Resource Allocation and Fluid Intelligence: Insights from Pupillometry

Elke van der Meer[1], Reinhard Beyer[1], Judith Horn[1], Manja Foth[1], Boris Bornemann[1], Jan Ries[1], Juerg Kramer[2], Elke Warmuth[2], Hauke R. Heekeren[3,4], Isabell Wartenburger[3,5]

[1]Department of Psychology at Humboldt University Berlin, [2]Department of Mathematics at Humboldt University Berlin, [3]Berlin NeuroImaging Center and Neuroscience Research Center Charité Berlin, [4]Max-Planck-Institute for Human Development Berlin, [5] Department of Linguistics, University of Potsdam

Corresponding author:

Elke van der Meer

Humboldt University Berlin

Department of Psychology

Rudower Chaussee 18

12489 Berlin

Phone: ++49-(0)30 – 20939390

Email: meerelke@cms.hu-berlin.de

Running title:

Resource Allocation and Fluid Intelligence

Abstract

Thinking is biological work and involves the allocation of cognitive resources. The aim of this study was to investigate the impact of fluid intelligence on the allocation of cognitive resources while processing low-level and high-level cognitive tasks. Individuals with high versus average fluid intelligence performed low-level choice reaction time tasks and high-level geometric analogy tasks. We combined behavioral measures to examine speed and accuracy of processing with pupillary measures, which indicate resource allocation. Individuals with high fluid intelligence processed the low-level choice reaction time tasks faster than normal controls. The task-evoked pupillary responses did not differ between groups. Furthermore, individuals with high fluid intelligence processed the high-level geometric analogies faster, more accurately, and showed greater pupil dilations than normal controls. This was only true, however, for the most difficult analogy tasks. In addition, individuals with high fluid intelligence showed greater pre-experimental pupil baseline diameters than normal controls. These results indicate that individuals with high fluid intelligence have more resources available and thus can solve *more demanding* tasks. Moreover, high fluid intelligence appears to be accompanied by more task-free exploration.

Key Words: fluid intelligence, resource allocation, geometric analogies, pupillary response

Resource Allocation and Fluid Intelligence: Insights from Pupillometry

Elke van der Meer, Reinhard Beyer, Judith Horn, Manja Foth, Boris Bornemann, Jan Ries, Jürg Kramer, Elke Warmuth, Hauke R. Heekeren, Isabell Wartenburger

Introduction

It has long been argued that all human reasoning, including logical inference, is essentially analogical and that the essence of intelligent insights lies primarily in making fluid analogies (French, 2002; Halford, 1992; Hofstadter, 1995; Holyoak & Thagard, 1995; James, 1890/1950; Klix, 1993; Mitchell, 1993). Fluid reasoning is one of the core components of fluid intelligence. Importantly, there is a strong relationship between fluid intelligence and the central executive of working memory (Duncan, 2003; Engle, Tuholski, Laughlin, & Conway, 1999; Gray, Chabris, & Braver, 2003). There is evidence that analogical reasoning requires specific executive processes, namely, selecting relevant and inhibiting irrelevant features, building and mapping relations, and providing interference resolution (Cho, Holyoak, & Cannon, 2007; Gentner, 1983; Holyoak & Thagard, 1995; van der Meer, 1996).

Across a broad range of cognitive tasks, individuals scoring high in fluid intelligence consistently perform better than individuals who score low. For example, individuals scoring high on the Raven Advanced Progressive Matrices (RAPM; Raven, 1958) show faster response times in reasoning tasks compared to individuals who score around the average (van der Meer, 1996; van der Meer & Klix, 1986).[1] Furthermore, Neubauer (1997) has observed a negative correlation (r = -.30) between psychometric intelligence and speed of information processing as indexed by the time required to perform elementary cognitive operations. Examples of these tasks are choice reaction time, reading rates, and coding of numbers or letters (e.g., Rindermann & Neubauer, 2001). This observation has lead to the proposal that individuals who score high in fluid intelligence use a limited set of fundamental cognitive operations more efficiently (Jensen, 1998; Neubauer, Freudenthaler, & Pfurtscheller, 1995;

Rypma, Berger, Prabhakaran, Bly, Kimberg, & Biswal et al., 2006; Vernon, 1983). These results, however, might also reflect the impact of other variables such as resource allocation.

Resource Allocation and Fluid Intelligence

Just, Carpenter, and Miyake (2003) have argued that cognition is biological work, which entails the consumption of resources. The concept of resources originally arose from Kahneman's (1973) capacity theory of attention and from the proposal by Just and Carpenter (1992) that defined resources as "the amount of activation *available* for information storage and processing" (p. 312) in the underlying cortical neural system. Importantly, the available pool of resources is assumed to be limited and to depend on (a) neurotransmitter functioning, (b) the various metabolic systems supporting the neural system, and (c) the structural connectivity of the neural system (Just et al., 2003). Variation within these systems is one source of individual differences in cognition. Similarly, Spearman (1904) suggested that fluid intelligence may correspond to the amount of "general mental energy" available to an individual. Another source of individual differences in cognition might be the *allocation* of resources — the amount of activation *actually invested* for information storage and processing. Just et al. (2003) verified three measures of activity as indices of resource allocation: functional brain imaging, event-related potentials, and pupil dilation.

One interesting question derived from this point of view refers to the relationship between the allocation of resources and fluid intelligence (cf. Ahern & Beatty, 1979). In the current work, we test resource allocation during the processing of cognitive tasks in individuals scoring high versus average in fluid intelligence.

Resource Allocation and Pupil Dilation

All cognitive efforts, like physical efforts and sensory stimuli, cause pupil dilation (Beatty & Kahneman, 1966; Hess & Polt, 1964; Kahneman & Beatty, 1967; Loewenfeld, 1993). Just et al. (2003) have demonstrated that the pupillary response reflects an overall aggregate of mental resource allocation that is not limited to a specific part of the cognitive system. According to this view, the pupil could thus be used to map the overall functional level of the cognitive system to the amount of activity in the underlying neural system. Beatty and Lucero-Wagoner (2000) argued that pupil dilation amplitude is a useful measure of task-evoked resource allocation. The more difficult a task, the more the pupil dilates (Nuthmann & van der Meer, 2005; Raisig, Welke, Hagendorf, & van der Meer, 2007; Verney, Granholm, & Marshall, 2004). For example, in a visual search task where different levels of search difficulty were contrasted, only the pupillary responses, but not response times, differentiated between conditions (Porter, Troscianko, & Gilchrist, 2007).

Of course, the amount of resources that are allocated to a task does not depend only on the cognitive demands of the task, but also on the intensity with which an individual engages in it. This intensity of task-engagement might also be reflected in pupil dynamics. It has long been known that pupil dilation increases with activation of the sympathetic nervous system (Loewenfeld, 1993). As the sympathetic nervous system regulates arousal, a higher pupillary dilation may indicate that an individual is applying his- or herself with more vigor to the task at hand (Ahern & Beatty, 1979). It has, however, been proposed by Yerkes and Dodson (1908) that performance increases with arousal only up to a point, and declines if this point is exceeded. This effect is in fact even stronger with increasing task difficulty, as Broadhurst (1959) points out.

This idea is also incorporated in a more differentiated view of task-related arousal that has been recently proposed by Aston-Jones and Cohen (2005). They propose that the activation of the cortex is strongly influenced by the locus coeruleus (LC), a structure in the dorsorostral pons that sends norepinephric projections to vast portions of the brain. In monkeys, LC activity is highly correlated with pupil dilation (Rajkowski, Kubiak, & Aston-Jones, 1993). For humans, this connection is not yet well established; however, studies by

Gilzenrat, Cohen, Rajkowski, and Aston-Jones (2003) tested predictions of Aston-Jones' and Cohen's LC-theory using pupillometry in humans, and found the predictions to be surprisingly well confirmed.

It is therefore reasonable to interpret pupil dilation in the light of Aston-Jones' and Cohen's theory of LC mediated task-engagement. In brief, the theory proposes two modes of activity: In the *tonic* mode, LC neurons exhibit a constantly high firing rate that renders the cognitive system sensitive to all kind of stimuli. This mode typically occurs when an individual is not bound to a particular task but rather "explores" its environment (low-task engagement). In the *phasic* mode, baserate firing is reduced and pronounced, punctual firings occur selectively in response to certain classes of stimuli. This mode typically occurs when the individual is engaged in a particular task and focuses on task-relevant stimuli while ignoring distracting environmental influences (high task-engagement). Bearing these theories in mind, we can use the pupillary responses to examine differences in resource allocation between individuals with high and average fluid intelligence that are due to different degrees of task-engagement.

Fluid Intelligence and Pupil Dilation

Pupil dilation also allows for discriminating between individuals who differ in fluid intelligence (for a review, see Beatty & Lucero-Wagoner, 2000). For example, Ahern and Beatty (1979) analyzed task-evoked pupil dilations in two groups of university students differing in intelligence (as indicated by their scores on the Scholastic Aptitude Test, SAT) while solving mental multiplication problems across three levels of difficulty. Individuals with higher SAT scores showed higher accuracy and smaller task-evoked pupil dilation than individuals with lower SAT scores. Since both groups did not differ in the magnitude of luminance-induced pupil dilations, the differences in the task-evoked pupillary responses were assumed to reflect differences in central brain processes, indicating that more intelligent individuals invested fewer resources.

