares?	40	0.6	0.6	3
12	Make a rectangle	40	0.7	0.5	2
81	Too many queens 2	40	0.7	0.7	2
102	Aces and the joker	40	0.7	0.5	1
61	Pin board shapes	40	0.7	0.4	2
67	How many sweets?	40	0.8	0.5	1
72	Truth and lies	40	0.8	0.5	2
62	A tricky inheritance	40	0.8	0.4	3
75	The wire cube	40	0.8	0.4	1
119	Red and blue 1	40	0.8	0.4	2
59	The longest path	50	0.7	0.4	1
42	The camera and case	50	0.8	0.5	3
44	Stamp stumper	50	0.8	0.5	2
112	My beloved	50	0.8	0.4	3
82	Too many queens 3	60	0.5	0.7	3
97	Princess in a box	60	0.7	0.7	3
78	Water pitchers	60	0.7	0.6	2
95	A magic square	60	0.8	0.5	1
100	Seven squares	70	0.6	0.9	3
98	Card order	70	0.6	0.8	1
99	33333!	70	0.6	0.8	2
94	Get the ball out! 4	70	0.8	0.5	3
83	Too many queens 4	80	0.5	0.7	1

Note: Puzzles have been ordered by maximum number of “picarats” (points) that can be obtained in each.

Materials

The same ability tests were administered, but only odd items because of time restrictions. Nevertheless, reliability coefficients showed acceptable values: AR = .71; VR = .80 and SR = .88.

Results

Table 6 includes descriptive statistics for ability measures, computed IQs and video game outcomes.

Except for the high kurtosis shown by DAT-SR and number of Solved Puzzles, the remaining variables show proper distributions (standard asymmetry and kurtosis values < 2.0). The leptokurtic distributions for DAT-SR and number of Solved Puzzles indicate that more than expected participants grouped around mean values. Ability scores were factor analyzed (principal axis factoring, PAF) to obtain a g score for each participant that was transformed into the IQ scale.

Table 4. Parallel analysis for the three ability tests and the three video game tests

Eigenvalues	Empirical	Mean	Percentile
1	3.646	1.593	1.859
2	1.383	1.284	1.455
3	0.457	1.062	1.193
4	0.296	0.871	0.995
5	0.147	0.691	0.819
6	0.072	0.498	0.643

Table 5. Factor solutions for one and two factors with the ability and video game tests

(a)

	Factor 1
Test 2 VG	.884
Test 1 VG	.871
Test 3 VG	.823
DAT-VR	.652
DAT-AR	.581
DAT-SR	.520

(b)

	Factor 1	Factor 2
Test 2 VG	.967	.466
Test 3 VG	.920	.416
Test 1 VG	.910	.470
DAT-AR	.370	.938
DAT-VR	.147	.729
DAT-SR	.335	.710

Note: KMO = .727; Bartlett = 158.74; $p < .001$.

Table 6. Descriptive statistics for abilities tests and video game outcomes (N = 20)

	M	SD	Z asymmetry	Z kurtosis
DAT-AR	11.70	3.04	0.18	-1.11
DAT-VR	13.50	3.39	-0.30	-1.30
DAT-SR	12.55	3.84	1.49	3.03
IQ	101	11.14	1.83	-0.03
Invested time	11.00	1.89	0.05	0.30
Puzzles found	98.65	8.14	1.16	2.06
Solved puzzles	88.85	10.96	1.96	3.43
Found puzzles per hour ¹	10.45	2.29	0.52	-0.61
Solved puzzles per hour ¹	9.19	2.54	0.89	0.01

Note: ¹these variables only refer to the first 7 hours of playing.

Table 7 shows descriptive statistics per hour for video games outcomes and the number of participants having completed the game hourly. In this second study participants required less time for completing the game than in the first study.

