Journal of Articles in Support of the Null Hypothesis
Vol. 9, No. 1

www.jasnh.com

Is Exposure to Visual Media Related to Cognitive Ability? Testing Neisser's Hypothesis for the Flynn Effect

Matthew J. Sigal
Stuart J. McKelvie
Bishop's University

This research was designed to investigate Neisser's (1997) hypothesis that the Flynn effect (rising intelligence scores during the 20th century) is related to increased exposure to the complex visual environment. In Study 1, 35 participants completed a test of mental rotation along with a survey of their exposure to various media activities. Test performance was positively related to 3-D video game exposure but not to the other predictors. In Study 2, 172 participants completed the Culture Fair Intelligence Test (CFIT) and the Wide Range Vocabulary Test (WRVT) along with another media survey. Test performance, particularly on the CFIT, was not related to any of the visual media variables. These nonsignificant results are inconsistent with Neisser's explanation of the Flynn effect.

Documenting the Flynn Effect

One of the most intriguing discoveries from research on intelligence is that there has been a rise in test scores during the 20th century (the "Flynn effect"). First reported by Flynn (1984), from whom the effect takes its name, and then discussed further by him at various times (e.g., Flynn, 1987, 1998a, 1998b, 1999), this trend has been found over an 80-year period in age groups from children to adults and in both industrialized and developing countries (Daley, Whaley, Sigman, Espinosa, & Neumann, 2003; Khaleefa, Abdelwahid, & Lynn, 2008; Lynn, 2009; Mingroni, 2007; Neisser, Boodoo, Bouchard, Boykin, Brody, & Ceci, et al., 1996; Nettlebeck & Wilson, 2004; Wicherts, Dolan, Hessen, Oosterveld, van Baal, Boosma, & Span, 2004).

Despite recent evidence that the increase in scores has been slowing or may even have ceased in certain European countries (Lynn, 2009; Mingroni, 2007; Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2005, 2008), the general trend remains robust. However, although the rise in scores has been found with various intelligence tests [e.g., Wechsler Adult Intelligence Scale (WAIS); Flynn, 1984], it is most clearly pronounced with tests of fluid intelligence (Colom, Andres-Pueyo, & Juan-Espinosa, 1998; Daley, et al., 2003; Flynn, 1998b; Greenfield, 1998; Lynn, 2009; Nettlebeck & Wilson, 2004; Wicherts et al., 2004). The effect has also appeared on tests of crystallized intelligence (Pietschnig, Voracek, & Formann, 2010), but is usually weaker or absent altogether (Colom, et al., 1998; Daley, et al., 2003; Flynn, 1998b; Lynn, 2009; Neisser et al., 1996; Nettlebeck & Wilson, 2004; Wicherts et al., 2004). This point is summarized nicely by the fact that gains are greater on the performance IQ than on the verbal IQ of the WAIS (Flynn, Shaughnessay, & Fulgham, 2012).

The distinction between these two kinds of test was first made by Cattell (1940), and developed in further writings (Cattell, 1963; Cattell, 1971, pp. 98-99; Cattell & Butcher, 1968, p. 19; Cattell & Horn, 1978). Fluid intelligence is the ability to adapt to new situations, particularly where relationships must be extracted without the use of a previously-stored solution. In contrast, crystallized intelligence involves judgments that have been encountered in the past. This knowledge, which often consists of accumulated facts, is acquired in part through the mechanism of fluid intelligence and can be further applied in the future, perhaps by verbatim recall. The best-known measure of fluid intelligence is Cattell's own Culture Fair Intelligence Test (CFIT) (Cattell, 1940; Cattell & Cattell, 1972; IPAT, 1973), although Raven's Progressive Matrices (RPM) is also recognized (Gregory, 2011, p. 167). Both tests are nonverbal and consist of visual diagrams from which patterns and relationships must be extracted. Notably, many studies have shown that the Flynn effect occurs with the Raven, on which scores have risen strongly over the years (Daley, et al., 2003; Lynn, 2009; Mingroni, 2007; Neisser et al., 1996). Although the recent levelling off (Sundet et al., 2004; Teasdale & Owen 2005, 2008) occurred with matrix-like items, the response format was not the same as on the standard RPM. It is surprising that the traditional measure of fluid intelligence (CFIT) has not figured more prominently in research on the Flynn effect, because it may be a more complete measure of fluid intelligence than the Raven (Cattell & Horn, 1978; Colom & Garcia-Lopez, 2003). One exception is a report from Spain showing the classic increase on the CFIT (Colom & Garcia-Lopez, 2003).

Issues Concerning the Flynn Effect

Although the increase in IQ scores has been widely documented, there is debate about whether the increase reflects a real change in intelligence itself or whether it only reflects a test-related activity (Flynn, 1984, 1987; Gregory, 2011, p. 270; Lynn, 2009; Neisser, 1997; Wicherts et al., 2004). Flynn (1984) himself regards the increase in scores as "semi-real." He argues that intelligence tests, particularly the Raven, only measure a weak correlate of actual intelligence, so that although test scores have been rising, intelligence itself may not (Flynn, 1987).

Of course, this still leaves open the question of what causes the Flynn effect. Intelligence may be highly heritable, but this does not mean that group differences (in this case, between generations) are highly heritable (Neisser et al., 1996). Indeed, most writers agree with Flynn (1984) that his effect is environmental rather than genetic (e.g., Mingroni, 2007; Neisser, 1997; Nettlebeck & Wilson, 2004). However, researchers have also proposed intermediate accounts, in which the increase in IQ scores reflects a reciprocal relationship between environment and genes (Dickens & Flynn, 2001) or an epigenetic relationship between improved nutrition and the fullest expression of genetic potential (Greiffenstein, 2011).

The specific cause or causes of the Flynn effect remain unclear, but there is little doubt that environment plays an important role, with as many as 10 specific suggestions on offer (Flynn, 1984; Mingroni, 2007; Neisser, 1997). Besides nutrition, examples include medical care, smaller families, child-rearing practices, education, and practice effects, which include test sophistication.

Neisser's Hypothesis

One fact that these hypotheses must explain is the frequently-reported different trend for tests of fluid and crystallized intelligence. Given that tests of fluid intelligence are nonverbal and, more specifically, visual, an interesting proposal is the idea that the rise in intelligence test scores may be due to increased exposure to visual media of various kinds (Neisser, 1997), collectively characterized by Neisser (1997, p. 446) as the "visual and technical environment."

Corresponding to the rise in intelligence test scores during the 20th century, Neisser (1997) argues that there has been a markedly increased exposure to many forms of visual media, which he identifies as pictures of various kinds (e.g., photographs, decorations on the wall, messages on cereal boxes and on plastic cups from fast-food restaurants), movies, television, puzzles, video games and computers. Neisser states that we view some of these media passively, but that we also view others actively in the sense that we analyze them, with consequences for other mental functions. According to Neisser (1997, p. 447), "exposure to complex visual media has produced genuine increases in a significant form of intelligence", which he terms "visual analysis". Moreover, Neisser claims that visual analysis is measured by tests like the Raven (RPM), on which the largest gains have been found. A similar position has been expressed by Johnson (2003, 2005), who argues that popular television programming has become more cognitively demanding over the years. Greenfield (1998) has also developed an argument like Neisser's, but in more detail for each medium. From her analysis of film, television, video games, and computers, she concludes that they emphasize “iconic imagery” over verbal processes. That is, representations tend to be analogue rather than symbolic or digital. Moreover, nonverbal intelligence tests, on which the largest Flynn effects have occurred are also iconic in nature. In particular, she observes that both 3-D action and 2-D puzzle video games require the exercise of visuo-spatial skills such as mental manipulation and visual perspective shifting, and that these to skills are employed on nonverbal tests. Indeed, it was noted above that the Raven, on which gains have been great, consists of visual diagrams from which relationships must be extracted. Finally, Flynn himself has suggested that increased leisure activities involving reading and puzzle games and modern devices such as television, computers and VCRs place strong demands on our cognitive capacities (Dickens & Flynn, 2001).

Research on Visual Media and Visual Cognition

The core of Neisser’s argument is that exposure visual media has provided, by unintentional proxy, training in visual analysis and that this training has led to the Flynn effect. This argument as two parts. The first is that exposure to complex visual media is positively associated with visual cognitive ability. The second is that the increase in intelligence test scores over the years, particularly for nonverbal intelligence, has been accompanied by a corresponding increase in exposure to visual media. The present paper is concerned with the first step in this argument.

Studies of spatial cognition. Neisser’s broad question connecting a wide variety of visual media to intelligence does not seem to have been empirically investigated. However, there is evidence of links between exposure to specific kinds of visual material and performance on visual cognitive tasks along the lines suggested by Greenfield (1998), and she argued that the cognitive processes stimulated by visual media are also required on nonverbal intelligence tests. If it can be established that there is a relationship between visual media and basic cognitive processes, this would provide a basis for the broader link to nonverbal tests.