Moreover, Heitz, Schrock, Payne, and Engle (2008) investigated the effect of incentives on working memory capacity in high- and low-span individuals. Individuals were presented a reading span task (consisting of sentence reading, letter encoding, and recall). High-spans exhibited larger pre-experimental and pre-trial pupil diameter baselines than low-spans. The incentive, however, affected recall performance in the reading span task equally for high- and low-span groups. Furthermore, task-evoked pupillary responses in the most demanding recall phase indicated that low-spans consumed more resources than high-span individuals. Taking into account the strong relationship between working memory and fluid intelligence (Duncan, 2003; Engle et al., 1999; Gray et al., 2003; Salthouse & Pink, 2008), these data also point to a negative correlation between fluid intelligence and resource allocation. A finding by Heitz et al. (2008), however, remains of special interest: High-spans exhibit larger pre-experimental and pre-trial baselines across all types of tasks. Following classical interpretations of pupil size as an indicator of sympathicus activity, this may indicate a higher general arousal in high-span individuals (cf. Granholm & Steinhauer, 2004). Following Aston-Jones' and Cohens' theory, however, the high-span individuals might be less engaged in the task as it is less challenging to them (tonic mode of LC activity; cf. Aston-Jones & Cohen, 2005).

Fluid Intelligence, Resource Allocation, and Pupil Dilation

The goal of the present study was twofold. First, we intended to replicate findings concerning the differential impact of fluid intelligence on the processing of easy and difficult cognitive tasks. Second, we aimed to shed light on the relation between resource allocation and fluid intelligence using a pupil dilation measure. To test this relation we adopted an extreme-groups approach: We compared individuals with high fluid intelligence scores (h-IQ) and average fluid intelligence scores (a-IQ, i.e., normal controls). We investigated performance (response times and error rates) in a cognitively low-level choice reaction time task that required a limited set of fundamental, yet simple cognitive processes (Neubauer, 1997) as compared to a cognitively high-level geometric analogy task that additionally required executive processes (Cho et al., 2007). In addition, we assessed pupil dilation as an index of resource allocation during the two cognitive tasks and during a task-free pre-experimental baseline condition.

There are three contrasting predictions about the effects of fluid intelligence and task difficulty on performance and pupil dilation (cf. Ahern & Beatty, 1979): *First*, if individuals with high fluid intelligence have more resources available and thus can solve *more demanding* tasks, they should only outperform normal controls in the most difficult analogy tasks (shorter or the same response times, lower or the same error rates, greater task-evoked pupil dilations) (*resource hypothesis*). *Second*, if individuals with high fluid intelligence *generally* invest more resources, we expect shorter or the same response times, lower or the same error rates, and *greater* tasked-evoked pupil dilations *across all types of tasks* compared to normal controls (*effort hypothesis*). *Third*, if individuals with high fluid intelligence use resources *more efficiently* than normal controls, a negative interindividual correlation between resource allocation and task-performance is expected. Consequently, we expect shorter or the same response times, lower or the same error rates, and *smaller* task-evoked pupil dilations *across all types of tasks* for individuals with high fluid intelligence compared to normal controls (*efficiency hypothesis*).

For the determination of resource allocation, the *pre-experimental pupil baseline* is of interest, too. The pre-experimental pupil baseline is assumed to index task-free exploration (Aston-Jones & Cohen, 2005). Furthermore, fluid intelligence is correlated with looking for new – that is, relevant or potentially interesting – information (Ackerman & Heggestad, 1997; Moutafi, Furnham, & Crump, 2003; Raine, Reynolds, Venables, & Mednick, 2002).

Therefore, we predict for individuals with high fluid intelligence a higher pre-experimental pupil baseline diameter compared to normal controls.

Method

Subjects

Thirty-seven students took part in the experiment (29 males and 8 females; age [mean ± SD]: 16.6 ± 0.6) and were paid for their participation. Their socio-economic backgrounds were controlled. All participants attended the 11th grade of three Berlin schools specializing in mathematics and natural sciences. All students were right-handed as assessed using the Edinburgh Handedness Inventory (Oldfield, 1971), had normal or corrected-to-normal vision, had no history of neurological or psychiatric diseases, and did not take any medications. The students and their parents gave written consent prior to investigation according to the Declaration of Helsinki (1964).

Three months prior to the experiment, all participants were screened for their fluid intelligence (F-IQ) by administering the RAPM (Heller, Kratzmeier, & Lengfelder, 1998; Raven, 1958). Participants were divided into two groups based on their RAPM scores (whole sample: F-IQ = 117.5 ± 16.9). Five female and 14 male participants were assigned to the average fluid intelligence group (F-IQ = 102.6 ± 8.5), whereas 3 female and 15 male participants were assigned to the high fluid intelligence group (F-IQ = 133.1 ± 4.7).

Tasks and Stimulus Material

The experiment consisted of three parts: the assessment of the pre-experimental pupil baseline diameter, the choice reaction time task, and the geometric analogy task.

Pre-experimental pupil baseline diameter task. This task was explained as a calibration procedure prior to any task instructions to avoid task-related expectancy effects. Participants were asked to fixate on a cross in black color on a light gray background appearing in the middle of the screen at intervals and for lengths of varying durations while we measured their baseline levels of pupil diameter.

Choice reaction time task. This low-level cognitive task consisted of 4 practice and 20 test trials. Participants were presented dots either to the left or right of a foveally displayed vertical line. The dots and the line were presented in black color on a light gray background. Participants had to decide as quickly and accurately as possible whether the dot was presented to the left or right of the vertical line.

Geometric analogy task. In this high-level cognitive task, participants were presented with stimuli quadruplets. Each quadruplet consisted of a source pair (A:A') and a target pair (B:B') of geometric chessboard-like patterns. Each pattern consisted of an 8 x 8 grid of squares with each square colored either white or black (Chipman, 1977; Offenhaus, 1983; cf. Figure 1). The stimuli quadruplets were presented on a light gray background. Six different patterns were used, each in four possible alignments: "normal," vertically mirrored, horizontally mirrored and diagonally mirrored. A pilot study had been conducted to select patterns of similar complexity. Three types of relation were applied: mirroring on the vertical, the horizontal, or the diagonal axis. These types of relation vary in difficulty (low [vertical] < medium [horizontal] < high [diagonal]; Offenhaus, 1983; Royer, 1981; van der Meer, 1996). The experiment consisted of 8 practice and 60 test items. Source pair and target pair had either the same type of relation (analogy items) or different types of relation (distractor items) (Figure 1). The same patterns were used in analogy items and distractor items. Participants had to decide as quickly and accurately as possible whether there was the same type of relation both in the source and the target pair.

Insert Figure 1 about here

Design

Choice reaction time task. The following independent variables were manipulated in the experiment: position of the dot (left vs. right of the vertical line; within subjects) and fluid intelligence (high vs. average; between subjects). Items were presented randomly. Response times (RTs, measured as the time between the dot onset and a response), error rates, and pupillary responses were recorded as dependent variables.

Geometric analogy task. The following independent variables were considered in the experiment: task difficulty based on the difficulty of the type of relation (low: mirroring on the vertical axis, medium: mirroring on the horizontal axis, high: mirroring on the diagonal axis; within subjects) and fluid intelligence (high vs. average; between subjects). Source pair and target pair had either the same type of relation (analogy items, 50%) or different types of relation (distractor items, 50%). For analogy items (n = 30), the type of relation between source and target was varied: mirroring on the vertical (n = 10), on the horizontal (n = 10), or on the diagonal axis (n = 10). Distractor items (n = 30) were included in the experiment so that participants would not only be exposed to analogy items. Items were presented in a randomized order. The following dependent variables were recorded: RTs (measured as the time between appearance of the item and the response), error rates, and pupillary responses. Note that only the data for correctly detected analogy items were further analyzed in detail, since we did not have specific hypotheses regarding the processing of distractor items.

Procedure

The experiment took place in a quiet and moderately illuminated room (background luminance 500 lux). All three phases of experimentation were performed automatically under the control of a laboratory interface system (see Apparatus). At the beginning of the experiment, participants filled out a questionnaire that ascertained demographic data as well as factors that are known to affect pupil dilation (e.g., psychiatric and neurological dysfunction, drug consumption, medication; cf. Loewenfeld, 1993). Following this background luminance adaptation, participants were seated comfortably in front of a computer screen (size of the display: 19”, display resolution: 1024 x 768) at a distance of approximately 100 cm.

Pre-experimental pupil baseline diameter. Participants were asked to fixate on a cross presented ten times with shuffled durations from 200 ms to 500 ms in steps of 50 ms resulting in a total fixation time of 3,500 ms. The interval between fixations varied between 700 ms and 1,000 ms. This procedure was repeated once after a self-paced blinking pause. The mean luminance of the stimuli was 49 cd/m^2. The individual average pupil diameter of the 2,450 ms of fixation was taken as a pre-experimental pupil baseline not influenced by any instructional effects.

Choice reaction time task. Each trial started with a fixation cross presented in the middle of the screen for 1,000 ms (pre-trial baseline phase). Following the fixation cross, a vertical line was shown in the middle of the screen. After 500 ms, a dot appeared either to the left or right of the vertical line. The mean luminance of the stimuli was 48.5 cd/m^2. Participants had to decide whether the dot was presented to the left or right of the vertical line. They were instructed to press the left button with the middle finger of the left hand if the dot appeared on the left and to press the right button with the index finger of the left hand if the dot appeared on the right. Immediately after pressing the button the next trial started. After every eight trials there was a self-paced blinking pause indicated by a smiley.