Figure 2 includes the correlations between IQ scores and video game performance (puzzles found and puzzles solved). The pattern of correlations shows an increasing value from the first hour to the 7th, for both found and solved puzzles. The difference between correlation values is statistically significant only for solved puzzles ($Z_{\text{found puzz.}} = 1.43 > Z_c = 1.96$; $Z_{\text{solved puzz.}} = 2.77 > Z_c = 1.96$) from hour 1 ($r_{\text{g-found puzz.}} = .388$; $r_{\text{g-solved puzz.}} = .308$) to hour 7 ($r_{\text{g-found puzz.}} = .514$; $r_{\text{g-solved puzz.}} = .564$). After the 7th hour, correlations decrease probably due to the sample size reduction caused by the fact that some participants have completed the video game. The correlation between IQ scores and invested time to complete the video game was $-.539$ ($p < .01$). Participants with high and low IQ scores (over 100 and below or equal 100) clearly differed in the time required for completing the video game ($M_{\text{High IQ}} = 9.36$, $SD = 1.62$; $M_{\text{Low IQ}} = 11.80$, $SD = 1.74$; $F(1, 45) = 13.113$, $p = .001$; $d = -1.45$). High IQ scorers required, on average, two hours less to complete the game. Importantly, the 95% Confidence Interval showed no overlap between groups (High IQ: 8.27 to 10.46; Low IQ: 10.84 to 12.76).

For each participant, scores from the puzzles test, as well as from the three equivalent versions constructed in the first study, were computed. Two Exploratory Factor Analyses (EFA), were run. Oblimin rotation was used for the two factors' solution. Table 8 includes both factor solutions. Explained common variance for the one factor solution was 48.51%. Explained common variance for the two factors solution was 69.97% (51.65 % for the first factor and 18.32 for the second); the correlation between the two factors was $r = .41$.

Discussion

Results observed in this second study replicate the findings of the first study. Therefore, seven years after the first study, and with a different group of participants more familiarized with video games, the same results were obtained: (a) video game performance shows an increasing correlation with IQ through practice; (b) intelligence predicts the amount of time required for completing the video game; in this second study individuals with higher IQ scores required, on average, two hour less to complete the video game; (c) video games and paper and pencil can be described either as measures of the same intelligence factor (factor loadings $> .45$) or as two intelligence correlated factors ($r = .41$).

Table 7. Descriptive Statistics

Hour (N)	Puzzles found		Puzzles Solved		Points (Picarats)	
	M	SD	M	SD	M	SD
1	9.15	3.10	8.55	3.05	207.40	90.25
2	18.45	6.63	16.85	6.67	436.15	193.06
3	28.50	8.68	25.55	8.83	684.80	247.04
4	40.50	12.81	36.05	12.44	951.90	351.61
5	52.25	14.31	45.95	14.92	1229.15	468.17
6	62.90	15.05	55.15	15.68	1501.50	530.09
7	73.15	16.06	64.35	17.77	1787.55	634.33
8 (19)	82.50	13.94	72.70	16.35	2070.85	636.30
9 (19)	89.50	12.30	78.85	16.20	2282.00	657.83
10 (15)	94.70	10.14	83.50	14.58	2471.00	620.89
11 (14)	97.05	8.69	86.85	12.14	2614.85	534.59
12 (9)	98.00	8.33	88.10	11.59	2685.10	492.30
13 (3)	98.60	8.15	88.70	11.07	2719.65	459.25
14 (2)	98.65	8.14	88.80	10.00	2725.60	453.37
15 (1)	98.65	8.14	88.85	10.96	2728.40	452.18

Note: Sample size decreases from the 7th hour, due to participants that had already completed the video game.

Comparison between studies

Participants from both studies were compared on the abilities measured and on the video game outcomes. 6 univariate ANOVAs were run. Table 9 summarizes the results. Group variances did not differ, except for the total points obtained in the video game. However, both groups obtained a similar average number of points. Groups differ in the number of puzzles found per hour ($d = -.43$) and in the number of puzzles solved per hour ($d = -.82$). In both instances, participants from the second study found and solved more puzzles. This suggests higher speed in the participants from the second study as a result of their previous experience with Layton series video games.