Indeed, Spence and Feng (2010) review studies of the relationship between playing video games and spatial cognition. Some of them have treated video game playing as a subject variable, comparing people who have had different amounts of exposure to video games. Studies of this kind have demonstrated positive relationships between video game playing and visual cognition. This approach captures the natural variations in game playing experience that constitute part of the visual environment that Neisser identifies as important for the Flynn effect but, as Spence and Feng point out, it cannot establish cause and effect. However, other studies, in which video game playing was treated as an experimental variable, have shown that people who have been trained to play the games perform better on tests of visual cognition than control groups (Feng, Spence, & Pratt 2007). For example, Subrahmanyam and Greenfield (1994) found that performance on two computerized tests of dynamic spatial skills (requiring judgments of objects in motion) improved after playing a dynamic video game but did not improve after playing a non-spatial word game.

In addition to this methodological distinction, studies have varied in the kind of spatial or visual skill that serves as the criterion test (dependent variable). For example, Dorval and Pepin (1986) investigated the effect of video game training on performance on the spatial relations subtest from the Differential Aptitudes Test (DAT). Here, participants must inspect a two-dimensional figure then mentally fold it to decide which three-dimensional figure is created. In Dorval and Pepin’s experiment, participants who had practice with a video game (Zaxxon) improved their scores on this subtest more than a control group that did not practice. Notably, in another investigation of the Flynn effect, a small gain in scores over time has been reported on this DAT subtest, although it was not as great as scores on the DAT abstract reasoning subtest, which measures fluid intelligence (Colom et al., 1998). However, Spence and Fang observe that, whereas studies like Dorval and Pepin's support a causal connection between video game playing and complex visual cognition, they do not demonstrate why the effect occurs. That is, they do not address the question of whether exposure to video games alters the fundamental sensory, attentional and memory processes that underlie performance on the more global cognitive spatial tasks. Fortunately, another set of studies has tackled these questions. For example, in what Spence and Feng identify as the most important early study of this issue, Green and Bavelier (2003; cited in Spence & Feng, 2010; see also Green & Bavelier, 2006, 2007) found that there was a positive effect of action video game practice on visual selective attention.

Studies of near transfer and far transfer. Although some of Green and Bavelier's (2003) results have not been replicated, suggesting a null relationship between playing video games and aspects of visual attention (Murphy & Spencer, 2009), these considerations bring Spence and Feng to a third issue: the extent to which exposure to video games has effects that generalize from performance on tasks that closely resemble these games (near transfer) to those that are less similar (far transfer). This issue is also relevant to the central issue of this paper: if transfer can be shown here, will transfer occur further to performance on nonverbal intelligence tests?

For example, in one study demonstrating far transfer with healthy senior citizens, training in an arcade-type video game produced increases on both the verbal and performance sections of the Wechsler Adult Intelligence Scale (Drew & Waters, 1986). However, in another study with training on a complex video game (Space Fortress; Boot, Basak, Erickson, Neider, Simons, Fabiani et al., 2010), there was a positive effect on criterion tests that resembled the game (near transfer) but not on other tests (far transfer). In particular, performance on the Raven's Progressive Matrices (RPM) was unaffected (see also Boot, Kramer, Simons, Fabiani, & Gratton 2008).

Consistent with Boot et al.'s (2010) experimental result on near and far transfer, Sims and Mayer (2002) conducted a non-experimental study comparing people who were skilled or unskilled in a video game (*Tetris*) that requires mental rotation. They found that the skilled group performed better on a task that also required mental rotation of shapes (near transfer), but did not perform better on other visual tasks (far transfer), even if they required rotation. Solidifying this relationship between playing *Tetris* and mental rotation, Okagaki and Frensch (1994) manipulated training with *Tetris* and showed that mental rotation times for shapes decreased for their experimental group but not the control group. Not surprisingly, they also found that performance on another visualization task that resembled the *Tetris* game improved more in the experimental group than the control group. More recently, Boot et al. (2008) confirmed that training with *Tetris* decreased mental rotation times, although it did not improve accuracy. However, in a second part of their study, neither mental rotation time nor accuracy were better for expert gamers compared to non-gamers. Nevertheless, in their review of this literature, Spence and Feng (2010) conclude that most studies show robust effects on near transfer but little or no effects on far transfer. The key determinant of transfer is the degree to which the basic processes affected by the game are also required to perform the criterion task. This also pertains to our central question, for which the issue is whether there are common processes between the visual media, in particular video games, and nonverbal intelligence tests.

Many of these issues are clearly illustrated by an investigation in which Feng and Spence themselves were authors (Feng, Spence, & Pratt, 2007). They investigated the relationship between action video game exposure and performance on both a basic visual task and on a more global cognitive visual task. In their first study, which was quasi experimental, they found that action video game players performed better than nonplayers on a task of spatial attention. In their second study, which was experimental, Feng et al. found that people who were given 10 hours of training on a 3-D first-person shooter game (*Medal of Honor: Pacific Assault*) performed better on the spatial attention task than a control group which played a 3-D puzzle game (Ballance). However, they also found that the trained participants were superior on a more complex visual mental rotation task in which they had to mentally manipulate three-dimensional objects on a two-dimensional perceptual plane. These stimuli were very similar to those constructed by Shepherd and Metzler (1971) in the original classic study of visual mental rotation but, of particular interest here, the requirement to imagine the stimuli rotated in three dimensions is similar to the cognitive process that is tapped by the spatial relations subtest of the DAT. As noted above, scores on this test have also been related to video game playing and have shown a Flynn effect. Similarly, for the central issue in this paper, an important question is whether the processes involved in mental rotation might also pertain to tests of nonverbal intelligence that are not simply assessing spatial ability.

Feng et al.'s finding that playing a video game affects mental rotation is consistent with the demonstration that direct training on mental rotation can improve performance on that task (Tarr & Pinker, 1989). However, whereas the latter result demonstrates near transfer, Feng et al.'s result demonstrates far transfer. In addition, Feng et al.'s result is similar to the findings of another experiment showing transfer of training on mental rotation. In this case, Widenbauer, Schmid, and Jansen-Osmann (2007) found that performance was improved by a computer simulation-training program. Far transfer was also seen in a study by Quaiser-Phol, Geiser, and Lehmann (2006) who found that boys who were classified as 3-D (action and simulation) computer game players performed better on mental rotation of Shepherd and Metzler-like shapes than 2-D (logic and skill training) game players or nonplayers. Unlike the study by Feng et al. (2007), this investigation was not experimental because people were classified as 3-D or 2-D players. However, the subject variable of video game experience, like Feng et al.'s (2007) experimental manipulation of video game training, is more relevant than direct mental rotation training or computer simulation training to the present issue of exposure to visual media. Together, the studies of Feng et al. and Quaiser-Phol et al. provide support for the idea that natural exposure to video games may alter the basic visual processes that lead to enhanced performance on a complex cognitive visual task. This lends credibility to Neisser's (1997) visual analysis hypothesis.

The Present Research

As just observed, superior mental rotation performance has been associated with 3-D video game preference (Quaiser et al.) and 3-D game training (Feng et al.). The first purpose of the present research was to conduct a small-scale study to supplement these results with another nonexperimental examination of the relationship between natural video game experience and mental rotation performance. Like Quaiser et al., information was obtained about experience with 3-D video games and 2-D video games, but additional information was also gathered about exposure to other visual media along the lines suggested by Neisser (1997). This is important to the present approach to the Flynn effect, because Neisser's hypothesis links performance on fluid intelligence to a variety of visual media. Based on the results of Quaiser et al. and of Feng et al., it was predicted that mental rotation performance would be positively related to 3-D video game experience. In addition, the study provided a test of Neisser's suggestion that natural exposure to the modern technical environment improves the skill of visual analysis. If this is true, it would be expected that mental rotation performance would also be positively related to experience with other visual media.

Study 1

Participants

Thirty-five Canadian undergraduate students received course credit for their participation. There were 18 males and 17 females, with a mean age of 21.9 years (*SD* = 3.09; range 17 to 32).

Materials

The materials consisted of a paper-and-pencil media exposure survey and a computerized mental rotation task based upon Shepard's original stimuli (Shepard & Mezler, 1971; Metzler & Shepard, 1974).

The survey began with questions for sex and age. They were followed by a question about reading and eight media exposure questions that were designed to capture Neisser's (1997) description of visual media. They were presented in the following fixed order with an example to illustrate seven of them: reading outside of class, television or movie watching, side-scrolling videogames (e.g., the *Super Mario Bros.* or *Contra*), sports-related video games (e.g., soccer, hockey or fighting simulations), driving or simulation video games (e.g., *Burnout* or *Microsoft Flight Simulator*), logic or skill-training video games (e.g., on-line *Sudoko*, *TextTwist* or *Tetris*), traditional role-playing video games (*Final Fantasy* or *Dragon Quest*), action/adventure videogames (e.g., *Quake*, *Myst* or *World of Warcraft*), and crosswords or other logic puzzles (e.g., paper-and-pencil *Sudokos*). The 3-D video game categories were sports-related, driving/simulation, and action/adventure; while the primarily 2-D categories consisted of side-scrolling games, fantasy role-playing games and logic/skill training puzzles.

