The Visual Implications of Inspection Time

Tyler W. Garaas and Marc Pomplun

Department of Computer Science, University of Massachusetts Boston
100 Morrissey Blvd., Boston, MA, USA 02125-3393

Abstract. The quest to define human intelligence has led researchers down a large range of paths. One such path has been the search for a single, basic psychometric measure that can be used to account for a large portion of the variance in human mental ability. Inspection Time (IT) has emerged at the forefront of these efforts and can be shown to account for approximately 25% of the variance in psychometric tests of intelligence (e.g., IQ). In this study, we attempt to gain an insight into the nature of IT as a psychometric measure by first contrasting individuals that are adept at performing the IT task (those with low ITs) with individuals that are not (those with high ITs) using oculomotor and task-performance measures recorded during two visual tasks. The results of the first experiment show that the current prevailing theory regarding IT, the integration theory, is incapable of accounting for the results found during the visual tasks. This leads us to introduce a novel theory of IT, the watered-tree theory, which places IT as a measure of information propagation. We then perform a second experiment to test the opposing predictions of the integration theory and the watered-tree theory and find that the results are in line with the predictions of the watered-tree theory. A discussion is presented on the implications of the proposed theory and the need for its future validation.

Keywords: Cognitive Informatics, Inspection time, Intelligence, Visual Search, Comparative Visual Search, IQ

1 Introduction

What defines human intelligence? Is it a simple measure of processing speed, or is it based on a person's ability to extrapolate knowledge? The search for the psychological basis of human intelligence has led researchers to attempt to answer this question using a large variety of methods. One such method dates back to the nineteenth century when Francis Galton tried to relate differences in psychometric measures of intelligence (e.g., IQ) to measures of basic human processes. Galton, in his book *Inquiries into Human*

Faculty [1], subjected a large number of human subjects to simple tests of sensory and motor quickness; the tests included measuring how quickly a participant could respond to the playing of a simple tone, tap on the skin, or flash of a light. Although these methods largely failed in their predictive ability of participants' scholastic achievement, their motivation, to relate a measure of processing speed to intelligence, has survived the century that followed. Galton's line of research was resurrected in the mid-twentieth century when various researchers attempted to relate IQ to simple reaction time tasks [2,3]. However, these measures also failed to gain widespread support due to their accounting for only a small percentage of the variance in mental ability and their theoretical intractability in explaining how they account for this variance [4,5].

One measure that has succeeded so far where reaction time tasks had failed is that of Inspection Time (IT). Although IT is now primarily used in studies that focus on investigating its link to intelligence, the original concept of IT was formulated a little differently. IT was first introduced in the seminal paper by Vickers, Nettelbeck, and Wilson [6] and was conceptualized as estimating an individual's rate of sampling from proximal situations [7]. It was not until later investigations [e.g., 8,9] that the relationship between IT and IQ was thoroughly explored. An important aspect of the theoretical rationale for the relationship between IT and intelligence is that the discrimination made in the IT task was designed to be fundamental enough as to be "relatively immune from influence by higher cognitive activities or by motivation and social factors" [10, p. 609]. Furthermore, since participants are allowed as long as necessary to respond to a single trial during the task used to measure IT, differences in physical ability cannot confound the measure as with the reaction time tasks. In this way, IT has a theoretical "leg up" in explaining how it accounts for the variance in mental ability.

Many studies have investigated the correlation between inspection time and intelligence using a wide range of psychometric tests of intelligence [e.g., 8,9,11-13], which has, in turn, led to no fewer than three meta-analysis papers [14-16]. These meta-analysis papers have each placed the correlation between IT and IQ at -0.50, or, put another way, IT accounts for approximately 25% of the variance in IQ across individuals; a substantial portion considering the elusive nature of intelligence. The success of IT to both account for a large part of the variance in mental ability, along with its consistent theoretical rationale for doing so, led some researchers to initially declare that individual differences in IT cause individual differences in IQ [e.g., 10,17]. However, more recently, Deary et al. [18,19] have expounded upon the dangers of following such a simple assumption.