Table 10 includes the correlations between IQ, ability measures, the three parallel tests made with

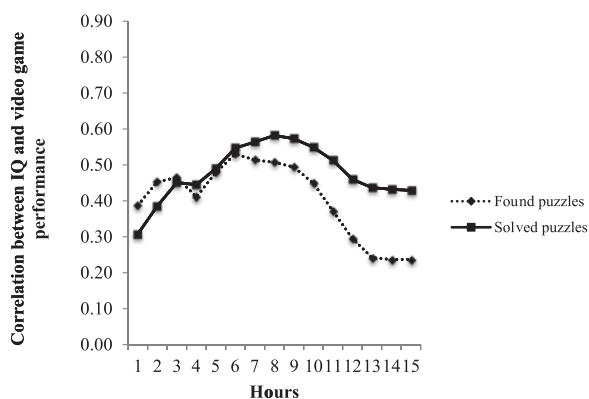


Figure 2. Correlations between g scores and video game performance through practice.

the video game and the whole test. These correlations were computed for 56 participants (36 from study 1 and 20 from study 2). The magnitude of these correlations ($.50 < r_{xy} > .35$) is similar to most of the convergent validity coefficients found for ability and intelligence tests (Pearson TalentLens, 2007).

Table 8. Factor solutions for one and two factors with the ability and video game tests (Second Study)

(a)		
	Factor 1	
Test 2 VG	.939	
Test 1 VG	.762	
Test 3 VG	.649	
DAT-VR	.448	
DAT-AR	.730	
DAT-SR	.541	
(b)		
	Factor 1	Factor 2
Test 2 VG	.959	.530
Test 3 VG	.766	.265
Test 1 VG	.888	.345
DAT-AR	.506	.857
DAT-VR	.242	.603
DAT-SR	.278	.862

Note: KMO = .645; Bartlett = 66.8; $p < .001$.

Table 9. IQ and video game outcomes per Study

	2009		2015		F_{Levene}	F	d
	M	SD	M	SD			
IQ	100	13.24	100	13.94	.006	.000	.00
Found puzzles	101.94	12.82	99.77	8.49	.969	.596	.20
Solved puzzles	91.28	15.80	85.26	19.85	.404	2.058	.34
Found puzzles per hour ¹	7.99	1.98	10.70	2.29	1.517	28.058***	-.43
Solved puzzles per hour ¹	7.04	1.88	9.08	2.95	3.723	13.312***	-.82
Points	2734	658.79	2715	415.58	4.161*	.018	.03

Note: ¹these variables only refer to the first 7 hours of playing. * $p < .05$; *** $p < .001$.

Table 10. Correlations between intelligence and ability measures, and tests constructed with the video game

	Test 1 Video Game ($k = 14$)	Test 2 Video Game ($k = 14$)	Test 3 video Game ($k = 13$)	Video Game Test ($k = 41$)
DAT-AR	.376**	.421**	.321*	.394**
DAT-VR	.313*	.327*	.291*	.328*
DAT-SR	.332*	.331*	.260	.326*
IQ	.434**	.450**	.361**	.439**

Note: * $p < .05$; ** $p < .01$.

General Discussion

Video games performance and intelligence assessment with a leisure video game

The first conclusion derived from the reported studies is this: Video game performance, measured either with Found Puzzles or Solved Puzzles, shows an increased correlation with intelligence through practice, showing that the selected video game requires the systematic use of participants' cognitive resources, as observed by Rabbitt et al. (1989) and Quiroga et al. (2011).

Second, time invested for completing the video game correlates with intelligence differences in both studies (-0.52 and -0.54). Specifically, we have shown that high intelligence participants complete the game from one to two hours before than low intelligence participants ($d = -0.92$ and $d = -1.45$). This converges with Ackerman's (1988) studies showing that time to learn a new task shows inverse correlations with intelligence.

Third, the most discriminative 41 puzzles' test, along with the three equivalent versions created, showed satisfactory reliability ($\alpha = .937$, $\alpha = .823$, $\alpha = .806$ and $\alpha = .840$, respectively), compared with the usual range of accepted values (Tavakol & Dennick, 2011; from .70 to .95).