Prior to the choice reaction time task, participants received written instructions presented on the computer monitor, and completed a practice session with similar stimulus material to become familiar with the task as well as with the experimental procedure. During the practice session, feedback on the correctness of the participant’s responses was given after each trial. Overall, it took about 5 minutes to finish the choice reaction time task.

Geometric analogy task. Each trial consisted of four phases. The trial started with a fixation cross, which was presented for 1,000 ms (pre-trial baseline phase). Then, the item was presented (stimulus presentation phase). The mean luminance of the stimuli was 34.5

cd/m^2. Participants had to decide as quickly and accurately as possible whether there was the same type of relation in both the source and the target pairs. If there was, they were instructed to press the right button with the index finger of the left hand; if there was not, they were instructed to press the left button with the middle finger of the left hand. As soon as a response button was pressed by the participant, the item disappeared from the screen (relaxation phase) to prevent subsequent processing or rumination. The item was followed by a mask with the same luminance as the test items for 2,000 ms. The mask was used to ensure that the pupillary response was not disrupted or affected by changing light conditions. After the relaxation phase a smiley appeared on the screen indicating that participants were now allowed to blink and could start the next trial by pressing one of the response buttons (blinking phase). During each trial, participants were asked not to move their heads and to restrict eye blinks if possible to the blinking phase at the end of the trial.

Prior to the analogy task, participants received written instructions presented on the computer monitor, and completed a practice session with similar stimulus material to become familiar with the task as well as with the experimental procedure. During the practice session, feedback on the correctness of the participant's responses was given after each trial. Overall, it took about 20 minutes to finish the geometric analogy task.

Apparatus

Stimuli were presented using the experimental control software Presentation 9.01 (Neurobehavioral Systems Inc, Albany, CA) running on a Microsoft® Windows® XP operating system. The computer used for stimulus presentation collected the behavioral data (RTs and error rates) and was connected with another computer for registration and storage of the pupil data for offline analyses. The connection of these two computers allowed a transmission of trigger signals to mark the beginning of every trial in the experiments.

Pupillary responses were continuously recorded using an iView system (SensoMotoric Instruments GmbH, Teltow, Germany). The pupillometer (i.e., an infrared light source with

λ = 700 – 1,049 nm and a video camera sensitive to infrared light) was mounted on a stand, which stabilized the participant's head. The light source and the camera were pointed at the participant's right eye. Pupil diameter was recorded at 240 Hz. The iView system measured pupil diameter in pixels. To relate this measure to absolute pupil size, however, we used the following calibration procedure: At the beginning and the end of the experiment, a black dot (5 mm in diameter) was placed on the closed lid of the participant's right eye. This procedure made it possible to convert pupil diameter from pixels to millimeters for each participant by determining the size of this artificial pupil in pixels.

Data Analysis

Behavioral data (RTs and error rates) were analyzed using the Statistical Package for the Social Sciences 14 (SPSS Inc., Chicago, USA). Incorrect responses were excluded from data analyses. The distribution of RTs of all remaining items was determined per subject. Trials with RTs less or greater than two standard deviations from the individual's mean were excluded from the statistical analyses. For the choice reaction time task, 4.76% of the trials were eliminated, and for the geometric analogy task, 4.94%.

Pupillary responses were analyzed using Matlab 7.1 (The MathWorks, Inc., MA, USA) and SPSS 14. Prior to statistical analyses data were cleaned following standard procedures (Beatty & Lucero-Wagoner, 2000; Granholm, Asarnow, Sarkin, & Dykes, 1996; Verney, Granholm, & Dionisio, 2001). Artifacts due to excessive blinking were removed. Pupillary artifacts were not systematically distributed across experimental conditions. Very small blinks were replaced by linear interpolation. In the end (after discarding errors, outliers, and artifacts) an average of 87.1% of choice reaction time trials (h-IQ: 86.6%, a-IQ: 87.6%) and 56.4% of geometric analogy trials (h-IQ: 60.2%, a-IQ: 55.7%) remained for statistical analyses.

For each trial (choice reaction time task, geometric analogy task), the average pupil diameter of the 200 ms preceding the stimulus onset was subtracted from the task-evoked pupil diameter (pre-trial baseline correction). We then computed stimulus-locked pupillary responses for each trial and averaged the responses for each condition and participant (cf. Beatty & Lucero-Wagoner, 2000). Data were smoothed using an unweighted five point moving average filter. For each participant and condition, peak dilation of the pupillary responses was defined as the maximal dilation obtained in the measurement interval of interest between 500 ms after stimulus onset and 1,000 ms after response. This measure has the advantage of being independent of the number of data points occurring in the measurement interval (Beatty & Lucero-Wagoner, 2000). Data were expressed as mm deviation from the pre-trial baseline (peak dilation).[2] This procedure was executed for the pupil dilation of each participant and each trial. Next, the data were averaged for each participant in each condition (cf. Granholm et al., 1996; Verney et al., 2004).

Repeated measure analyses of variance (ANOVAs) for RTs, error rates, and pupillary responses were conducted after testing for normal distributions (Kolmogorov-Smirnov test). Significant main effects were further analyzed by separate t-tests. A rejection criterion of $p < .05$ (two-tailed) was chosen for all analyses (corrected for multiple comparisons).

Behavioral Results

Choice reaction time task

Individuals with high fluid intelligence (h-IQ) responded faster (RT: means (M) and standard errors (SE): M = 317.70 ms, SE = 9.51 ms) and with higher accuracy (error rate: M = 1.11%, SE = 0.50%) than individuals with average fluid intelligence (RT: M = 356.37 ms, SE = 15.79 ms; error rate: M = 1.58%, SE = 0.86%). A one-way repeated measures ANOVA with group (h-IQ vs. a-IQ) as a between-subjects factor was performed. The analysis revealed a statistically significant main effect for the group (h-IQ vs. a-IQ), $F(1,35) = 4.279$, $MSE = 3{,}229$, $p = .046$, $\eta^2 = .109$, indicating that RTs were shorter for the h-IQ group than for the a-IQ group. Error rates indicated that this result was not due to a speed/accuracy trade-off.

Geometric analogy task

Descriptive statistics are displayed in Table 1 including means (M), and standard errors (SE) of RTs and error rates for the geometric analogy task.

Insert Table 1 about here

A 2 (group: h-IQ vs. a-IQ) x 3 (task difficulty: low, medium, high) repeated-measures ANOVA on RTs and error rates was performed. The RT analysis revealed statistically significant main effects of task difficulty [$F(2,34) = 44.807$, $MSE = 9378588$, $p < .001$, $\eta^2 = .561$] and group [$F(1,35) = 5.175$, $MSE = 30343166$, $p = .029$, $\eta^2 = .129$], as well as for the interaction of task difficulty x group [$F(2,35) = 4.790$, $MS = 44923379$, $p = .035$, $\eta^2 = .120$]. RTs increased for more difficult analogy tasks, and the h-IQ group was faster than the a-IQ group (Figure 2).

Insert Figure 2 about here

Participants with high fluid intelligence, however, outperformed normal controls only for the more difficult tasks. That is, they did not solve the easiest tasks (i.e., mirroring on the vertical axis) significantly faster than normal controls, $t(35) = 1.484$, $p = .147$, $\eta^2 = .059$. Only the more difficult geometric analogy tasks were processed faster by participants with high fluid intelligence than by normal controls—mirroring on the horizontal axis: $t(35) = 2.099$, $p = .043$, $\eta^2 = .112$; mirroring on the diagonal axis: $t(35) = 2.319$, $p = .026$, $\eta^2 = .133$.

In general, our data concerning task difficulty replicate a number of recent studies (Bornstein & Krinsky, 1985; Ferguson, 2000; Offenhaus, 1983; Royer, 1981; Palmer & Hemenway, 1978; van der Meer, 1996), that is, mirroring on the diagonal axis appeared to be the most difficult type of relation, and mirroring on the vertical axis was the easiest type of relation.

The analysis of error rates revealed significant main effects of task difficulty [F(2,34) = 46.787, MSE = 191.375, p < . 001, η^2 = .572] and group [F(1,35) = 7.122, MSE = 334.47, p = .011, η^2 = .169] as well as a significant interaction effect [F(2,35) = 7.903; MS = 1512.36, p = . 008, η^2 = .184]. In general, performance accuracy decreased with increasing task difficulty. The h-IQ group made fewer errors than the a-IQ group. However, the h-IQ group only made significantly fewer errors than the a-IQ group when processing more difficult tasks (mirroring on the horizontal axis: t(35) = 2.158, p = .038, η^2 = .117; mirroring on the diagonal axis: t(35) =2.639, p =.012, η^2 = .166). For the easiest tasks, that is, mirroring on the vertical axis, error rates in participants with high fluid intelligence and normal controls did not differ, t(35) = -0.141, p = .889, η^2 = .001. These data confirm the RT results. Importantly, error rates indicated that there was no speed/accuracy trade-off in the data.

Pupillary Responses

Pre-experimental pupil baseline diameter task

Individuals with high fluid intelligence exhibited a larger pre-experimental baseline pupil diameter (M = 4.961 mm, SE = .183 mm) than individuals with average fluid intelligence (M = 4.509 mm, SE = .107 mm). A one-way ANOVA revealed a significant effect of group (h-IQ vs. a-IQ), F(1,35) = 4.679, MSE = 0.405, p = .037, η^2 = 0.118, indicating that the h-IQ group has a greater pre-experimental pupil baseline diameter than the a-IQ group. No sex differences within the two groups were found, (a-IQ: F (1,17) = 1.009, MSE = 0.219, p = .329, η^2 = 0.056; h-IQ: F (1,16) = 0.052, MSE = 0.568, p = .823, η^2 = 0.003).