The standard IT task begins by cueing participants' attention to the location of the impending stimulus using a simple cue figure (see Figure 1a). The cue is immediately followed by the presentation of the stimulus figure, which consists of two vertical lines adjoined at their tops by a horizontal line (see Figure 1b); the stimulus is commonly referred to as the pi-figure because of its resemblance to the Greek letter Π. The stimulus figure is randomly presented to the user in one of two possible forms: one where the left leg is longer and one where the right leg is longer. Immediately following the presentation of the stimulus, a backward pattern mask is presented to disrupt any iconic processing of the stimulus (see Figure 1c). After the presentation of the cue-stimulus-mask triplet, it is the participants' task to manually report which of the pi-figure legs they believed to be longer (or shorter); participants are under no time pressure to respond. This procedure is repeated using a range of stimulus durations, referred to as stimulus onset asynchronies (SOAs; period of time between the onset of the stimulus and the onset of the mask), until an SOA can be identified with some confidence for which the participant responds with some preset level of accuracy, for instance, 75% or 90%. The SOA identified in this manner is operationally defined as that participant's IT. IT is sometimes referred to as the critical stimulus onset asynchrony (CSOA) for which participants can respond at a certain accuracy level [e.g., 20].

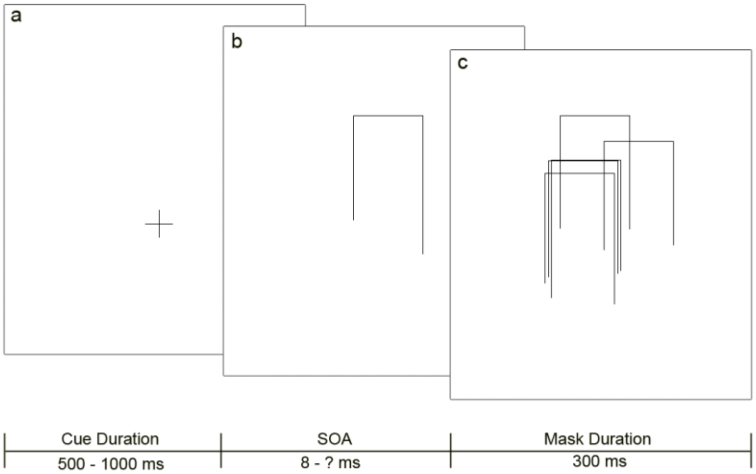

Fig. 1 Figures used for the IT task: (a) the cue figure presented just prior to stimulus onset to focus attention, (b) the IT stimulus figure, commonly referred to as the 'pi-figure', and (c) the backward-mask figure presented after stimulus presentation to prevent any iconic-image processing.

Despite the significant success that has been achieved using IT, a handful of unresolved issues still remain. For instance, the mask originally used by Vickers, Nettelbeck, and Wilson [6] has been shown to exhibit a mask-breaking effect that can be used to undermine the measurement of IT [e.g., 21,22]. This mask-breaking effect was often present in the form of apparent motion, which could be used by participants to artificially lower their IT. Furthermore, at present, there is some dissent over what IT is actually measuring. As mentioned before, participants are not under any time pressure to make a decision, and their ability to answer accurately rests solely on whether they could discriminate the longer leg during the allotted SOA. For this reason, IT has often been described as a measure of participants' speed of information intake or speed of sensory processing [6]. Some researchers have disputed this and instead claim that IT is a measure of participants' temporal resolution of visual perception [20,23,24]. Still others have regarded IT as a measure of general processing speed [25]. Perhaps the most pressing unresolved matter relating to IT is the causal direction between IT and IQ. Various researchers' initial claims, that differences in IT cause differences in mental ability, have since been discarded and if IT is to remain as the foremost psychometric test used to account for variance in mental ability, this issue, above all, must be resolved.