Fourth, EFA analyses showed that the one or two factor solutions account for a medium to high percentage of shared common variance. The one factor solution underscores that the constructed video game tests could be good measures of a g factor (loadings from .82 to .88 in the first study, and from .65 to .94

in the second study). The two-factor solution shows that the latent factor measured with the video game tests is correlated with the g factor in both studies (.48 and .41). These values depart from those reported by Quiroga et al. (2015) for brain games, but are close to those obtained from Baniqued et al. (2013) for casual games.

Fifth, participants from both studies were similar in their IQ scores as well as in the obtained outcomes for the video game (found puzzles, solved puzzles and points) but they clearly differed in the efficiency measures computed for video game performance: found puzzles per hour and solved puzzles per hour ($d = -0.43$ and $d = -0.82$). Participants from the second study are more efficient, finding and solving almost 30% more puzzles per hour. This result was surprising; nothing from the participants' characteristics led us to expect this huge difference in the speed for completing the task. This might be a consequence of having played video games since childhood. Indeed, for the first study (run in 2009) it was easy to find naïve participants, but it was impossible for the second study (run in 2015). Actually, Boot (2015) underscored the need of more precise measures about video game playing experience, including the history of game play across their lifetime. Results reported here support his demand.

Sixth, for the whole sample (participants from both studies that completed the 41 puzzles, $N = 56$) the correlations between IQ, ability measures and video

games tests are in the medium range ($.46 < r_{xy} > .30$) showing proper convergence. Note that previous studies correlating video games performance and ability measures obtained similar correlations values: (1) McPherson and Burns (2008) using two specifically designed games to measure processing speed and working memory (*Space Code* and *Space Matrix*) obtained correlations ranging from .21 to .54 between games performance and tests' scores (Digit Symbol, Visual Matching, Decision Speed, Picture Swaps and a short form of the Advanced Progressive Matrices); (2) Shute, Ventura, and Ke (2015) used the commercial video game Portal 2 and obtained correlations ranging from .27 to .38 between video game performance and tests' scores (*Verbal Insight Test*, *Mental Rotation Test*, *Spatial Orientation Test* and *Visual Spatial Navigation Assessment*); (3) Baniqued et al. (2013) used 20 casual games to obtain 5 components that correlated with the 5 latent factors obtained from 14 intelligence and ability tests; the obtained correlations ranged from .18 to .65. In this latter study, all game components correlated highly with the fluid intelligence latent factor (from .27 for visuo-motor speed games to .65 for the working memory games); and (4) Buford and O'Leary (2015) using the *Puzzle creator* from Portal 2 developed a video game test that correlates .46 with the Raven Standard Progressive Matrices.

In summary, here we have shown that intelligence can be assessed with existing commercial video games. Also, these video games' tests converge with ability tests and IQ scores at the same level than ability tests do among themselves. However, an important question remains to be answered: is it feasible to use commercial video games for testing intelligence or abilities? Currently, our answer is negative because: (1) available video games are less efficient given that they require a lot of time for obtaining the same reliability and validity coefficients obtained with paper and pencil tests; (2) researchers lack control regarding stimuli and dependent variables; (3) outcomes are not saved in any type of data set, requiring inefficient data registration, and only small samples can be considered.

However, in exceptional cases, commercial video games could provide adequate estimations of cognitive abilities. For example, for applicants to high-level jobs (managers, etc.) completing a video game with no time limits is a novel and unexpected situation that can provide good estimations and valuable information about the ability to solve problems across an extended period of time.

What about the influence of previous experience playing video games? Could the assessment of intelligence be biased if video games are used? In this regard, Foroughi, Serraino, Parasuraman, and Boehm-Davis (2016) have shown that when the video games measure Gf, previous experience does not matter, so having

applicants familiar or unfamiliar with video games in general wouldn't be a problem for the assessment of those groups.

We finish this report providing answers to the next crucial question: is there any future for video games as tools for assessing intelligence?