Choice reaction time task

Individuals with high fluid intelligence exhibited a larger pre-trial baseline pupil diameter (M = 4.831 mm, SE = .167 mm) and a larger pupil peak dilation (M = .317 mm, SE

= .034 mm) than individuals with average fluid intelligence (pre-trial baseline pupil diameter: M = 4.416 mm, SE = .105 mm; pupil peak dilation: M = .282 mm, SE = .027 mm). Figure 3 illustrates the pupillographic waveforms for the choice reaction time task.

Insert Figure 3 about here

A one-way ANOVA revealed a significant effect of group (h-IQ vs. a.IQ) for the mean pre-trial baseline pupil diameter, $F(1,35) = 4.530$, $MSE = 0.351$, $p = .040$, $\eta^2 = 0.115$. The h-IQ group had a greater pre-trial baseline pupil diameter than the a-IQ group. For the peak dilation, however, the ANOVA revealed no effect of group, $F(1,35) = 0.647$, $MSE = 0.017$, $p = .427$, $\eta^2 = 0.018$.

Geometric analogy task

Descriptive statistics are displayed in Table 2 and include means (M), and standard errors (SE) for pupil diameter (pre-trial baseline, peak dilation) in this high-level cognitive task.

Insert Table 2 about here

A one-way ANOVA revealed no significant effect of group (h-IQ vs. a-IQ) on the mean pre-trial baseline pupil diameter, $F(1,35) = 0.806$, $MSE = 1.137$, $p = .375$, $\eta^2 = 0.023$. That is, for the geometric analogy task, the h-IQ group and the a-IQ group did not differ in pre-trial baseline diameters.

For pupil dilation in analogy-items, a 2 (group: h-IQ vs. a-IQ) x 3 (task difficulty: low, medium, high) repeated-measures ANOVA was performed. There was a significant main effect of group, $F(1,35) = 8.453$, $MSE = 0.063$, $p = .006$, $\eta^2 = 0.195$, that is, pupil peak dilation was greater in the h-IQ group than in the a-IQ group. There was no effect of task difficulty, $\underline{F}(2,34) = 0.135$, $\underline{MSE} = 0.002$, $\underline{p} = .874$, $\eta^2 = 0.004$, and no group x task difficulty interaction, $\underline{F}(2,35) = 1.182$, $\underline{MSE} = 0.017$, $\underline{p} = .313$, $\eta^2 = 0.033$.

In line with our hypotheses, we examined the group x task difficulty interaction more closely. The *resource hypothesis* predicted that group differences would be most pronounced on the most difficult trials. We therefore analyzed the different levels of task difficulty separately, using one-way ANOVAs. The analysis yielded a significant effect with higher peak dilation for the most difficult trials (mirroring on the diagonal axis) for the h-IQ group compared to the a-IQ group, $\underline{F}(1,35) = 11.703$, $\underline{MSE} = 0.027$, $\underline{p} = .002$, $\eta^2 = 0.251$. For the easier trials —mirroring on the vertical and on the horizontal axis— the group differences did not reach significance (mirroring on the vertical axis: $\underline{F}(1,35) = 3.884$, $\underline{MSE} = 0.025$, $\underline{p} = .057$, $\eta^2 = 0.100$; mirroring on the horizontal axis: $\underline{F}(1,35) = 3.785$, $\underline{MSE} = 0.040$, $\underline{p} = .060$, $\eta^2 = 0.0098$]. Figure 4 illustrates these findings. Taken together, the pupil data show that the h-IQ group allocated more resources than the a-IQ group in solving the most difficult geometric analogies (mirroring on the diagonal axis): Higher processing load is reflected in higher peak dilation for the h-IQ group.

Furthermore, we controlled for sex differences and differences in the early periods of pupil dilation. First, no sex differences regarding task-evoked pupillary dilations within the two groups were found; h-IQ: $\underline{F}(1,16) = 0.150$, $\underline{MSE} = 0.089$, $\underline{p} = .703$, $\eta^2 = 0.009$; a-IQ: $\underline{F}(1,17) = 1.069$, $\underline{MSE} = 0.043$, $\underline{p} = .316$, $\eta^2 = 0.059$. Thus, our findings are independent of sex. Second, we ran a Principal Component Analysis (PCA) to examine effects in the first 2 seconds of the geometric analogy task separately from the later period. For each level of task difficulty (low, medium, high), the analysis revealed five factors. A 2 (group: h-IQ vs. a-IQ) x 5 (factor) repeated-measures ANOVA for each level of task difficulty was performed. The analysis yielded no significant effect of factor or of group, and no significant interaction.[3] Thus, the lack of group differences in the early periods of pupil dilation indicates that our findings reflect cognitive processing rather than spontaneous emotional responses to the stimuli (cf. Compton, Banich, Mohanty, Milham, Herrington, & Miller et al., 2003; Liddell, Brown, Kemp, Barton, Das, & Peduto et al., 2005; Ochsner & Feldman Barrett, 2001; Phelps, 2006; Prehn, Heekeren, Blasek, Lapschies, Mews, & van der Meer, 2008).

Insert Figure 4 about here

Changes in Pupil Baseline Diameters.

There is a decrease in pupil baseline diameter from the beginning to the end of the whole test session which differs between the h-IQ and the a-IQ groups. In the h-IQ group the mean geometric analogy pre-trial baseline diameter is significantly smaller than the pre-experimental baseline diameter, $t(17) = 4.713$, $p = .000$, and significantly smaller than the mean choice reaction time pre-trial baseline diameter, $t(17) = 4.336$, $p = .000$. In the a-IQ group the geometric analogy pre-trial baseline diameter is significantly smaller than the pre-experimental baseline diameter, $t(18) = 2.789$, $p = .012$. Interestingly, the decrease in baseline diameter from the pre-experimental condition to the geometric analogy task is significantly higher in the h-IQ group than in the a-IQ group, $F(1,35) = 0.647$, $MSE = 0.149$, $p = .023$, $\eta^2 = 0.139$ (note that there is no difference between groups in pre-trial baseline pupil diameter in the geometric analogy task, see above).

Discussion

We used a choice reaction time task and a geometric analogy task to investigate the processing of low-level (easy) versus high-level (difficult) cognitive tasks in individuals with high fluid intelligence compared to normal controls. We recorded behavioral data (i.e., RTs and error rates) indexing speed and accuracy of task processing as well as phasic changes in pupil diameter indexing task-evoked mental resource allocation. Additionally, we examined the pre-experimental pupil baseline diameter indexing the general task-free resource allocation (cf. Aston-Jones & Cohen, 2005; Granholm & Steinhauer, 2004).

The study yielded the following main findings. First, individuals with high fluid intelligence processed the cognitively low-level choice reaction time task faster than normal controls. Task-evoked pupillary responses, however, did not differ between the groups. Second, we found that individuals with high fluid intelligence processed the cognitively high-level geometric analogy task faster, more accurately, and with greater pupillary responses than normal controls in the more difficult task conditions only. Furthermore, individuals with high fluid intelligence showed greater pre-experimental pupil baseline diameters than normal controls. Taken together, our results demonstrate that individuals with high fluid intelligence allocate more resources than normal controls in processing the most difficult cognitive problems. Additionally, individuals with high fluid intelligence seem to allocate more resources toward exploring the given environment, even if there is no task at hand and the experimental tasks are not yet introduced.

Impact of fluid intelligence on processing low-level vs. high-level cognitive tasks

The first goal of the present study was to replicate findings concerning the differential impact of fluid intelligence on processing a low-level cognitive task (inspection time; Neubauer, 1997) as compared to a high-level cognitive task, namely, solving geometric analogies. The finding that individuals with high fluid intelligence performed the simple choice reaction time task significantly faster than normal controls did achieve this. This result points to a higher processing efficiency in individuals with high fluid intelligence. In a recent study, Salthouse and Pink (2008) asked for the critical factor in the relationship between fluid intelligence and working memory. Because strong influences were apparent in the simplest versions and on the initial trials in their working memory tasks, the critical factor was not assumed to be related to how much storage and processing was required, or to processes associated with successive trials in these tasks. Instead, the critical factor might be to quickly adapt to a new task and to perform effectively, “even in situations that have minimal demands for simultaneous storage and processing” (Salthouse & Pink, 2008, p. 370). Barrouillet, Lépine, and Camos (2008) extended this view in presenting empirical evidence that any elementary attention-demanding processing step is sensitive to variations in working memory capacity. The differences between individuals differing in working memory capacity observed on complex cognitive activities were exactly proportionate to those elicited by elementary activities. That is, the time to perform each processing step is assumed to “depend on a basic general capacity, conceived as the amount of available attention needed to activate relevant items of knowledge and procedures” (Barrouillet et al., 2008, p. 533). This conclusion corresponds with our findings in the choice reaction time task. This low-level cognitive task requires the participant to quickly detect the position of a critical stimulus. The effect of high fluid intelligence appears to make the accessing of items faster, that is, more efficient.