For each question, participants were asked to consider the past year and to choose how often, on average, they engaged in the activity, choosing from five frequency options: once or more per day, 2-6 times per week, 1-7 times per month, less than once a month and never. This was followed by an open-ended question about the number of minutes for a typical session. A score was obtained for each category by coding the frequency categories from 5 (once or more per day) to 1 (never), by coding the session duration into number of hours (with appropriate decimals) and then by multiplying the two numbers.

The dependent variable was the score on the mental rotation task that was designed by John Krantz from Hanover College and is freely available for use from his web site. (http://psych.hanover.edu/JavaTest/CLE/Cognition/Cognition.html) The test consisted of Metzler and Shepard's (1974) original three-dimensional stimuli, with variables preset by the author.

Participants compared two three-dimensional objects and, through keyboard input, stated if they were the same image or mirror images. No special instructions were given apart from telling participants to choose the correct answer. A single test session included the presentation of the stimuli from 10 different angles (0°, 20°, 40°,…, 160°, 180°), with each angle given 10 trials, yielding 100 assessments in total. Two measures were taken: proportion correct (accuracy) and response time (in seconds). The dependent variables used in the analysis were accuracy and weighted accuracy (accuracy divided by average response time).

Procedure

Participants were tested individually. They completed a consent form, the media exposure survey and then the mental rotation task, which was administered on an Apple 12" G4 Titanium PowerBook at a resolution of 1024 by 768 pixels. Sessions typically lasted between 15 and 25 minutes followed by debriefing.

Results

Pearson product-moment correlation coefficients between each predictor and both dependent variables were calculated, after which multiple regression analyses were conducted. Tabachnick and Fidell (1996, pp. 146-156; see also http://www.palgrave.com/pdfs/0333734718.pdf) distinguish three kinds of multiple regression: standard, where all predictors are entered simultaneously, hierarchical, where order is specified by the researcher, and statistical (stepwise), where statistical considerations guide the manner in which predictors are considered. Because we had a clear prediction for one relationship (3-D video game exposure and mental rotation), a hierarchical regression analysis was performed. Here, exposure to 3-D video games was entered at the first step, followed by the three other visual media variables (2-D video games, TV/movies and puzzles) at the second step, and then finally reading at the third step. However, a stepwise regression analysis was also conducted, because it is thought to lead to the "best" (Tabachnick & Fidell, p. 150) or "most parsimonious" (http://www.palgrave.com/pdfs/0333734718.pdf) prediction equation. Here, all four predictor variables were entered together. Alpha was set at .05.

Accuracy

Descriptive statistics and Pearson product-moment correlation coefficients are shown in Table 1. For accuracy (proportion correct), one was significant. Mental rotation performance was positively related to 3-D video game exposure.

In the hierarchical regression analysis (see Table 2), R^2 at each step was .194, .261, and .261, the first of which was significant. Although the second one approached significance, neither of the changes in R^2 was significant. In the first step, 3-D video games was significant, $\beta = .440$, and it remained significant by itself in the second and third steps ($\beta = .556$, $\beta = .554$, respectively). In the stepwise regression analysis, which is presented here in full (not in Table 2), again only one predictor was significant in the model: 3-D video games, $R^2 = .194$, $F(1, 33) = 7.94$, $p = .008$, $\beta = .440$.

Table 1: *Descriptive Statistics and Correlations (r) for Mental Rotation Performance and Predictors in Study 1*

Variable	*M*	*SD*	*r* with Accuracy	*p*	*r* with Weighted Accuracy	*p*
Accuracy	0.82	0.11				
Reaction time (RT)	5.03	2.45				
Weighted Accuracy	0.19	0.06				
Age	21.91	3.09				
3-D Weighted	1.86	1.84	.440	.008	.365	.031
2-D Weighted	1.57	1.40	.189	.276	.103	.557
TV/movie weighted	5.76	3.87	-.236	.173	.192	.270
Puzzle weighted	1.54	1.49	.023	.896	.092	.598
Reading weighted	3.60	2.81	.073	.679	.018	.918

Note: $n = 35$. Weighted accuracy = accuracy/RT

Table 2: *Hierarchical Regression Analyses for Prediction of Mental Rotation Performance in Study 1*

Model	R^2	*F*	df_1	df_2	*p*	R^2 Change	*F*	df_1	df_2	*p*	β	*t*	*p*
						Accuracy							
1	.194	7.94	1	33	.008	.194	7.94	1	33	.008	.440*	2.82	.008
2	.261	2.65	4	30	.053	.067	0.91	3	30	.449	.556*	2.73	.010
3	.261	2.05	5	29	.101	.000	0.002	1	29	.969	.554*	2.63	.014
						Weighted Accuracy							
1	.133	5.06	1	33	.031	.133	5.06	1	33	.031	.365*	2.25	.031
2	.199	1.87	4	30	.142	.066	0.83	3	30	.488	.448*	2.30	.028
3	.210	1.54	5	29	.208	.011	0.39	1	29	.559	.514	2.36	.025

Note: Predictors were 3-D video (Model 1), 3-D video, 2-D video, TV, Puzzles (Model 2), 3-D video, 2-D video, TV, puzzles, reading (Model 3); *significant predictor 3-D video.

Weighted Accuracy

For weighted accuracy (accuracy divided by average response time; see Table 1), there was one significant correlation coefficient. Once more, mental rotation performance was positively related to 3-D video game exposure. In the hierarchical regression analysis (see Table 2), R^2 at each step was .133, .199, and .210, only the first of which was significant. In addition, neither of the changes was significant. In the first step, 3-D video games was significant, β = .365, and it remained significant by itself in the second and third steps (β = .448, β = .514, respectively). In the stepwise regression analysis, which is presented here, the only significant predictor was, again, 3-D video games, R^2 = .133, $F(1, 33) = 5.06$, $p = .031$, β = .365.

Although sex and age data were gathered, neither variable was included as a predictor in the main regression analyses. With only 35 participants in total, the sample size was too small to justify additional predictors. In addition, for sex, the number of participants (n = 35 in total; 18 men, 17 women) was deemed to be insufficient for a robust comparison. For age, there was restriction of range because the participants were university students. However, when the multiple regression analyses were repeated with sex and age included, the results were the same: the only significant predictor was 3-D video game experience.

Discussion

The results of Study 1 are clear. Performance on the mental rotation task was positively related to reported experience with 3-D video games. This contrasts with Boot et al.'s (2008) finding that expert gamers and non-gamers did not differ in mental rotation time or accuracy. On the other hand, the present results replicate Quaiser-Phol, et al.'s (2006) finding that mental rotation was associated with 3-D game preference. These two nonexperimental results do not demonstrate a causal relationship. However, when they are taken in conjunction with Feng et al.'s (2007) demonstration that experimental training on a specific 3-D video game led to improved mental rotation performance, they suggest that 3-D video game experience has effects that generalize beyond the game, in this case to the ability to rotate shapes in three dimensions. In their seminal study of mental rotation, Shepard and Metzler (1971) found that time to respond to the rotated figure increased linearly with the amount of physical rotation required to answer the question. They argued that to accomplish the task, participants formed a visual image of the figure and rotated that image mentally. Perhaps there was a relationship between 3-D game experience and mental rotation performance because to play these games participants use imagery in an active manner, allowing them to transfer this skill to the mental rotation task.

In contrast, mental rotation performance was not related to amount of reading or to any of the other visual variables (2-D video games, TV/movies, puzzles) that were chosen to represent the visual environment as described by Neisser (1997). Although the dependent variable in Study 1 was mental rotation performance and not intelligence, this finding potentially weakens his case that exposure to the complex visual environment in general might be responsible for the Flynn effect. Perhaps the processes that are developed by exposure to the various kinds of visual media are not important for performance on nonverbal tests of fluid intelligence that demand the extraction of relationships from visual figures, contrary to Neisser's visual analysis hypothesis.

On the other hand, it is possible that a more specific version of Neisser's hypothesis might have validity. That is, active experience with the 3-D visual environment might be a factor in the Flynn effect. In particular, the significant relationship between 3-D video game exposure and mental rotation performance suggests that the increase in scores over time on the spatial relations subtest of the DAT (Colom et al., 1998) might be also be linked to this factor, because performance on this subtest seems to involve visual imagery. Divesta, Ingersoll and Sunshine (1971) found that spatial relations scores loaded on an imagery factor, and they suggest that this kind of mental paper folding is requires the subject to hold the percept in memory and rotate it. In addition, Tapley and Bryden (1977) found that accuracy on this DAT subtest was positively correlated on accuracy with Shepard and Metzler's figures.

At the same time, Colom et al. (1998) found that the Flynn effect was greater on the abstract reasoning test (measuring fluid intelligence) of the DAT than on the spatial relations test (measuring mental rotation ability). This implies that the relationship between playing 3-D video games and spatial relations (mental paper folding) might account for part of the Flynn effect. However, it would also have to be shown that there is a relationship between playing 3-D video games and performance on other tests. Colom et al.'s stronger effect on the more direct test of fluid intelligence is consistent with other research showing that the Flynn effect has occurred most strongly on tests of fluid intelligence, particularly the Raven, and not on other tests, particularly tests of crystallized intelligence. Therefore, the question remains whether exposure to the visual and technical environment, particularly the 3-D environment, develops the skill of "visual analysis" (Neisser, 1997) and whether this skill can be transferred to tests of fluid intelligence. The purpose of Study 2 was to investigate this possibility.