Future research on intelligence measured with video games

Thirty years ago, concluding remarks from Hunt and Pellegrino (1985) about the use of computerized tests in the next future distinguished economic and psychological reasons. Economic considerations underlie the use of computerized tests if the main reasons for their use are easy administration (psychologists are free to observe participants' behavior because the computer saves performance), or greater accuracy (saving not only response, but also response time and type of response, from which to compute the derived scores needed). These economic reasons explain why thirty years later from this forecast we still lack a generalized use of computerized assessments of the intelligence factor. It has been very expensive to develop computerized tests even when their practical advantages have been recognized. However, commercial video games are beginning to include "mods" (abbreviation for modifications) for free allowing researchers to design his tests, as Buford and O'Leary (2015) or Foroughi et al. (2016) have done with Portal 2 to test fluid intelligence. Nevertheless, even being free it is not easy to master the mod; multidisciplinary group are strongly needed.

Psychological issues arise when the new tool is intended to provide a better measure of intelligence than paper and pencil tests (Hunt & Pellegrino, 1985). Could video games be better measures of intelligence and abilities? Recent reviews (Boot, 2015; Landers, 2015) suggest that testing with video games is more motivating, because they include an engaging narrative and usually consider new challenges or elements as players complete levels in the game. Engaging narratives are good because players solving problems or puzzles want to know more about the story. In other words, the narrative has to include key aspects for the game; otherwise, players will play the game without reading the story (as an example *Cradle of Empires*®, from Awem Studio®, includes a narrative not very engaging that allow players to avoid reading the story). Another important feature of recent video games is that motivation is managed through the essentials of the psychology of learning (*Candy Crush*®, now property of Activision Blizzard®, is an excellent example of how to use this discipline for designing a very addictive video game, see for example Hopson, 2001; Margalit, 2015).

Using video games for testing intelligence and cognition could also reduce test anxiety in those cases where assessment is perceived as a threatening situation. However, we think the most important point is that a video game can include the rules and components Primi (2014) has designed to create an assessment tool criterion-reference based, to overcome the question of arbitrary metrics on the normative based test scores (Abad, Quiroga, & Colom, 2016). Thus, the next step for measuring intelligence is to develop tests that include an objective scale metric relying on cognitive complexity and essential processes underlying intelligence. Video Games could be the tool for fulfilling this goal. As noted by Primi (2014) “the scale proposed that the levels of fluid intelligence range from the ability to solve problems containing a limited number of bits of information with obvious relationships through the ability to solve problems that involve abstract relationships under conditions that are confounded with an information overload and distraction by mixed noise” (p. 775).

This goal can be certainly accomplished in the next future with video games, but available commercial video games are far from these two points, even new generation video games such as *Portal 2*® (Valve Corporation, 2011) or *Witness*® (Tekla Inc., 2016; <http://the-witness.net/news/>). *Witness*®, a 3D puzzle video game just released on January 2016, includes neither instructions nor time limits. The player has to deduce what to do to complete the game, it is a continuous visuo-spatial reasoning test, but the player has to explore and solve all the difficulty levels to complete the game. Thus, it is not an adaptive test and this point should be compulsory for the new tools. The same is true for *Portal 2*®. We need adaptive video games including mods for psychologists to develop their own tests. Also, it is crucial for video games to provide accuracy and speed scores, because this will allow computing efficiency measures. Recently, we have reviewed more than 40 games looking for ones saving both criteria, accuracy and speed, and it has been difficult to find more than 5. The most common situation is to have a game providing separate accuracy and speed scores. Even worst, a mixture expressed as “obtained points”, probably computed with an algorithm that remains unknown for users.

To conclude, video games might be the basis for the next generation of assessment tools for intelligence and abilities. They are T data in terms of Cattell’s (1979) classification, and, therefore, are objective, and can be completed even without the supervision of a psychologist. Remote assessments are possible. Games are engaging, motivating and attractive, because people like to play. But they must be also psychometrically as sound as classic tests. In this sense, video games research requires validity studies (Landers, 2015;

Landers & Bauer, 2015). This issue is still on its infancy but progress is made. Psychologists have to develop their own video games and contribute to the design of commercial video games.