High fluid intelligence also leads to shorter response times and lower error rates in processing the high-level geometric analogy task. However, this was only found to be significant for the more difficult analogy trials (mirroring on the diagonal axis). This finding suggests that individuals with high fluid intelligence do not necessarily clearly outperform normal controls in a cognitive problem, which is easily managed by individuals with average fluid intelligence, too. There are two explanations: First, as might be expected for the easiest trials (mirroring on the vertical axis; cf. Offenhaus, 1983; Royer, 1981; van der Meer, 1996) the groups do not differ in applying the global set of fundamental cognitive processes required in analogical reasoning (cf. Cho et al., 2007). This explanation, however, contradicts the findings of Salthouse and Pink (2008) and Barrouillet et al. (2008). Therefore, a second explanation should be taken into account. Considering the remarkable variances in RTs between participants, we assume that potential group differences in processing the easier trials may have been masked by different strategies that individuals use to perform the task (cf. van der Meer, 1996). For example, Vigneau, Caissie, and Bors (2006) explored strategic influences on performance in a fluid intelligence task in more detail. They presented individuals differing in fluid intelligence a selection of items from the RAPM. Latency and eye-movement data showed that individuals differed in terms of speed, but also in terms of strategies. Consequently, the impacts of visual scanning strategies on performance in visually presented cognitive tasks should be considered in more detail in future research.

Modulation of resource allocation in individuals with high versus average fluid intelligence

The second goal of the present study was to investigate the modulation of resource allocation in individuals with high versus average fluid intelligence while performing cognitively low-level choice reaction time tasks as compared to cognitively high-level geometric analogy tasks. Pupillometrics was used to shed light on resource allocation. Given the large group differences in fluid intelligence indicated by the RAPM scores, the pupillary response was expected to differentiate between the hypotheses outlined in the introduction, namely the *resource*, *effort*, and *efficiency hypotheses*. Measures of both *phasic* and *tonic* pupil dilation helped in contrasting these hypotheses.

Phasic pupillary response

The phasic pupillary responses indicated that individuals with high fluid intelligence allocated more resources than normal controls only for the most difficult task. Thus, the *effort hypothesis* was not supported. For the low-level choice reaction time task, the task-evoked pupillary responses did not differ between normal controls and individuals with high fluid intelligence. Since individuals with high fluid intelligence had significantly shorter RTs than normal controls, this finding points to a higher efficiency in h-IQ individuals compared to a-IQ individuals during the choice reaction time task. Here, we argue that component processes of this task might be more automated in individuals with high fluid intelligence. For the cognitively high-level geometric analogy task, high fluid intelligence also led to shorter RTs and lower error rates. However, this was only found to be significant for the most difficult trials (mirroring on the diagonal axis). Most importantly, higher speed and accuracy of h-IQ individuals corresponded with stronger task-evoked pupillary responses. That is, individuals with higher fluid intelligence allocate more resources compared to individuals of average fluid intelligence only when processing the most difficult geometric analogies. Consequently, our data clearly support the *resource hypothesis*. Individuals scoring high in fluid intelligence appear to have more resources available and thus perform better on more demanding tasks. These findings also correspond to the neuroimaging results found by Duncan (2003) and others (Gray et al., 2003; Lee, Choi, Gray, Cho, Chae, & Lee et al., 2006; O'Boyle, Cunnington, Silk, Vaughan, Jackson, & Syngeniotis et al., 2005) who found a positive correlation between regional brain activation and intelligence (but see also Rypma et al., 2006, for a critical discussion).

Our results are, however, not consistent with findings by Ahern and Beatty (1979) who reported smaller pupillary responses in more intelligent individuals. These different patterns of results may occur for a number of reasons (cf. Rypma et al., 2006). First, Ahern and Beatty (1979) presented multiplication tasks of differing complexity. For students, these tasks are highly overlearned, that is, the component processes of arithmetic are assumed to be more automatic than the fluid processing required in the newly encountered analogical reasoning task. This is especially true for the most difficult geometric analogies, which appeared to best distinguish between individuals differing in fluid intelligence. Also note that the stimuli in the Ahern and Beatty study were presented acoustically and sequentially, whereas our stimuli were presented visually and simultaneously.

Second, the differences between experimental populations may have influenced the results. In contrast to our approach, Ahern and Beatty (1979) divided their participants based on SAT scores. The SAT is a standardized test for college admission in the U.S.A. that does not purely measure fluid intelligence, but rather measures proficiencies (for example in mathematics and writing). This supports our assumption that superior performance of the “high group” in the Ahern and Beatty study was partly due to better trained skills (and thus more to automatic processes) rather than to fluid intelligence. Moreover, as SAT scores are influenced by training and preparation, the fluid intelligence of the “high group” may actually have been lower than that of the h-IQ group in our study that was selected by a fluid intelligence test (RAPM).

Altogether, we see sufficient evidence to argue that Ahern and Beatty’s (1979) favoring of the *efficiency hypothesis* is the result of a considerably different experimental design. We believe that during the process of learning there could be a larger increase in efficiency in intelligent/proficient subjects, whereas superior performance on an unknown task (such as ours) is initially administered by additional allocation of resources. In line with this assumption, in a pre/post training design Neubauer, Grabner, Freudenthaler, Beckmann, and Guthke (2004) reported a negative correlation between fluid intelligence and prefrontal brain activation during the post-test *only*. We deem the impact of learning on resource allocation to be an interesting area for future research.

Another conclusion of our pupillary data refers to the **interaction** between fluid intelligence and subjective task difficulty as indicated by phasic pupillary responses. As mentioned before, the most difficult trials —mirroring on the diagonal axis— induced the largest difference in pupil dilation between individuals with high and average fluid intelligence. Granholm et al. (1996) used pupillometric recordings during a digit span recall task that differed in processing load. The authors found that pupillary responses increase systematically with increasing processing demands that are below resource limits, change little during active processing at or near the resource limits, and decline when processing demands exceed available resources. Similarly, in our study, the most difficult trials of the analogy task may have overstrained the resources of the individuals in the a-IQ group. This is suggested by both the dramatically increased error rate and a decrease in pupillary dilation as compared to the easier trials. Note that these findings add further support to the *resource hypothesis*. They are also of great value in explaining the differential findings of Ahern and Beatty (1979) as their tasks probably did not exceed the cognitive capacities of the individuals in their “low group.”

Tonic pupillary response

Tonic pupil size also proved to be sensitive to fluid intelligence. This concerns the **pre-experimental pupil baseline**, which was larger for individuals with high fluid intelligence. The contributions of the autonomic nervous system to pupil dilation have been known for some time (cf. Loewenfeld, 1993), and this suggests an interpretation in terms of general arousal: The dilation of the pupil is mediated by activation of the sympathetic dilator muscle as well as inhibition of the parasympathetic sphincter. Accordingly, a tonically dilated pupil is typically associated with wakefulness and activation. Related psychological concepts (e.g., stimulation-seeking or the personality trait “openness to experience”) have been shown to be positively correlated with fluid intelligence (Ackerman & Heggestad, 1997; Moutafi et al., 2003) and even to promote the development of cognitive abilities (cf. the longitudinal study by Raine et al., 2002). Following this line of reasoning, the larger baseline can be seen as an indicator of a more pronounced tendency toward task-free exploring and scanning of the environment in the h-IQ group. This finding is comparable to the findings of Heitz et al. (2008), who report greater pupil baseline diameters for individuals with high working memory capacity as compared to individuals with low working memory capacity.

A related though more elaborate view on the interplay of arousal and performance has been proposed by Aston-Jones and Cohen (2005). Their theory allows for a differentiated dealing with overall activation and task-performance, as it also accounts for individual differences in task-engagement, and elegantly incorporates the Yerkes-Dodson (Yerkes & Dodson, 1908) relationship as discussed in more detail in the introduction section (p. 5). Still, in case of pre-experimental baseline differences, this leads to a similar interpretation for the understanding of arousal in terms of activation in the autonomic nervous system: According to Aston-Jones and Cohen (2005), a large tonic pupil diameter reflects *exploratory behavior*, that is, the individual is scanning the environment for possible sources of reward.

Interestingly, the **pre-trial pupil baselines** (i.e., baseline measurements recorded before the beginning of each trial) show a striking difference between the high intelligence group and the normal controls, too. In the low-level choice reaction time task we found a significantly enlarged pre-trial baseline for h-IQ individuals. However, we found a downward trend in pre-trial pupillary data: In the geometric analogy task, tonic pupillary baselines were similar for both groups. Aston-Jones' and Cohen's (2005) theory provides a satisfactory explanation for this result: It is only the difficult task that is demanding enough for the h-IQ group to display a comparably strong task-engagement as the a-IQ individuals. Since the order of the tasks was not permutated in our study, we cannot exclude the possibility that the decrease in pupil baseline in the h-IQ group was due to a drop in autonomic arousal over the course of the experimental procedure irrespective of the administered tasks. Following this explanation, one would still have to explain why this drop was more pronounced in the h-IQ group than in the a-IQ group. Future studies should consider this in their experimental design.

Conclusion

Our study makes the crucial point that the combination of pupillometrics with traditional behavioral measures is promising as a way to assist our understanding of fluid intelligence and resource allocation in cognitive processing. Our results support the *resource hypothesis,* that is, highly fluid intelligent individuals have more resources available than averagely fluid intelligent individuals and allocate them if the tasks become sufficiently demanding. This finding is consistent with Heitz et al. (2008), and speaks against the *effort hypothesis:* Highly intelligent individuals do not invest more resources across all types of tasks, but only in those that are demanding enough to require the allocation of additional resources. Our results contradict the classical findings from Ahern and Beatty (1979) who found highly intelligent individuals to allocate less resources in solving cognitive tasks (*efficiency hypothesis*). We have argued that these differential findings might be explained by differences in the employed cognitive tasks and in the investigated populations. Finally, high fluid intelligence is in line with higher tonic pupil size in situations and tasks without or with only limited processing requirements.