Study 2

It has been argued throughout this paper that for transfer to occur from one activity to another, there have to be processes that are common to the two activities. Given the evidence that exposure to 3-D video games is related to mental rotation, which is also involved in tests of spatial ability (e.g., on the DAT), the following question arises: are there common processes between tests of spatial ability and nonverbal tests of intelligence? In particular are there common processes between tests of spatial ability the Cattell Culture Fair Test of Intelligence (CFIT), which is the measure of fluid intelligence adopted in Study 2?

First, it should be reiterated that Cattell regarded his test as measuring Gf, general fluid intelligence (IPAT, 1973). Second, it has been empirically demonstrated from factor analysis that the g-loadings of the CFIT are commensurate in size (greater than .75) with the g-loadings from other test batteries (Johnson, te Nijenhuis, & Bouchard, 2008). Of most interest are the similar g-loadings (.48 to .58) for the four specific subtests of the CFIT and for other subtests that involve spatial manipulation (the Blocks test from the Twents Institute of Business Psychology test, the three-dimensional space test from the General Aptitude Test Batter, and the Figures test from the Groninger Intelligence Test). Third, it has been shown the mental folding tests and spatial ability tests are among the best measures of Gf (Lohman, 1993). Fourth, and most specifically, the correlation between scores on the general spatial ability marker (Gv) and scores on the CFIT is .6 Colom, Contreras, Botella, & Santacreu (2001). Together, these results support the contention that there common processes between tests of spatial ability involving mental rotation and the CFIT.

Effects of Training on Intelligence Test Performance

In the introduction, it was shown that mental rotation could be improved with direct instruction (close transfer; Tarr & Pinker, 1989) but also with experimental exposure to 3-D video games (far transfer; Feng et al., 2007). There is also evidence that scores on intelligence tests can be improved with direct instruction. As early as 1927, Gilmore showed that following a systematic coaching plan, subjects improved their scores on a general intelligence test (the Otis) more than others in a control group. In a meta-analytic review of coaching studies, Kulik, Bangert-Downs, and Kulik (1984) found that the average standard effect size for coaching was 0.33 on the Scholastic Aptitude Test which, despite its name, contains items that are similar to those on a general intelligence test. In fact, Frey and Detterman (2004) provide evidence that it is a measure of general intelligence ("g"). In addition, Kulik et al. found that the average effect size for coaching was 0.43 on other tests such as the Otis and the DAT. Of more interest here, Stankov (1986) conducted a large-scale study in which high school students were given exercises in creative problem-solving over a three-year period. He reported that effect sizes for treatment were fairly small (from 1 to 8 IQ points depending on the way that the data were analyzed), but were larger for tests of fluid intelligence than for tests of crystallized intelligence.

Denny and Heindrich (1990) point out that other studies have also demonstrated gains in fluid intelligence scores following training, particularly with older participants. For example, relative to a control group with no specific training, Baltes, Sowarka and Kliegel's (1989) participants, who received cognitive training on a figural relations test, performed better on another test of figural relations (direct training effect) and on a culture fair test (close transfer). In contrast, they did not perform better on other tests, including the Raven, that were less similar to the figural relations test (far transfer). However, given the importance of the Raven test in the Flynn effect, it is notable that Denny and Heindrich (1990) compared RPM performance for younger, middle-aged and older participants in a control group with others in an experimental group that was trained in a strategy for solving matrix problems. For all three age groups, there was a significant positive effect of training (near transfer). In addition, with matrix-type items, Freund and Holling (2011) found that scores on a re-test increased with practice (standardized effect size = 0.48) but increased more when training was given (effect size = 0.97) (direct transfer).

Together, these studies demonstrate that systematic training can improve scores on intelligence tests in general and on fluid intelligence tests that include the Raven in particular. These results are similar to the effects of direct training on mental rotation and show near transfer. However, the question here is whether or not there is far transfer between natural experience with visual media, specifically 3-D video games, and performance on a test of fluid intelligence.

Exposure to Media and Intelligence Test Performance

There appears to be only one study that has examined this question directly. Boot et al. (2008) compared expert gamers and non-gamers on the RPM. Performance did not differ between the two groups. However, in the same study, the performance of the two groups did not differ significantly on a number of basic attentional and visual tasks or on mental rotation. The lack of relationship with attentional tasks also been reported by others (Murphy & Spence, 2009), but Boot et al.'s results are inconsistent with other findings of significant positive relationships (e.g., Feng et al., 2007; Green & Bavelier, 2003). The authors themselves noted that small procedural variations among the studies may account for inconsistent results. Consequently, it remains an open question whether there is a transfer effect from experience with video games to performance on tests of fluid intelligence.

In addition, although the Raven test is the one on which the Flynn effect has, perhaps, been strongest and most reliable, it has been suggested that Cattell's CFIT is a better measure of fluid intelligence because it has a greater variety of items (Colom & Garcia-Lopez, 2003). It has a subtest with matrices like the RPM, but it also has three other nonverbal subtests. In addition, as noted earlier, the classic Flynn effect has appeared on the CFIT (Colom & Garcia-Lopez, 2003). Consequently, the CFIT was employed as the measure of fluid intelligence in Study 2. Moreover, for comparison purposes, a measure of crystallized intelligence was also included. In this case, a vocabulary test, the Wide Range Vocabulary Test (WRVT) was employed. Such tests are widely accepted as appropriate measures of crystallized intelligence (Lynn, 2009; Neisser, 1997). Indeed, in a factor-analytic study, Cattell (1971, p. 97) reports that vocabulary is the test with the highest loading on crystallized intelligence and the lowest loading on fluid intelligence.

The remainder of the methodology in Study 2 was similar to that of Study 1, with participants completing another survey tapping their exposure to various kinds of media. However, this survey also included a question about exposure to music. This widens the scope of the media being examined, but also provides a test of discriminative validity because it was assumed that simply listening to music (an auditory medium) would be unlikely to predict performance on fluid intelligence or on crystallized intelligence. Although it has been claimed that listening to Mozart increases spatial intelligence, the effect was immediate (Rauscher, Shaw, & Ky,1993). Moreover, given problems with replication, it may only occur under restricted conditions (Rideout, Dougherty, & Wernert, 1998) or even not at all (Chabris, 1999; Steele, Bass, & Crook, 1999). This contrasts with actively performing music, which does seem to increase IQ scores (Schellenberg, 2004). In addition, sex, age, and years of education were added as predictors. In the small scale Study 1, participants were too few to provide a valid comparison of men and women, and both age and education did not vary much because the participants were all undergraduates. In Study 2, numbers were greater and the general population was also sampled.

Hypotheses

First, from the results of Study 1, and from the theoretical analysis of common processes between mental folding and the CFIT, it was hypothesized that the most likely predictor of performance on the CFIT (fluid intelligence) would be experience with 3-D video games. However, given that this variable tapped visual imagery, it was not expected to predict performance on the WRVT (crystallized intelligence). Second, from Neisser's (1997) proposal that the Flynn effect, particularly on fluid intelligence, is due to exposure to the visual and technical environment, which increases the skill of visual analysis, it was predicted that performance on the CFIT, but not on the WRVT, would be related to experience not only with 3–D video games, but also with 2-D video games, TV/movie-watching, and with crosswords and other logic puzzles. Of course, if this result was obtained, it would not demonstrate the causal relationship envisaged by Neisser. However, the connection of performance on the CFIT to natural differences in exposure would be a first step towards testing this hypothesis. Of course, it is also possible that Neisser is mistaken in his conjecture. In particular, there may be insufficient overlap in the processes that are common to the visual and technical environment on the one hand and performance on the CFIT on the other. If this is the case, these relationships would not be significant.

As noted above, given that it tapped audition, exposure to music was not expected to be related to performance on either test. The necessary visual component for predicting CFIT scores is absent and the necessary verbal component for prediction of WRVT scores is absent. However, reading experience, in which words would be central, was expected to predict scores on the WRVT.

Study 2

Participants

One hundred and seventy-two Canadian participants took part in the study. They were recruited from a variety of sources including internal announcements within the university undergraduate student population and online requests sent to friends and colleagues, who were also asked to forward the test to their acquaintances. Because there was particular interest in 3-D video game exposure, the experimenter also posted a request for participants on two video-game related online message boards. Each survey took between thirty minutes to one hour to complete. Eighteen participants submitted incomplete data and were not included in the analyses.

Of the 154 participants included in the final sample, 122 were male and 32 were female, with a mean age of 28.4 years (SD = 8.9; range = 16 to 57). On average, the participants had completed 16.0 years (SD = 2.8; range = 11 to 25) of education.

Materials

The materials consisted of another paper-and-pencil media survey, a test of fluid intelligence and a test of crystallized intelligence.