Video games for assessment should include (1) an engaging story from which many levels could be derived, with and without distraction; (2) items of different complexity implemented in an adaptive way for testing each participant with the small number of items required to better discriminate his abilities profile; these items should differ in the number and difficulty of the rules needed to solve them; (3) measures of accuracy and speed from which efficiency measures could be derived; accuracy and speed must be automatically saved for each assessed intelligence process; (4) hierarchical levels of the essential processes that underlie intelligence (working memory and processing speed) that match the different developmental levels; (5) video games must be designed for avoiding automation through practice (psychomotor abilities should not be essential for video game outcomes) and (6) no time limit for the whole video game, but speed modules.

Finally, research studies on existing video games must follow the gold standard for psychological research as recently summarized by R. N. Landers addressing video games’ programmers: “Rigorous experimental designs, large sample sizes, a multifaceted approach to validation, and in-depth statistical analyses should be the standard, not the exception” (Landers, 2015; p. iv).

References

- Abad F., Quiroga M. A., & Colom R. (2016). Intelligence assessment. In *Encyclopedia of applied psychology. Online reference database titled Neuroscience and biobehavioral psychology*. Oxford, UK: Elsevier Ltd.
- Ackerman P. J. (1988). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Sernberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and practice* (pp. 165–217). New York, NY: W.H. Freeman and Company.
- Arce-Ferrer A. J., & Martínez-Guzmán E. (2009). Studying the equivalence of computer-delivered and paper-based administrations of the raven standard progressive matrices test. *Educational and Psychological Measurement*, 69, 855–867. <http://dx.doi.org/10.1177/0013164409332219>
- Baniqued P. L., Lee H., Voss M. W., Basak C., Cosman J. D., DeSouza S., ... Kramer A. F. (2013). Selling points: What cognitive abilities are tapped by casual video games? *Acta Psychologica*, 142, 74–86. <http://dx.doi.org/10.1016/j.actpsy.2012.11.009>
- Bennett G., Seashore H., & Wessman A. (1990). *DAT-5. Test de aptitudes diferenciales. Manual*. Madrid, Spain: TEA.
- Boot W. R. (2015). Video games as tools to achieve insight into cognitive processes. *Frontiers in Psychology*, 6, 1–2. <http://dx.doi.org/10.3389/fpsyg.2015.00003>