Future studies will need to investigate the impact of task type and learning on the allocation of mental resources in more detail. In particular, learning induced improvements and automatization of cognitive functions might be crucial for the relationship of task performance and resource allocation (cf. Neubauer et al., 2004; Poldrack, Desmond, Glover, & Gabrieli, 1998), as the coming-into-effect of such mechanisms could mark a transition from *resource* to *efficiency* explanations. Finally, individual processing strategies should be taken into account. In future fMRI studies, multiple cognitive tasks in individuals differing in fluid intelligence should be employed to further investigate the dynamic nature of resource allocation and the contribution of specific neural networks in the service of resource modulation and cognition (Critchley, Tang, Glaser, Butterworth, & Dolan, 2005; Grabner, Ansari, Reishofer, Stern, Ebner, & Neuper, 2007; Krueger, Spampinato, Pardini, Pajevic, Wood, & Weiss et al., 2008; O'Boyle et al., 2005; Satterthwaite, Green, Myerson, Parker, Ramaratnam, & Buckner, 2007).

References

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219–245.

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science, 205*(4412), 1289–1292.

Alderton, D. A., & Larson, G. E. (1990). Dimensionality of Raven's Advanced Progressive Matrices items. *Educational & Psychological Measurement, 50*, 887–900.

Aston-Jones, G., & Cohen, J. D. (2005). An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience, 28*, 403–450.

Barrouillet, P., Lépine, R., & Camos, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review, 15*(3), 528–534.

Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: the relationship of inspection time and openness to g and Gc. *Intelligence, 31*(3), 275–287.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5*(10), 371–372.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology (2nd ed.)* (pp. 142–162). New York, NY US: Cambridge University Press.

Bornstein, M. H., & Krinsky, S. J. (1985). Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes. *Journal of Experimental Child Psychology, 39*(1), 1–19.

Bors, D. A., & Stokes, T. L. (1998). Raven´s Advanced Progressive Matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement*, *58,* 382.

Broadhurst, P. L. (1959). The interaction of task difficulty and motivation: The Yerkes Dodson law revived. *Acta Psychologica, Amsterdam, 16*, 321–338.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404–431.

Chipman, S. F. (1977). Complexity and structure in visual patterns. *Journal of Experimental Psychology: General, 106*(3), 269–301.

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition, 35*(6), 1445–1455.

Choi, Y. Y., Shamosh, N. A., Cho, S. H., DeYoung, C. G., Lee, M. J., & Lee, J.-M., et al. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *Journal of Neuroscience., 28*(41), 10323–10329.

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., & Miller, G. A., et al. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. *Cognitive, Affective & Behavioral Neuroscience, 3*(2), 81–96.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., & Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *Neuroimage, 27*(4), 885–895.

Duncan, J. (2003). Intelligence tests predict brain response to demanding task events. *Nature Neuroscience, 6*(3), 207–208.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, Andrew R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309–331.

Ferguson, R. W. (2000). Modeling orientation effects in symmetry detection: The role of visual structure. In *Proceedings of the 22nd Conference of the Cognitive Science Society*. Hillsdale, New Jersey: Erlbaum.

French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Sciences*, *6*(5), 200–205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science: A Multidisciplinary Journal, 7*(2), 155–170.

Gilzenrat, M.S., Cohen, J.D., Rajkowski, J., & Aston-Jones, G. (2003). Pupil dynamics predict changes in task engagement mediated by locus coeruleus. *Society of Neurosciences, Abstracts*, 515.19.

Grabner, R. H., Ansari, D., Reishofer, G., Stern, E., Ebner, F., & Neuper, C. (2007). Individual differences in mathematical competence predict parietal brain activation during mental calculation. *Neuroimage, 38*(2), 346–356.

Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology, 52*(1), 1–6.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*(4), 457–461.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*(3), 316–322.

Haier, R. J., Sternberg, R. J., Lautrey, J., & Lubart, T. I. (2003). Brain imaging studies of intelligence: Individual differences and neurobiology. In *Models of Intelligence for the New Millennium* (pp. 185–193). Washington, DC: American Psychological Association.

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development, 35*(4), 193–217.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45*(1), 119–129.

Heller, K. A., Kratzmeier, H., & Lengfelder, A. (1998). *Matrizen-Test-Manual (Band 2).: Ein Handbuch mit deutschen Normen zu den Advanced Progressive Matrices von J.C.Raven.* Weinheim: Beltz-Test GmbH.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*(3611), 1190–1192.

Hoeks, B., & Ellenbroek, B. A. (1993). A neural basis for a quantitative pupillary model. *Journal of Psychophysiology, 7*(4), 315–324.

Hofstadter, D. R. (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought.* New York, NY, US: Basic Books.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought.* Cambridge, MA: The MIT Press.

James, W. (1890/1950). *The principles of psychology.* New York: Dover.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT, US: Praeger Publishers/Greenwood Publishing Group.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science, 4*(1), 56–88.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1967). Pupillary responses in a pitch-diskrimination task. *Perception & Psychophysics, 2*(3), 101–105.

Klix (1993). *Erwachendes Denken.* Heidelberg: Spektrum.

Krueger, F., Spampinato, M. V., Pardini, M., Pajevic, S., Wood, J. N., & Weiss, G. H., et al. (2008). Integral calculus problem solving: An fMRI investigation. *Neuroreport: For Rapid Communication of Neuroscience Research, 19*(11), 1095–1099.

Lee, K. H., Choi, Y. Y., Gray, J. R., Cho, S. H., Chae, J.-H., & Lee, S., et al. (2006). Neural correlates of superior intelligence: Stronger recruitment of posterior parietal cortex. *Neuroimage*, *29*(2), 578–586.

Liddell, B. J., Brown, K. J., Kemp, A. H., Barton, M. J., Das, P., & Peduto, A., et al. (2005). A direct brainstem-amygdala-cortical ‘alarm’ system for subliminal signals of fear. *Neuroimage*, *24*(1), 235–243.

Loewenfeld, I. E. (1993). *The pupil.* Iowa State University Press: Ames.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*(2), 107–127.

McCrory, C., & Cooper, C. (2005). The relationship between three auditory inspection time tasks and general intelligence. *Personality and Individual Differences, 38*(8), 1835–1845.

Mitchell, M. (1993). *Analogy-making as perception: a computer model.* Cambridge, MA: The MIT Press.

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the Neo personality inventory and the Myers-Briggs type indicator. *European Journal of Personality, 17*(1), 79–94.

Neubauer, A. C. (1997). The mental speed approach to the assessment of intelligence. In J. Kingma, W. Tomic, J. Kingma, & W. Tomic (Eds.), *Advances in cognition and* *educational practice: Reflections on the concept of intelligence, Vol. 4* (pp. 149–173). US: Elsevier Science/JAI Press.

Neubauer, A. C., Grabner, R. H., Freudenthaler, H. H., Beckmann, J. F., & Guthke, J. (2004). Intelligence and individual differences in becoming neurally efficient. *Acta Psychol.(Amst), 116*(1), 55–74.

Neubauer, A., Freudenthaler, H. H., & Pfurtscheller, G. (1995). Intelligence and spatiotemporal patterns of event-related desynchronization (ERD). *Intelligence, 20*(3), 249–266.

Nuthmann, A., & van der Meer, E. (2005). Time's arrow and pupillary response. *Psychophysiology, 42*(3), 306–317.

O'Boyle, M. W., Cunnington, R., Silk, T. J., Vaughan, D., Jackson, G., & Syngeniotis, A., et al. (2005). Mathematically gifted male adolescents activate a unique brain network during mental rotation. *Cognitive Brain Research, 25*(2), 583–587.

Ochsner, K. N., & Feldman Barrett, L. (2001). A multiprocess perspective on the neuroscience of emotion. In T. J. Mayne, G. A. Bonanno, T. J. Mayne, & G. A. Bonanno (Eds.), *Emotions: Currrent issues and future directions* (pp. 38–81). New York, NY US: Guilford Press.

Offenhaus, B. (1983) *Analyse des analogen Schlussprozesses unter differentiellem Aspekt. Dissertation.* Berlin: Humboldt-Universität.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Palmer, S. E., & Hemenway, K. (1978). Orientation and symmetry: Effects of multiple, rotational, and near symmetries. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 691–702.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology, 57*, 27–53.

Poldrack, R. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1998). The neural basis of visual skill learning: An fMRI study of mirror reading. *Cerebral Cortex*, *8*(1), 1–10.

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology, 60*(2), 211–229.

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural Substrates of Fluid Reasoning: An fMRI Study of Neocortical Activation during Performance of the Raven's Progressive Matrices Test. *Cognitive Psychology, 33*(1), 43–63.

Prehn, K., Heekeren, H. R., Blasek, K., Lapschies, K., Mews, I., & van der Meer, E. (2008). Neuroticism influences pupillary responses during an emotional interference task. *International Journal of Psychophysiology, 70*(1), 40–49.

Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher g-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence, 33*(2), 203–213.

Raine, A., Reynolds, C., Venables, P. H., & Mednick, S. A. (2002). Stimulation seeking and intelligence: A prospective longitudinal study. *Journal of Personality and Social Psychology, 82*(4), 663–674.

Raisig, S., Welke, T., Hagendorf, H., & van der Meer, E. (2007). Investigating dimensional organization in scripts using the pupillary response. *Psychophysiology, 44*(6), 864–873.

Rajkowski J., Kubiak P., Aston-Jones G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society of Neurosciences, Abstracts*, 19:974

Raven, J. C. (1958). *Advanced progressive matrices.* London: Lewis.