To determine frequency of media usage, a slightly-modified version of the media exposure and demographics survey used in Study 1 was employed. The survey began with the same questions for sex and age, but added one for years of education. These were followed by the reading question and five media exposure questions (TV or movies, 3-D video games, listen to music or play an instrument, 2-D games, and crosswords or other logic puzzles). The difference from Study 1 was that there was a new music question but only one general question for each of 3-D video games and 2-D video games. The 3-D games were illustrated by the examples of first person shooters and action/adventure and the 2-D games were illustrated by *Super Mario Bros.*, *Tetris* and *Minesweeper*. The crossword/puzzle question was illustrated by Sudoko, Solitaire and anagrams.

For each question, participants were asked to consider the past year and to choose how often, on average, they engaged in the activity using five frequency options: every day, 2-6 times per week, 1-7 times per month, less than once a month and never. This was followed by 6 response options for the length of a typical session: not applicable, 15 minutes, 30 minutes, 1 hour, 2 hours and 3 hours+. A score was obtained for each category by coding the frequency categories from 5 (every day) to 1 (never), by coding the session duration from 1 (not applicable) to 6 (3 hours+) and then by multiplying the two numbers. The range of scores was from 1 to 30.

The dependent variables were scores on a test of fluid intelligence and on a test of crystallized intelligence. The first one was an untimed version of Cattell's Culture Fair Intelligence Test: Scale 3, Short Form A (CFIT; Cattell & Cattell, 1972; IPAT, 1973), which contained 47 of the original 50 stimuli from the four subtests. The three missing questions were omitted so that each subtest could be fitted neatly on to a single page. However, they were the last (most difficult) two questions from Subtest 2 and the last (most difficult) question from Subtest 3. Although the CFIT is normally administered with strict time limits for each subtest (total time = 12.5 min.), it was administered here without a time limit to accommodate online testing. Notably, Gregory (2011, p. 222) mentions that Scale 2 can be given as an untimed power test, although he adds the caveat that norms are limited. It was reasoned here Scale 3 could also be given as a power test, particularly because this has also been done with at least one other highly speeded intelligence test (Wonderlic, 1983). In addition, norms do not matter in the present case because only raw scores were analyzed.

As noted above, the CFIT was chosen because it is a recognized test of fluid intelligence. In addition, it has historical relevance given that Cattell (1940) first made the distinction between fluid and crystallized intelligence. Scale 3 was chosen because it is appropriate for university students and educated adults (Cattell & Cattell, 1972; IPAT, 1973). In addition, it has adequate psychometric properties. Internal consistency reliability is .74 and test-retest reliability (for times up to one week) is .69 (IPAT, 1973). Although these values are smaller than the benchmark of .80, which is often invoked as a standard, they are acceptable for research purposes, where standards are lower than for individual test interpretation (Gregory, 2011, p. 103). Validity is supported by a correlation of .85 with the "g" factor and by an average criterion validity coefficient of .66 with other mainstream tests (IPAT, 1973).

The second dependent variable, crystallized intelligence, was measured with the Wide Range Vocabulary test (WRVT; Atwell & Wells, 1972). To facilitate the ease of administration and testing time, the official 100-item WRVT was divided into two 50-item tests, designated Form A and Form B, that were equated for item difficulty. Participants were then randomly assigned to receive either one of the forms. The WRVT was chosen because, although crystallized intelligence involves accumulated knowledge in a number of areas (Cattell & Horn, 1978), vocabulary tests are considered the most valid short measure of crystallized intelligence (Gregory, 2011, p. 168; Lynn, 2009).

Unfortunately, there is no psychometric information in the WRVT manual (Atwell & Wells, 1972). However, it has a correlation of .73 with the Shipley-Institute of Living Scale (Martin, Blair, & Vickers, 1979), which the researchers interpret as evidence of validity, particularly with college students. In addition, the content of the WRVT is similar to the verbal sub-scale of Bennet, Bennet, Wallace and Wesman's (1961) College Qualification Test, for which the manual reports high reliability and validity coefficients.

Procedure

The demographic and media survey and the two intelligence tests were prepared in digital PDF format, which most participants completed on line. The others were tested individually in person with a paper version. After reading and signing the consent form, each participant completed the media exposure survey and then the CFIT and WRVT tests in a counterbalanced order. Once finished, participants submitted their answers and were given a debriefing form.

Results

As in Study 1, correlations were calculated followed by hierarchical and stepwise multiple regression analyses. However, this time, the predictors were reading, five media predictors (TV/movies, 3-D video games, 2-D video games, music, and puzzles) and three demographic variables (sex, age, and years of education). The two dependent variables were scores on the CFIT and on the WRVT. There was a small but significant correlation between their scores, $r = .302, p < .001$.

In Study 1, the results of both the hierarchical and stepwise analyses were reported, and were found to be the same. In this study, both analyses again yielded the same results for the CFIT and for the WRVT, but only the hierarchical analyses will be presented because they are based directly on predicted relationships.

Descriptive statistics and correlations are shown in Table 3. For the CFIT, the only significant predictor of performance was years of education, although age approached significance.

Two hierarchical regression analyses were performed (see Table 4). In the first case, it was reasoned that exposure to 3-D video games should be entered first because it predicted mental rotation performance in Study 1. At the second step, the other three

Table 3: *Descriptive Statistics and Correlations (r) for for CFIT and WRVT Performance and for Predictors in Study 2*

Variable	*M*	*SD*	*r* with CFIT	*p*	*r* with WRVT	*p*
CFIT	31.08	5.94				
WRVT	38.31	4.80				
Sex			.039	.634	.020	.805
Age	28.41	8.96	-.151	.062	.338	<.001
Education	16.06	2.90	.175	.030	.162	.044
3-D Weighted	11.40	8.39	.032	.694	-.033	.685
2-D Weighted	6.77	5.74	.010	.900	.036	.661
TV/movie weighted	19.63	5.90	-.114	.161	.071	.382
Puzzle weighted	7.08	5.54	.082	.311	.181	.024
Music weighted	18.93	8.05	-.087	.282	-.096	.238
Reading weighted	14.38	6.29	.099	.220	.299	.001

Note: $n = 154$.

visual media variables (2-D video games, TV/movies, puzzles) were added. At the third step, music and reading were added and at the final step the three demographic variables (sex, age, education) were included. R2 at each step was .001, .024, .038, and .100. None of these values was significant, although the last one came close. However, the change in R2 from the third to the fourth step was significant. In the final regression equation, two predictors were significant: age, β = -.204, and education, β = .188.

In the second hierarchical regression analysis, all the visual media variables were

Table 4: *Hierarchical Regression Analyses for Predictors of CFIT and WRVT Performance in Study 2*

Model	R^2	F	df_1	df_2	p	R^2 Change	F	df_1	df_2	p	β	t	p
CFIT Analysis 1													
1	.001	0.16	1	152	.694	.001	0.16	1	152	.694			
2	.024	0.91	4	149	.461	.023	1.16	3	149	.328			
3	.038	0.98	6	147	.441	.015	1.12	2	147	.329			
4	.100	1.79	9	144	.076	.062	3.31	3	144	.022	-.204*	2.28	.024
											.188**	2.25	.026
CFIT Analysis 2													
1	.024	0.91	4	149	.461	.024	0.91	4	149	.008			
2	.038	0.97	6	147	.441	.015	1.12	2	147	.449			
3	.100	1.79	9	144	.076	.062	3.31	3	144	.969	-.204*	2.27	.024
											.188**	2.25	.026
WRVT Analysis													
1	.089	14.90	1	152	<.001	.089	14.90	1	152	<.001	.299*	3.86	<.001
2	.113	3.78	5	148	.003	.024	1.00	4	148	.411	.300*	3.64	.001
3	.139	3.95	6	147	.001	.026	4.35	1	147	.039	.310*	3.70	<.001
											-.165**	2.09	.039
4	.214	4.35	9	144	<.001	.075	4.58	3	144	.004	.261*	3.17	.002
											.274***	3.26	.001

Note: Predictors for the CFIT analyses were 3-D video, 2-D video, TV, puzzles (Model 1), 3-D video, 2-D video, TV, puzzles, music, reading (Model 2), 3-D video, 2-D video, TV, puzzles, music, reading, sex, age, education (Model 3). Significant predictors were *age, **education.

Predictors for the WRVT analysis were reading (Model 1), reading, 3-D video, 2-D video, TV, puzzles (Model 2), reading, 3-D video, 2-D video, TV, puzzles, music (Model 3), reading, 3-D video, 2-D video, TV, puzzles, music, sex, age, education (Model 4). Significant predictors were *reading, **music, ***age.

entered together first, because Neisser (1997) did not distinguish among them when he discussed the complex visual and technical environment. At the second step, music and reading were added and at the third and final step the three demographic variables (sex, age, education) were included. R2 at each step was .024, .028, .100. Again, none of these values were significant, although the last one was close. In addition, the change from the third to the fourth step was again significant. In the final regression equation, the same two predictors were significant: age, β = -.204, and education, β = .188.