- Buford C. C., & O'Leary B. J.** (2015). Assessment of fluid intelligence utilizing a computer simulated game. *International Journal of Gaming and Computer-Mediated Simulations*, 7, 1–17. <http://dx.doi.org/10.4018/IJGMS.2015100101>
- Burke B.** (2014). *Gartner redefines gamification*. Stamford, CT: Gartner, Inc. Retrieved from http://blogs.gartner.com/brian_burke/2014/04/04/gartner-redefines-gamification/
- Byrne A. J., Hilton P. J., & Lunn J. N.** (2007). Basic simulations for anaesthetists. A pilot study of the ACCESS system. *Anaesthesia*, 49, 376–381. <http://dx.doi.org/10.1111/j.1365-2044.1994.tb03466.x>
- Cattell R. B.** (1979). Adolescent age trends in primary personality factors measured in T-data: A contribution to use of standardized measures in practice. *Journal of Adolescence*, 2(1), 1–16. [http://dx.doi.org/10.1016/S0140-1971\(79\)80002-0](http://dx.doi.org/10.1016/S0140-1971(79)80002-0)
- Detterman D.** (1979). A job half-done: The road to intelligence testing in the year 2000. *Intelligence*, 3, 295–306. [http://dx.doi.org/10.1016/0160-2896\(79\)90024-2](http://dx.doi.org/10.1016/0160-2896(79)90024-2)
- Entertainment Software Association** (2014). *Essential facts about the computer and video game industry*. Washington, DC: Author. Retrieved from http://www.theesa.com/wp-content/uploads/2014/10/ESA_EF_2014.pdf
- ETS, Pearson & CollegeBoard.** (2010). *Some considerations related to the use of adaptive testing for the common core assessments*. Boulder, CO: Author.
- Foroughi C. K., Serraino C., Parasuraman R., & Boehm-Davis D. A.** (2016). Can we create a measure of fluid intelligence using puzzle creator with Portal 2? *Intelligence*, 56, 58–64. <http://dx.doi.org/10.1016/j.intell.2016.02.011>
- Granic I., Lobel A., & Engels R. C. M. E.** (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66–78. <http://dx.doi.org/10.1037/a0034857>
- Gray W. D.** (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly*, 2, 205–227.
- Hays R. T., Jacobs J. W., Prince C., & Salas E.** (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology*, 4, 63–74. http://dx.doi.org/10.1207/s15327876mp0402_1
- Heim A. W.** (1967). *AH4 group test of intelligence*. London, UK: National Foundation for Educational Research.
- Hopson J.** (2001). *Behavioral game design*. New York, NY: Gamasutra. Retrieved from http://www.gamasutra.com/view/feature/131494/behavioral_game_design.php
- Horn J. L.** (1979). Trends in the measurement of intelligence. *Intelligence*, 3, 229–239. [http://dx.doi.org/10.1016/0160-2896\(79\)90019-9](http://dx.doi.org/10.1016/0160-2896(79)90019-9)
- Hunt E. B.** (2011). *Human intelligence* (pp. 31–63). New York, NY: Cambridge University Press.
- Hunt E., & Pellegrino J.** (1985). Using interactive computing to expand intelligence testing: A critique and prospectus. *Intelligence*, 9, 207–236. [http://dx.doi.org/10.1016/0160-2896\(85\)90025-X](http://dx.doi.org/10.1016/0160-2896(85)90025-X)
- Interactive Software Federation of Europe/ IpsosMediaCT** (2015). *Game track quarterly digest*. Brussels, Belgium: ISFE. Retrieved from <http://www.isfe.eu/industry-facts/statistics>
- Landers R. N.** (2015). Guest editorial preface. Special issue on assessing human capabilities in video games and simulations. *International Journal of Gaming and Computer-Mediated Simulations*, 7, 4–8.
- Landers R. N., & Bauer K. N.** (2015). Quantitative methods and analyses for the study of players and their behaviour. In P. Lankowski & S. Bjork (Eds.), *Research methods in game studies* (pp. 151–173). Pittsburg, PA: ETC Press. Retrieved from http://press.etc.cmu.edu/files/Game-Research-Methods_Lankowski-Bjork-et-al-web.pdf
- Lavie N.** (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Science*, 9, 75–82. <http://dx.doi.org/10.1016/j.tics.2004.12.004>
- Lunce L. M.** (2006). Simulations: Bringing the benefits of situated learning to the traditional classroom. *Journal of Applied Educational Technology*, 3(1), 37–45.
- Mané A., & Donchin E.** (1989). The space fortress game. *Acta Psychologica*, 71, 17–22.
- Margalit L.** (2015). *Why are the Candy crushes of the world dominating our lives?* New York, NY: Psychology Today. Retrieved from <https://www.psychologytoday.com/blog/behind-online-behavior/201508/why-are-the-candy-crushes-the-world-dominating-our-lives>
- McPherson J., & Burns N. R.** (2007). Gs invaders: Assessing a computer game-like test of processing speed. *Behavior Research Methods*, 39, 876–883. <http://dx.doi.org/10.3758/BF03192982>
- McPherson J., & Burns N. R.** (2008). Assessing the validity of computer-game-like tests of processing speed and working memory. *Behavior Research Methods*, 40, 969–981. <http://dx.doi.org/10.3758/BRM.40.4.969>
- Miyake A., Friedman N. P., Rettinger D. A., Shah P., & Hegarty M.** (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, 130, 621–640. <http://dx.doi.org/10.1037/0096-3445.130.4.621>
- Pearson TalentLens** (2007). *Core abilities assessment. Evidence of reliability and validity*. Retrieved from: https://www.talentlens.co.uk/assets/legacy-documents/36917/caa_evidence_of_reliability_and_validity.pdf
- Primi R.** (2014). Developing a fluid intelligence scale through a combination of rasch modeling and cognitive psychology. *Psychological Assessment*, 26, 774–788. <http://dx.doi.org/10.1037/a0036712>
- Quiroga M. A., Colom R., Privado J., Román F. J., Catalán A., Rodríguez H., ... & Ruiz J.** (2009, December). Video games performance and general intelligence. *Poster presented at the Xth Annual ISIR Conference*, Madrid, Spain.
- Quiroga M. A., Escorial S., Román F. J., Morillo D., Jarabo A., Privado J., ... Colom R.** (2015). Can we reliably measure the general factor of intelligence (g) through commercial video games? Yes, we can! *Intelligence*, 53, 1–7. <http://dx.doi.org/10.1016/j.intell.2015.08.004>
- Quiroga M. A., Herranz M., Gómez-Abad M., Kebir M., Ruiz J., & Colom R.** (2009). Video-games: Do they require general intelligence? *Computers & Education*, 53, 414–418. <http://dx.doi.org/10.1016/j.compedu.2009.02.017>
- Quiroga M. A., Román F. J., Catalán A., Rodríguez H., Ruiz J., Herranz M., ... Colom R.** (2011). Videogame performance