Rindermann, H., & Neubauer, A. C. (2001). The influence of personality on three aspects of cognitive performance: Processing speed, intelligence and school performance. *Personality and Individual Differences, 30*(5), 829–842.

Royer, F. L. (1981). Detection of symmetry. *Journal of Experimental Psychology: Human Perception and Performance, 7*(6), 1186–1210.

Rypma, B., Berger, J. S., Prabhakaran, V., Bly, B. M., Kimberg, D. Y., & Biswal, B. B., et al. (2006). Neural correlates of cognitive efficiency. *Neuroimage, 33*(3), 969–979.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review, 15*(2), 364–371.

Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., & Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *Neuroimage, 37*(3), 1017–1031.

Schweizer, K., Goldhammer, F., Rauch, W., & Moosbrugger, H. (2007). On the validity of Raven's matrices test: Does spatial ability contribute to performance? *Personality and Individual Differences, 43*(8), 1998–2010.

Snow, R. E., Kyllonen, C. P., Marshalek, B., & Sternberg, R. J. (1984). The topography of ability and learning correlations. In *Advances in the psychology of human intelligence* (pp. 47–103). Hillsdale, NJ: Erlbaum.

Spearman, C. (1904). 'General intelligence,' objectively determined and measured. *American Journal of Psychology, 15*(2), 201–293.

Tamez, E., Myerson, J., & Hale, S. (2008). Learning, working memory, and intelligence revisited: Proceedings of the Meeting of the Society for the Quantitative Analyses of Behavior (SQAB 2007), 2007, San Diego, CA, USA. *Behavioural Processes, 78*(2), 240–245.

Thoma, R. J., Yeo, R. A., Gangestad, S. W., Halgren, E., Sanchez, N. M., & Lewine, J. D. (2005). Cortical volume and developmental instability are independent predictors of general intellectual ability. *Intelligence*, *33*(1), 27–38.

van der Meer, E. (1996). Memory and Analogical Reasoning. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M. K. Johnson (Eds.), *Basic and Applied Memory Research: Practical Applications* (pp. 139–151). Mahwah NJ: Lawrence Erlbaum Associates.

van der Meer, E., & Klix, F. (1986). Mathematical Giftedness - Its Nature and Possible Early Identification. In A. J. Cropley, K. K. Urban, H. Wagner, & W. Wieczerkowski (Eds.), *Giftedness: A continuing worldwide challenge.* New York: Trillium Press.

Verney, S. P., Granholm, E., & Dionisio, D. P. (2001). Pupillary responses and processing resources on the visual backward masking task. *Psychophysiology, 38*(1), 76–83.

Verney, S. P., Granholm, E., & Marshall, S. P. (2004). Pupillary responses on the visual backward masking task reflect general cognitive ability. *International Journal of Psychophysiology, 52*(1), 23–36.

Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence, 7*(1), 53–70.

Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence, 34,* 261–272.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology, 18*, 459–482.

Acknowledgments

This research was supported by grants from the German Federal Ministry for Education and Research (NIL-I/II and Berlin NeuroImaging Center). IW is funded by the Stifterverband für die Deutsche Wissenschaft, Claussen-Simon-Stiftung. This work benefited from conversations with Herbert Hagendorf, Frank Krueger, and Steffen Landgraf.

The authors wish to acknowledge the assistance of Martin Brucks, Franziska Preusse, Roman Purkhart, Susanne Raisig, Torsten Schiemann, and Dorothea Ullwer.

Footnotes

[1] Raven´s Advanced Progessive Matrices (RAPM) are frequently used as a measure of fluid intelligence (e.g., Bates & Shieles, 2003; Haier, Sternberg, Lautrey, & Lubart, 2003; McCrory & Cooper, 2005; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997; Prokosch, Yeo, & Miller, 2005; Tamez, Myerson, & Hale, 2008; Thoma, Yeo, Gangestad, Halgren, Sanchez, & Lewine, 2005). According to Carpenter, Just, and Shell, (1990), the RAPM assesses analytical intelligence, which equals Cattell's concept of fluid intelligence as the "ability to reason and solve problems involving new information" (Carpenter et al., 1990, S.404). Furthermore, Schweizer, Goldhammer, Rauch, and Moosbrugger (2007) have analyzed whether the RAPM measures fluid intelligence exclusively or also partially measures spatial ability. By means of structural equation modelling they confirmed that RAPM "can be considered as a marker of fluid intelligence as well as of figural reasoning" (p. 2009).

Various studies have shown that the RAPM has the highest loading on Spearman's general factor of intelligence (Alderton & Larson, 1990; Bors & Stokes, 1998; Marshalek, Lohman, & Snow, 1983; Snow, Kyllonen, Marshalek, & Sternberg, 1984). Nevertheless, we do not assume that they are identical. We argue that fluid intelligence and g are closely linked, but with respect to neuroscientific findings (Choi et al., 2008) they cannot be considered identical. By investigating neural correlates of intelligence at the structural and functional level, Choi, Shamosh, Cho, DeYoung, Lee, and Lee et al. (2008) pointed out that different components of g, in particular fluid and crystallized components, are distinguishable in brain function and structure.

[2] The evidence indicates that the extent of the pupil dilation evoked by cognitive processing is independent of baseline pupillary diameter for baseline values smaller than 7 mm (Hoeks & Ellenbroek, 1993). Still, as a control analysis, the relative peak dilation was also calculated in our study. It yielded the same results compared to absolute peak dilation.

[3] For each level of task difficulty (low, medium, high), the PCA revealed five factors. A 2 (group: h-IQ vs. a-IQ) x 5 (factor) repeated-measures ANOVA for each level of task difficulty was performed. The analysis yielded no significant effect of factor or of group, and no significant interaction (low task difficulty: factor [$F(2,34) = 0.000$, $MSE = 1.018$, $p = 1.000$, $\eta^2 = .000$], group [$F(1,35) = 3.145$, $MSE = 0.920$, $p = .060$, $\eta^2 = .106$], factor x group [$F(2,35) = 0.379$, $MS = 1.018$, $p = .823$, $\eta^2 = .011$]; medium task difficulty: factor [$F(2,34) = 0.001$, $MSE = 1.004$, $p = 1.000$, $\eta^2 = .000$], group [$F(1,35) = 0.127$, $MSE = 1.025$, $p = .724$, $\eta^2 = .127$], interaction factor x group [$F(2,35) = 0.840$, $MS = 1.004$, $p = .502$, $\eta^2 = .023$]; high task difficulty: factor [$F(2,34) = 0.001$, $MSE = 1.002$, $p = 1.000$, $\eta^2 = .000$], group [$F(1,35) = 1.428$, $MSE = 0.988$, $p = .240$, $\eta^2 = .039$], factor x group [$F(2,35) = 0.938$, $MSE = 1.002$, $p = .444$, $\eta^2 = .026$]).

Table 1. Geometric Analogy Task. Means (M), and Standard Errors (SE) of Response Times (in Milliseconds, ms), and Error Rates (RF: Relative Frequencies in %) Dependent on Fluid Intelligence (Measured by RAPM) and Task Difficulty (Dependent on Type of Relation: Low = Mirroring on the Vertical Axis, Medium = Mirroring on the Horizontal Axis, High = Mirroring on the Diagonal Axis). h-IQ: Individuals with High Fluid Intelligence; a-IQ: Individuals with Average Fluid Intelligence.

	Fluid intelligence			
	h-IQ		a-IQ	
Condition:	Analogy items	Distractor items	Analogy items	Distractor items
Response times				
M (ms)	8003	6804	10382	8745
SE (ms)	495.98	347.55	902.77	654.34
Error rates				
RF (%)	12.34	7.65	21.61	9.34
SE (%)	2.39	1.27	2.51	1.37

	h-IQ			a-IQ		
Task difficulty	Low	Medium	High	Low	Medium	High
Response times						
M (ms)	6658	7483	9867	7695	9432	14021
SE (ms)	412.06	536.25	783.77	556.25	746.91	1576.50
Error rates						
RF (%)	6.79	10.49	19.75	6.37	21.05	37.42[1]
SE (%)	1.82	2.45	4.71	2.29	4.15	4.47

Table 2. Geometric Analogy Task. Means (M), and Standard Errors (SE) for Pre-trial Pupil Baseline Diameter and Pupil Peak Dilation Dependent on Fluid Intelligence (Measured by RAPM), Condition (Analogy-Items vs. Distractor-Items), and Task Difficulty (Dependent on Type of Relation). h-IQ: Individuals with High Fluid Intelligence; a-IQ: Individuals with Average Fluid Intelligence.

	Fluid intelligence			
	h-IQ		a-IQ	
Condition:	Analogy items	Distractor items	Analogy items	Distractor items
Peak dilation				
M (mm)	0.544	0.477	0.401	0.374
SE (mm)	0.040	0.031	0.028	0.045
Pre-trial baseline				
Pupil diameter				
M (mm)	4.451	4.468	4.269	4.262
SE (mm)	0.172	0.174	0.110	0.114

	h-IQ			a-IQ		
Task difficulty	Low	Medium	High	Low	Medium	High
Peak dilation						
M (mm)	0.526	0.531	0.575	0.425	0.403	0.389
SE (mm)	0.041	0.048	0.049	0.032	0.046	0.026
Pre-trial baseline						
Pupil diameter						
M (mm)	4.434	4.441	4.478	4.224	4.284	4.299
SE (mm)	0.167	0.179	0.173	0.122	0.104	0.109

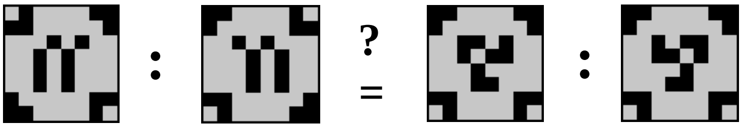
Figure Captions

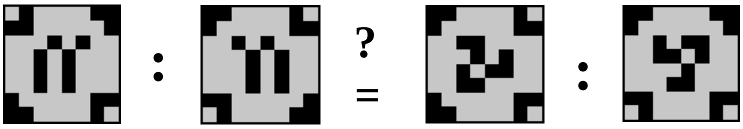
Figure 1. Geometric analogies. Examples of an analogy item (mirroring on the vertical axis) and a distractor item (mirroring on the vertical axis vs. mirroring on the diagonal axis).