Because the 3-D video game experience was the CFIT performance predictor that was identified as being most likely to be significant in this study, two supplementary analyses were run in which it was entered along with sex of participant and then along with education level. In the first case, neither 3-D video game experience or sex were significant. In the second case, 3-D video game experience was again not significant but, consistent with the analyses above, education was significant. However, the interaction between 3-D video game experience and education level approached significance, $p = .085$. Inspection of the data suggested that the positive relationship between education and CFIT performance was greater at a lower level than at moderate and higher levels of 3-D game experience. Another way of looking at this interaction is that at a lower level of education (less than 16 years), CFIT performance was lower at lower 3-D game experience levels than at moderate or higher levels. However, at higher levels of education, CFIT performance was unrelated to 3-D game experience.

For the WRVT, there were four significant correlations. Performance was predicted positively by age, education, puzzles and reading.

For the WRVT, one hierarchical regression analysis was performed. Reading was entered first because it was expected to predict performance. At the second step, the four visual media variables (3-D video games, 2-D video games, TV/movies, puzzles) were added. At the third step, music was added and at the final step the three demographic variables (sex, age, education) were included. R2 at each step was .089, .113, .139, and .214, all of which were significant. In addition, two changes in R2 were significant: from the second to the third step, and from the third to the fourth step. In the first step, reading was significant, β = .299, and it remained significant by itself in the second step (β =.300). In the third step, reading (β = .310) was joined by music, β = -.165. In the fourth and final regression equation, music disappeared, and reading, β = .261, was joined by age, β = .274.

Discussion

Results for Main Question

The main purpose of Study 2 was to investigate whether exposure to 3-D video games and perhaps also other visual media variables, would be positively related to fluid intelligence, which has implications for Neisser's (1997) hypothesis that the Flynn effect may be due to increased exposure to a complex visual and technical environment. If Neisser is correct, exposure to visual media should correlate positively with the scores on the test of fluid intelligence but not on a test of crystallized intelligence. Further, from the results of the first study, it was hypothesized that the relationship between the visual media and fluid intelligence would be most strongly fuelled by exposure to 3-D video games.

The most important outcome in the present study is that there were no significant relationships between the visual media variables and scores on the CFIT (fluid intelligence).

That is, the results seem to support the null hypothesis. This is inconsistent with Neisser's proposal that the Flynn effect, and particularly the increase in fluid intelligence, is due to increased exposure to the complex visual environment. In addition, the results of Study 2 show that the relationship between 3-D video game exposure and mental rotation performance that was found in Study 1 did not generalize to performance on a test of fluid intelligence. The lack of any significant relationship between 3-D video game experience and scores on the CFIT is also consistent with the results obtained by Boot et al. (2008), who found that there was no difference in performance on the Raven test between gamers and non-gamers.

The only exception to these clearly nonsignificant results was a slight positive relationship between 3-D game experience and CFIT performance at lower levels of education. However, this trend only approached significance and was examined in isolation as a supplementary analysis following the main hierarchical regression. It might be examined more systematically in future research with care taken to ensure a wide variety of education levels.

Other Results

Exposure to puzzles was significantly and positively correlated with WRVT performance, but it was not significant in any of the regression analyses. In addition, although exposure to music appeared in the hierarchical regression analysis as a negative predictor when it was added by itself at the third step, it disappeared in the fourth and final step when the demographic variables were included and it did not appear in the stepwise regression. In short, none of the media variables predicted performance on the CFIT or on the WRVT in the final regression analyses. For music exposure, this was expected. Music does not possess the visual or verbal components that would be important for predicting performance on the CFIT and WRVT respectively.

In addition, and consistent with Neisser's hypothesis, exposure to visual media was not related to crystallized intelligence (WRVT). Notably, however, scores on the test of crystallized intelligence were positively correlated with age, level of education, reading exposure and puzzle exposure, although only age and reading remained in the regression analysis. The increase in crystallized intelligence with reading was expected, because exposure to text should help people increase their vocabulary. Although the significant positive correlation between crystallized intelligence and education did not appear in the regression analysis, the initial relationship is consistent with the relationship with reading, because more education should also help people increase their vocabulary. In addition, the increase with age is consistent with previous studies showing that crystallized intelligence, particularly vocabulary (Bowles & Salthouse, 2008) increases with age (Beier & Ackerman, 2005; Gregory, 2011, p. 265; Horne & Cattell, 1967). Given the agreement with past research, this finding also supports the use of the WRVT as a valid measure of crystallized intelligence.

Critical Discussion of Methodology

One factor in Study 2 that may have concealed a relationship between exposure to visual media and performance on the CFIT is that self-reports on a survey may not provide an accurate measure of the participants' actual experience with the media. However, this approach to evaluating a person's use of media has been used in previous research (e.g., Feng et al., 2007; Quaiser-Phol et al., 2006; Sims & Meyer, 2002). In addition, questions might be raised about the specific manner in which experience was quantified here. People chose a frequency option from daily to never and then estimated the time for each session from specified lengths of time. These two choices were then multiplied. However, with the exception of the time for each session, which was open ended in Study 1, the method of quantifying experience on the survey in Study 2 was very similar to the one used in Study 1 in which the previous relationship between exposure to 3-D video games and mental rotation performance (Feng et al., 2007) was replicated. It might have been better in Study 2 to have participants give precise estimates for both the number of sessions and for the time in these sessions, so that a single time estimate for the amount of exposure the over the past year could be calculated. However, overall, the lack of a significant relationship between media exposure and CFIT performance in Study 2 is not easily attributable to weaknesses in the manner in which media exposure was assessed in the survey. Consequently, it is unlikely that faults in the measuring instrument itself seriously undermine the contribution of the present results to the evaluation of Neisser's visual media hypothesis. On the other hand, the results from the survey were combined for the majority of the people who took it on line and for the others who were tested individually. There was less control over testing conditions in the first case, which means that the administration procedure was not completely standardized. This may have introduced some degree of random error in the results.

Another factor that may have weakened the methodology of Study 2 is that the CFIT was administered as a pure power test without a time limit. Although the test can be given in this manner (Gregory, 2011, p. 222), the standard instructions require that each subtest is strictly timed, with a total of 12.5 minutes permitted for the four parts of the test. In other words, the CFIT is highly speeded. It is possible that exposure to visual media, and particularly to fast action 3D video games, might have been related to performance on a timed CFIT, the condition under which a Flynn effect was found with the test (Colom & Garcia-Lopez, 2003). If this issue were pursued in the future, the test could be timed or, if participants were not given a time limit, their time could be noted and included as a factor in the regression analysis.

On the other hand, it has been shown that scores on timed and untimed versions of intelligence tests are correlated (Chandra, 1956; McKelvie, 1994), and that criterion validity coefficients are similar for the timed and untimed versions of an intelligence test (the Wonderlic Personnel Test; McKelvie, 1994). In addition, the Raven test, with which the Flynn effect has been repeatedly demonstrated, is untimed under standard administration conditions (Raven, 1952; Raven, Court, & Raven, 1996). Therefore, it is unlikely that giving the CFIT without a time limit seriously undermines the contribution of the present results to the evaluation of Neisser's visual media hypothesis.

Finally, whenever nonsignificant relationships are obtained, and particularly if an argument is being made in favour of the null hypothesis, it must be considered whether the study has sufficient power and whether the variables might have been subject to restriction of range.

With regard to power, the main determinant is sample size. In Study 2, there were 154 participants. In multiple regression analysis, a minimum of 10 cases for each predictor has been commonly recommended as a rule of thumb (Knofczynski & Mundfrom, 2008; Maxwell, 2000). With nine predictors, this implies that there should be at least 90 participants. The present sample size exceeds that and is therefore adequate. In Study 1 there were five predictors, which implies at least 50 participants. The sample size of 35 is less than that and is therefore less than adequate. However, despite this restriction, there was a significant relationship between 3-D video exposure and mental rotation performance. If such a relationship existed for the CFIT in Study 2, the sample size was adequate to detect it. Taken together, these considerations suggest that the sample size in Study 2 was sufficient to detect a significant relationship if it had been present.

On the other hand, the common rule of thumb suggestions for sample size have been criticized as too low (Knofczynski & Mundfrom, 2008; Maxwell, 2000; Tabacnik & Fidell, 2001), particularly when the purpose of the regression is taken into account and when the expected value of R2 can be specified. For example, Knofczynski and Mundfrom claim that if R2 is expected to be .20, at least 44 cases per predictor are required if there are nine predictors. This would imply that almost 400 participants should have been tested, which far exceeds the present number.

However, it should be recalled that the correlation coefficients between performance on the CFIT and three of the predictors relevant to Neisser's hypothesis (3-D video games, 2-D video games, puzzles,) did not approach significance (ps > .600) and were close to 0. The remaining one (TV) was closer to significance ($p = .16$), but was actually negative (see Table 3). The values of β in the regression equations are only reported in Table 4 for the significant predictors, but the results for the visual media variables (βs < .13 for 3-D video games, 2-D video games and puzzles) are commensurate with these correlations. Therefore, it is extremely unlikely that a larger sample size would have shown that these relationships were significant.