- (not always) requires intelligence. *International Journal of Online Pedagogy and Course Design*, 1, 18–32. <http://dx.doi.org/10.4018/ijopcd.2011070102>
- Rabbitt P., Banerji N., & Szymanski A.** (1989). Space fortress as an IQ test? Predictions of learning and of practised performance in a complex interactive video game. *Acta Psychologica*, 71, 243–257.
- Resnick L. B.** (1979). The future of IQ testing. *Intelligence*, 3, 241–253.
- Rubio V. J., & Santacreu J.** (2003). *TRASI. Test adaptativo informatizado para la evaluación del razonamiento secuencial y la inducción como factores de la habilidad intelectual general*. [TRASI: Computerized adaptive test for the assessment of sequential reasoning and induction as factors of general mental ability]. Madrid, Spain: TEA.
- Salkind N. J.** (1999). *Métodos de investigación* [Exploring research]. México, México: Prentice Hall.
- Santacreu J., Shih P. Ch., & Quiroga M. A.** (2011). *DiViSA. Test de Discriminación simple de árboles* [DiViSA. Simple visual Discrimination Test of trees]. Madrid, Spain: TEA Ediciones.
- Shute V. J., Ventura M., & Ke F.** (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers and Education*, 80, 58–67. <http://dx.doi.org/10.1016/j.compedu.2014.08.013>
- Sitzmann T.** (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, 64, 489–528. <http://dx.doi.org/10.1111/j.1744-6570.2011.01190.x>
- Szumal S.** (2000). How to use problem-solving simulations to improve knowledge, skills and teamwork. In M. Silberman & P. Philips (Eds.), *The 2000 team and organizational development sourcebook*. New York, NY: McGraw Hill.
- Tavakol M., & Dennick R.** (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <http://dx.doi.org/10.5116/ijme.4dfb.8dfd>
- Tekla Inc.** (2016). *Witness*. Lexington Park, MD: Author. Retrieved from <http://the-witness.net/news/>
- Thompson B.** (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Valve Corporation** (2011). *Portal 2*. Bellevue, WA: Author. Retrieved from <http://www.valvesoftware.com/games/portal2.html>
- Ventura M., Shute V. J., Wright T., & Zhao W.** (2013). An investigation of the validity of the virtual spatial navigation assessment. *Frontiers in Psychology*, 4, 852. <http://dx.doi.org/10.3389/fpsyg.2013.00852>
- Weiss D. J., & Kingsbury G. G.** (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21, 361–375. <http://dx.doi.org/10.1111/j.1745-3984.1984.tb01040.x>
- Wilding J., Munir F., & Cornish K.** (2001). The nature of attentional differences between groups of children differentiated by teacher ratings of attention and hyperactivity. *British Journal of Psychology*, 92, 357–371. <http://dx.doi.org/10.1348/000712601162239>