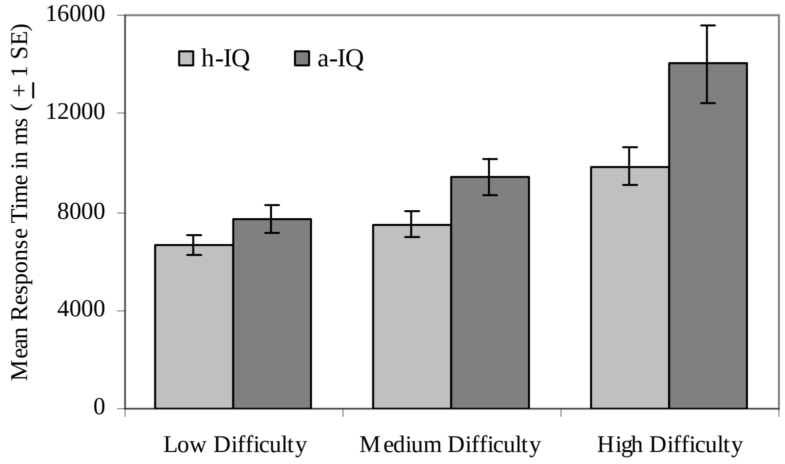
Figure 2. Geometric analogies. Effect of task difficulty (low: mirroring on the vertical axis, medium: mirroring on the horizontal axis, high: mirroring on the diagonal axis) on mean response times (RTs, in milliseconds, ms) and error rates (in %). SE: standard error.

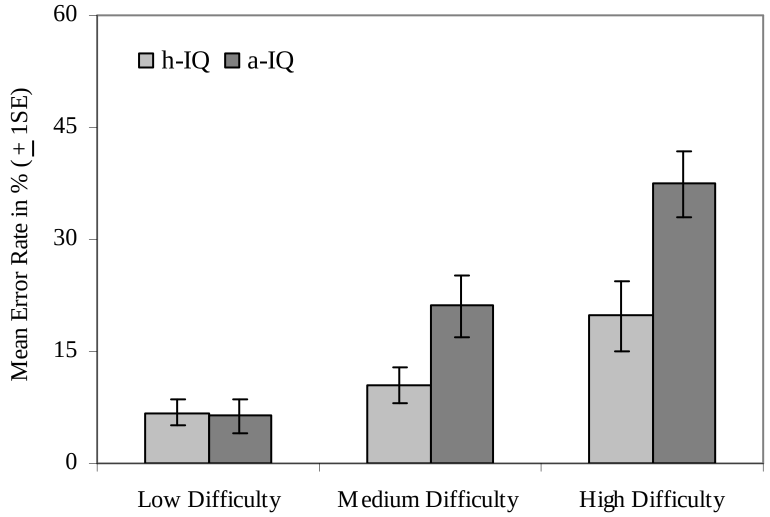
Figure 3. Choice reaction time. Effect of fluid intelligence (measured by RAPM) on mean response times (vertical lines, in sec) and mean pupillary responses (pupil dilation, in mm). h-IQ: individuals with high fluid intelligence; a-IQ: individuals with average fluid intelligence.

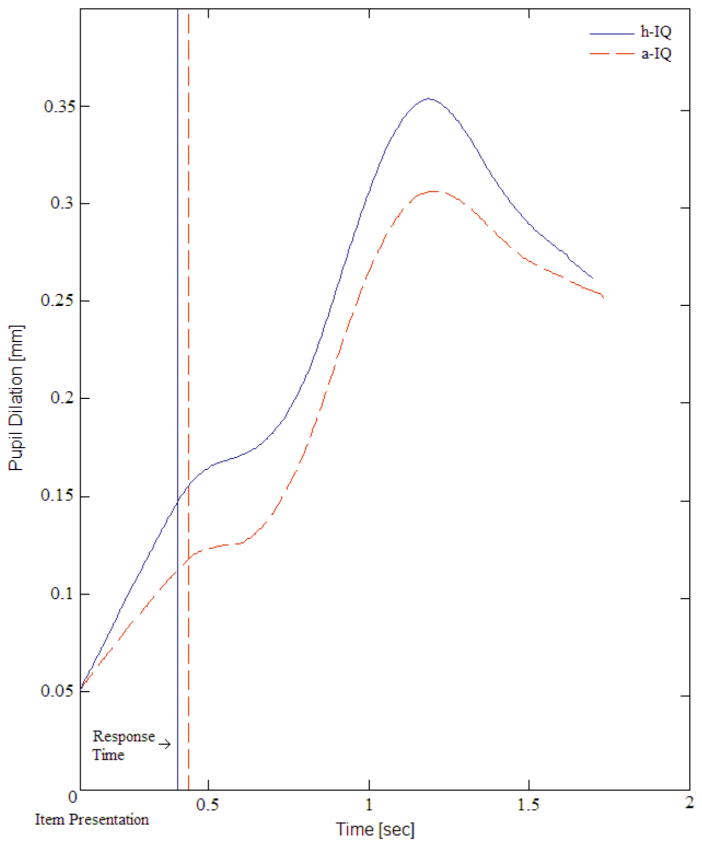
Figure 4. Geometric analogies. Effect of fluid intelligence (measured by RAPM) on mean response times (vertical lines, in sec) and mean pupillary responses (pupil dilation, in mm) depending on task difficulty (low: mirroring on the vertical axis, medium: mirroring on the horizontal axis, high: mirroring on the diagonal axis). h-IQ: individuals with high fluid intelligence; a-IQ: individuals with average fluid intelligence.

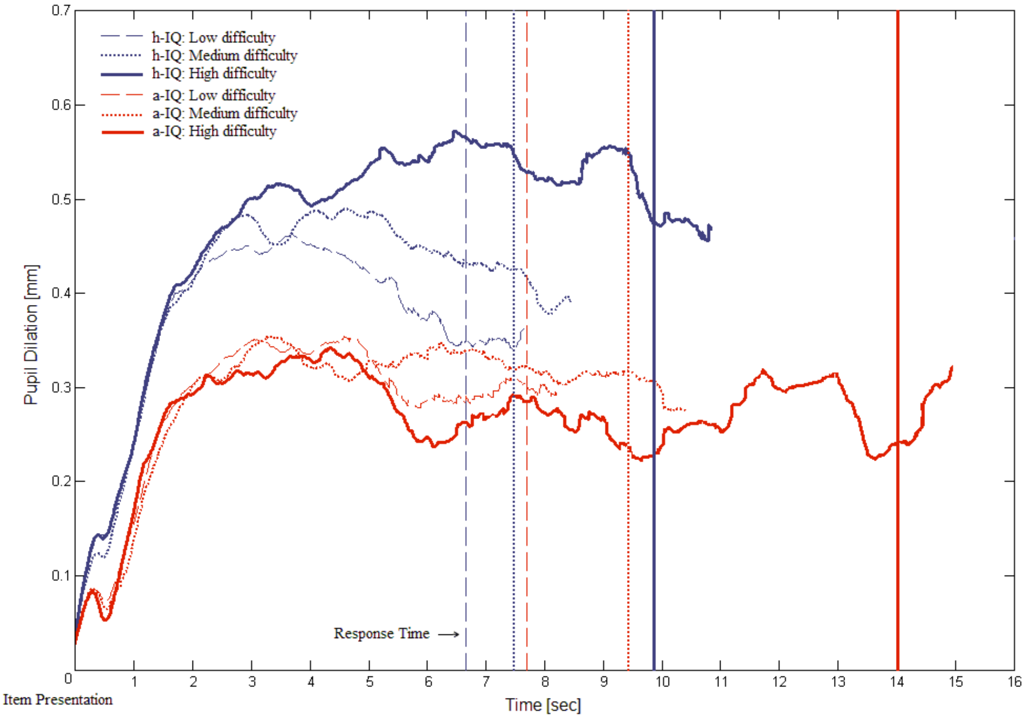





Mean Response Time in ms (+ 1 SE)
16000
12000
8000
4000
0
h-IQ
a-IQ
Low Difficulty
Medium Difficulty
High Difficulty


Mean Error Rate in % (+ 1SE)
60
45
30
15
0
h-IQ
a-IQ
Low Difficulty
Medium Difficulty
High Difficulty


h-IQ
a-IQ
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Pupil Dilation [mm]
Response Time
0.5
1
1.5
2
Item Presentation
Time [sec]


h-IQ: Low difficulty
h-IQ: Medium difficulty
h-IQ: High difficulty
a-IQ: Low difficulty
a-IQ: Medium difficulty
a-IQ: High difficulty
Pupil Dilation [mm]
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Response Time
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Item Presentation
Time [sec]