With regard to restriction of range, if the full range of scores is not present in the sample, a significant relationship might not be detected. Given the successful use of the survey in Study 1, the restriction of range hypothesis will be considered here with reference to the two dependent variables, the CFIT and the WRVT, particularly because restriction of range is a potential problem when maximum performance tests are administered to selected groups such as university undergraduates. The range of scores for the 2-D and 3-D games predictors were also examined.

In Study 2, there were university undergraduates, but there were also other participants with varied ages and varied levels of education. In addition, both tests were demanding and therefore were likely to discriminate among those who were university students. Third, the possible minimum and maximum scores on these two tests are 0 to 47 and 0 to 50 respectively. Because the obtained ranges were 12 to 42 and 20 to 48, the scores were well spread throughout the range. The means and standard deviations for each test also support this argument (see Table 3). For the CFIT, $M = 31.01$ and $SD = 5.94$. Assuming normality (see below), which means that the scores were spread from about 3 *SD*s below the mean to 3 *SD*s above the mean, the implied range is 31+/-15 or 13 to 39, which is close to the actual range. For the WRVT, $M = 38.31$ and $SD = 4.80$, giving an implied range of 38+/-15 or 23 to 53, which is close to the actual range. Finally, inspection of their distributions confirmed that they were both close to normal. Consequently, it does not seem likely that the absence of a relationship between visual media variables and the CFIT and WRVT is due to restriction in range.

With regard to the survey that assessed exposure to video games, the minimum and maximum values were 1 to 30 for both the 2-D games weighted scores and the 3-D games weighted scores. Because the obtained ranges were also 1 to 30 in both cases, these scores were well spread throughout the range. Consequently there was no restriction in range for the two video games variables.

Moreover, although fluid intelligence scores on the CFIT were unrelated to any of the media variables, in the regression analysis they were predicted negatively by age and positively by years of education. In fact, for age, the standard deviation and range were both much greater in Study 2 than in Study 1. These results are consistent with other studies showing that fluid intelligence declines with age (Beier & Ackerman, 2005; Gregory, 2011, pp. 267-268; Horne & Cattell, 1967), particularly on the CFIT (Rabbitt, Lunn, Ibrahim, & McInnes, 2009), and increases with education level (Kaufman, Kaufman, Liu, & Johnson, 2009; Murphy, Cassimjee, Nafisa, & Schur, 2011). Although fluid intelligence is independent of acquired knowledge through acculturation, the skills required can be learned in other ways (Cavanagh & Blanchard-Fields, 2008, pp. 268-269). For example, with more education, people are likely to have more experience with abstract logic-based relationships, which would assist them on the CFIT. This possibility is consistent with the evidence that scores on tests of fluid intelligence can be improved by programmes in which people are trained directly to use appropriate strategies (Baltes et al., 1989; Denny & Heindrich, 1990).

Given their agreement with past research, these two findings on age and education add further support for the use of the untimed version of the CFIT as a measure of fluid intelligence. In addition, because crystallized intelligence is formed in part via fluid intelligence (e.g., Cattell & Butcher, 1968, p. 19), they should be positively correlated. This has been found, with an overall correlation of .50 (Gregory, 2011, p. 168). Of particular relevance here, correlations between fluid intelligence and crystallized intelligence have been obtained when they were measured, respectively, by the Raven and by the vocabulary test from the Wechsler adult intelligence scale. In one study, the correlations were .672 for younger people and .386 for older people (Cunningham, Clayton, & Overton, 1975), and in a second study they were .35, .34, and .39 for younger, middle-aged and older people, respectively (Stoner, 1982). In the present case, the correlation between scores in the CFIT and scores on the WRVT was .302. Although this is lower than the general results and the results of the first study for younger people, it is similar to the results of the second study and provides additional support for the validity of CFIT and WRVT as measures of fluid intelligence and crystallized intelligence.

General Discussion

The main purpose of these two studies was to investigate Neisser's (1997) hypothesis that the Flynn effect can be at least partly accounted for by the increased complexity of the visual and technical environment. The first step was to replicate previous reports that experience with 3-D video games would be associated with performance on a classic mental rotation task, and to investigate if this relationship might generalize to other visual media as described by Neisser. The replication was successful, but the association between mental rotation performance and 3-D video game exposure did not generalize to other visual media.

The second step was to examine Neisser's hypothesis more directly. This was accomplished by investigating if exposure to visual media in general, and to 3-D video game experience in particular, would be associated with performance on a test of fluid intelligence, which has been shown to exhibit a strong Flynn effect, but not with performance on a test of crystallized intelligence, which has been shown to exhibit either a weaker Flynn effect or none at all. Although exposure to visual media was not associated with scores on crystallized intelligence, as expected, it was also not significantly associated with scores on fluid intelligence. Together, these results are inconsistent with Neisser's visual media hypothesis.

To explain the present findings, it may be useful to again consider the issue of near transfer and far transfer and the question of the processes that are common to the training variables and to the test variables. First, as noted in the discussion of Study 1, far transfer between exposure to 3-D video game exposure and mental rotation performance can perhaps be accounted for by the common process of active imagery. Second, the absence of a far transfer effect in Study 2 may be due to the fact that, although mental rotation tasks and spatial ability both load on Gf, the dynamic imagery that is dominant in 3-D video games may not be directly relevant for solving problems on the CFIT. Similarly, the kinds of skill developed by playing 2-D video games, by watching TV and movies, and by completing crosswords and other logic puzzles may also not be useful for solving problems on the CFIT.

This raises the question of what processes are tapped by the CFIT. According to Cattell and Horn (1978), the essence of fluid intelligence is "analytic ability", which requires that participants discover rules that will help them understand relations. To use Spearman's terminology, tests of fluid intelligence "should be concerned with the eduction of relations and correlates" (Cattell & Horn, 1978, p. 143). They state that the Raven test taps these processes, but does so only with a single kind of item (matrices). The CFIT has a matrix subtest, but also contains three others - series, classification, and topology. In addition, they state that although these kinds of items are presented visually and spatially, the key process that they tap is the extraction of relations.

With this in mind, and as mentioned above, it seems that the dynamic visual processes involved in playing 3-D and even 2-D video games may not require extraction of logical relations and therefore may not be important for solving problems on the CFIT. Similarly, although solving crossword puzzles may involve detecting relationships, they also depend highly on verbal knowledge, which is not important for solving problems on the CFIT. In contrast, the processes tapped by other logic puzzles might be more relevant to the CFIT. A problem in the present study may have been that the question about exposure to this kind of puzzle was combined with the question about crosswords. Notably, although the relationship did not appear in the regression analysis, there was a significant correlation between exposure to puzzles and performance on the WRVT. This association may have been due to the crossword part of the question. Completing crosswords is likely to improve a person's lexicon, giving them an advantage on the WRVT. Given that level of education predicted CFIT performance, it is reasonable to assume that higher levels of education demand greater use of analytic thinking that is similar to the kind of process tapped by logic puzzles. The importance of exposure to visual logic puzzles for the CFIT and to crossword puzzles for the WRVT could be investigated further with separate questions.

However, even if performance on the CFIT was related to exposure to logic puzzles, such a discovery would not provide much support for Neisser's (1997) general hypothesis, which is centred on the idea that the complex visual and technical environment develops the skill of "visual analysis" and that this skill is a factor in the rising scores on intelligence tests, particularly fluid intelligence. The present argument, and particularly the null relationship between exposure to visual media and performance on the CFIT, implies that the main cause of the Flynn effect lies elsewhere among the other explanations that have been offered (Flynn, 1984; Mingroni, 2007; Neisser, 1997). Consequently, future research should focus on these other factors ranging from increases in nutrition and smaller families to improved test-taking sophistication, schooling, and child-rearing practices.

References

Atwell, C.R., & Wells, F.L. (1972). *Wide range vocabulary test.* The Psychological Corporation: New York, N.Y.

Baltes, P.B., Sowarka, D., & Kliegl, R. (1989). Cognitive training research on fluid intelligence in old age: What can older adults achieve by themselves? *Psychology and Aging, 4*(2), 217-221.

Beier, M. E & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging, 20,* 341-355.

Bennett, G.K., Bennett, M.G., Wallace, W.L., & Wesman, A.G. (1961). *College qualification test manual*. The Psychological Corporation: New York, N.Y.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiana, M., & Gratton, G. (2008). The effect of video game playing on attention, memory, and executive control. *Acta Psychologia, 129,* 387-398.

Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiana, M., et al. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychologia, 135,* 349-357.

Bowles, R. P., & Salthouse, T. A. (2008). Vocabulary test format and differential relations to age. *Psychology and Aging, 23,* 366-376.

Cattell, R. B. (1940). A culture free intelligence test. I. *Journal of Educational Psychology, 31,* 161-179.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1-22.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action.* New York: Houghton Mifflin.

Cattell, R. B., & Butcher, H. J. (1968). *The prediction of achievement and creativity.* New York: the Bobbs-Merrill Company, Inc.

Cattell, R.B., & Cattell, A.K. (1972). *Test of "g": Culture fair.* The Institute for Personality and Ability Testing: Champaign, IL.

Cattell, R. B. & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement, 15,* 139-164.

Cavanaugh, J.C., & Blanchard-Fields, F. (2008). *Adult development and aging* (1st Canadian ed.). Toronto: Nelson.

Chandra, J. M. (1956). A study of intelligence scores with and without time-limit. *Indian Journal of Psychology, 31*, 27-29.

Chabris, C. F. (1999). Prelude or requiem for the 'Mozart effect'? *Nature, 400,* 826-827.

Colom, R., Andres-Pueyo, A., & Juan-Espinosa, M. (1998). Generational IQ gains: Spanish data. *Personality and Individual Differences, 25,* 927-933.

Colom, R., Contreras, Ma J., Botella, Juan, & Santacreu J. (2001). Vehicles of spatial ability. *Personality and Individual Differences, 32,* 903-912.

Colom, R., & Garcia-Lopez, O. (2003). Secular gains in fluid intelligence : Evidence from the culture-fair intelligence test. *Journal of Biosocial Science, 35,* 33-39.

Cunningham, W. R., Clayton, V., & Overton, W. (1975). Fluid and crystallized intelligence in young adulthood and old age. *Journal of Gerontology, 30,* 53-55.

Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P., & Neumann, C. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science, 14,* 215-219.

Denney, N.W., & Heidrich, S.M. (1990). Training effects on Raven's Progressive Matrices in young, middle-aged, and elderly adults. *Psychology and Aging, 5*(1), 145.

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108,* 346-369.

Di Vesta, F. J., Ingersoll, G., & Sunshine, P. (1977). A factor analysis of imagery tests. *Journal of Verbal Learning and Verbal Behavior, 10,* 471-479.

Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62,* 159-162.

Drew, B., & Waters, J. (1986). Video games: Utilization of a novel strategy to improve perceptual motor skills and cognitive functioning in the non-institutionalized elderly. *Cognitive Rehabilitation, 4*, 26-31.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video gamer reduces gender differences in spatial cognition. *Psychological Science, 18,* 850-855.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171-191.

Flynn, J. R. (1998a). WAIS-III and WISC-III: IQ gains in the United States from 1972 to 1995: How to compensate for obsolete norms. *Perceptual and Motor Skills, 86,* 1231-9.

Flynn, J. R. (1998b). IQ gains over time. In U. Neisser (Ed.), *The rising curve* (pp. 25-66). Washington, DC: American Psychological Association.

Flynn, J. R. (1999). Searching for justice: The discover of IQ gains over time. *American Psychologist, 54,* 5-20.

Flynn, J., Shaughnessay, M. F., & Fulgham, S. W. (2012). An interview with Jim Flynn about the Flynn effect. *North American Journal of Psychology, 14,* 25-38.

Freund, P. A., & Holling, H. (2011). How to get really smart: Modeling retest and training effects in ability testing using computer-generated figural matrix patterns. *Intelligence, 39,* 233-243.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the Scholastic Assessment Test and General Cognitive Ability. *Psychological Science, 15,* 373-378.

Gilmore, M. E. (1927). Coaching for intelligence tests. *Journal of Educational Psychology, 18,* 118-121.

Green, C.S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology, 32*(6), 1465-78.

Green, C.S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science, 18*(1), 88-94.

Greenfield, P.M. (1998). The cultural elevation of IQ (pp. 81-123). In U. Neisser (Ed.), *The rising curve* (pp. 25-66). Washington, DC: American Psychological Association.

Gregory, R.J. (2011). *Psychological testing: History, principles, and applications* (6th ed.). Boston: Allyn & Bacon.

Greiffenstein,M. F. (2011). Secular increases by epigenesist? The hypothesis of cognitive genotype optimization. *Psychological Reports, 109,* 353-366.

Horne, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence, *Acta Psychologia, 26,* 107-129.

IPAT (1973). *Measuring intelligence with the Culture Fair Tests: Manual for Scales 2 and 3.* Champaign, IL: Institute for Personality and Ability Testing.

Johnson, S. (2003). *Everything bad is good for you: How today's popular culture is making us smarter.* New York: Riverhead Books.

Johnson, S. (2005). Dome improvement. Retrieved May 5, 2005 from http://www.wired.com/wired/archive/13.05/flynn_pr.html.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J. Jr. (2008). Still just 1 g : Consistent results from five test batteries. *Intelligence, 36,* 81-95.

Kaufman, Alan S., Kaufman, J. C., Liu, X., & Johnson, C. K. (2009). How do educational attainment and gender relate to fluid intelligence, crystallized intelligence, and academic skills at ages 22–90 years? *Archives of Clinical Neuropsychology, 24,* 153-163.

Khaleefa, O., Abdelwahid, S. B., Abdulradi, F., & Lynn, R. (2008). The increase of intelligence in Sudan 1964-2006. *Personality and Individual Differences, 45,* 412-413.

Kulik, J. A., Bangert-Downs, R. L., & Kulik, C.-L. C. (1984). Effectiveness of coaching for aptitude tests. *Psychological Bulletin, 95,* 179-188.

Lohman, D. F. (1993). Spatial ability and G. Paper presented at the first Spearman Seminar, University of Plymouth, July 21.

Lynn, R. (2009). Fluid intelligence but not vocabulary has increased in Britain, 1979-2008. *Intelligence, 37,* 249-255.

Martin, J. D., Blair, G. E., & Vickers, D. M. (1979). Correlation of the Quick Test and Wide Range Vocabulary Test with the Shipley Institute of Living Scale. *Educational and Psychological Measurement, 39,* 935-937.

McKelvie, S. J. (1994). Validity and reliability findings for an experimental short form of the Wonderlic Personnel Test in an academic setting. *Psychological Reports, 75,* 907-910.

Metzler, J., & Shepherd, R. N. (1974). Transformational studies of the internal representation of three-dimensional objects. In R. L. Solso (Ed.), *Theories of Cognitive Psychology: The Loyola Symposium* (pp. 147-201). Potomac, MD. Lawrence Earlbaum Associates.

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review, 114,* 806-829.

Murphy, K., & Spencer, A. (2009). Playing video games does not make for better video attention skills. *Journal of Articles in Support of the Null Hypothesis, 6,* 1-20.

Murphy, R., Cassimjee, N., & Schur, C. (2011). Influence of socio-demographic factors on SRAVEN performance. *Journal of Psychology in Africa, 21*, 91-102.

Neisser, U. (1997). Rising scores on intelligence tests. *American Scientist, 85,* 440-447.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77-101.

Nettlebeck, T., & Wilson, C. (2004). The Flynn effect: Smarter not faster. *Intelligence, 32*, 85-93.

Okagaki, L., & Frensch, P.A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15,* 33-58.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Pervasiveness of the IQ rise: A cross-temporal meta-analysis. *PLoS ONE, 5,* 1-6.

Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences, 40*(3), 609-619.

Rabbitt, P., Lun, M., Ibrahim, S. & McInnis, L. (2009). Further analyses of the effects of practice, dropout, sex, socio-economic advantage and recruitment cohort differences during the University of Manchester longitudinal study of cognitive change in old age. *The Quarterly Journal of Experimental Psychology 62,* 1859-1872.

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature, 365,* 611.

Raven, J. C. (1956) *Guide to using Progressive Matrices* (1938). London: H. K. Lewis.

Raven, J. C., Court, J. H., & Raven, J. (1996). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 3, Standard Progressive Matrices 1996 Edition.* Oxford: Oxford Psychologists Press.

Rideout, B. E., Dougherty, S., & Wernert, L. (1998). Effect of music on spatial performance: A test of generality. *Perceptual and Motor Skills, 86,* 512-514.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15,* 511-514.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701-703.

Sims, V.K., & Mayer, R.E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology, 16*(1), 97-115.

Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Review of General Psychology, 14*, 93-104.

Stankov, L. (1986). Kvashchev's experiment: Can we boost intelligence? *Intelligence, 10,* 209-230.

Steele, K. M., Bass, K. E., & Crook, M. D. (1999). They mystery of the Mozart effect: Failure to replicate. *Psychological Science, 10,* 366-369.

Stoner, S. B. (1982). Age differences in crystallized and fluid intellectual abilities. *The Journal of Psychology, 110,* 7-10.

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

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence, 32,* 349-362.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*, 4th ed. Boston: Allyn and Bacon.

Tapley, S. M., & Bryden, M. P. (1977). An investigation of sex differences in spatial ability: Mental rotation of three-dimensional objects. *Canadian Journal of Psychology, 31,* 122-130.

Tarr, M.J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233-282.

Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn effect in reverse. *Personality and Individual Differences, 39,* 837-843.

Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scors: A reversal of the Flynn effect. *Intelligence, 36,* 121-126.

Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boosma, D. I., & Span, M. M. (2004). Are intelligence test measurements invariant over time? Investigating the nature of the Flynn effect. *Intelligence, 32,* 509-537.

Widenbauer, G., Schmid, J., & Jansen-Osmann, P. (2007). Manual training of mental rotation. *European Journal of Cognitive Psychology, 19,* 17-36.

Wonderlic, E. F. (1983). *Wonderlic Personnel Test Manual.* Northfield, IL: E. F. Wonderlic.

Received: 4.10.2012
Revised: 5.21.2012
Accepted: 5.22.2012