It has been known for quite some time that oculomotor measures (measures related to eye movements) can be used to capture an insight into an observer's underlying cognitive processes. These measures focus around two of the general oculomotor states that can be identified during observers' eye movements in static scenes, fixations and saccades. Fixations refer to the period of time when the observers' gaze is primarily stable in order to gather visual input. On the other hand, saccades refer to the period of time when the observers' gaze is being quickly shifted to a new visual area. These two periods can be broken down into two primary measures that have been shown to provide insight into an observer's current cognitive processes, fixation duration and saccade length [30]. We have identified three possible hypotheses regarding how inspection time may affect these oculomotor measures: (1) Participants with a low IT (higher processing speed or intake speed) may be able to take in the same amount of visual information in a shorter period of time, which would lead to significantly shorter fixation durations. (2) Participants with a low IT may be able to take in more information from a wider visual area during a fixation, which would lead to significantly longer saccades. (3) Participants with a low IT may take in the same amount of information in the same period of time, which would lead to no difference in fixation duration or saccade length; however, this information may be subjected to a larger amount of cognitive processing, which would allow low-IT participants to direct their saccades more efficiently. If the latter is the case, we would expect participants with a low-IT to make significantly fewer saccades and to complete the tasks significantly more quickly.

Another measure derived from video-based eye tracking - though not technically oculomotor - that has proven to be useful in revealing participants' cognitive processes is pupil size. One common example is the use of pupil size as an indicator of cognitive processing load. Aside from pupillary changes caused by environmental factors, consistent changes in pupil size have also been shown to be caused by task-related factors such as processing difficulty and working memory load [26-29]. For instance, Kahneman and Beatty [26] presented participants with a string of three to seven digits at a rate of one per second. After a two-second pause, subjects were asked to repeat the string that was presented to them. It was found that the pupil diameter increased with the presentation of each digit, reaching a maximum in the pause prior to repeating the string. During the report, the pupillary diameter decreased with each digit spoken all the way to the baseline following the report of the final digit. In the present study, we estimated participants' cognitive load by measuring relative pupil variance, which is computed as the difference between a participant's minimum and maximum pupil size divided by their minimum pupil size during an experimental trial [26-28].

In an attempt to determine which of the three hypotheses is correct, we tested participants in an initial experiment using three tasks. In the first task, participants' ITs were recorded using a slight variation of the standard IT task. Following this, participants were tested using two visual tasks. The visual tasks were designed to test participants' visual-attentional control and visual working memory. Specifically, participants performed (1) a visual search task in which they searched a display for a target object amid distracters that varied in color or orientation from the target and (2) a comparative visual search task in which participants attempted to locate a difference between two spatially-separated sets of objects. We found that the results of this initial experiment were inconsistent with the currently prevailing theory of IT, referred to as integration theory, which led us to form a new theory regarding the nature of IT.

The proposed theory, the watered-tree theory, places IT as a measure of information propagation. After briefly discussing the watered-tree theory, we devise and perform a second experiment, about which the integration theory and the watered-tree theory make opposing predictions. The results of this experiment follow the predictions made by the watered-tree theory. A discussion about the implications of the watered-tree theory concludes this chapter.

2 Experiment 1A: Inspection Time Task

To record participants' ITs, we devised an IT task modeled after the standard, backward-masked IT task with two slight changes: (1) The mask used originally by Vickers, Nettelbeck, and Wilson [6] has been shown to exhibit a mask-breaking effect, and to reduce this effect, we designed a mask similar to that used by Knibb [21]; simply put, the mask attempts to "overload" participants' visual field, thus preventing any processing of an iconic image. (2) The standard IT task requires participants to perform the same number of trials across a wide range of stimulus onset asynchronies (SOAs) regardless of their relevance to a particular participant. In an attempt to reduce the number of irrelevant trials (and therefore the overall length of the task), we have created an algorithm that actively pursues only the SOAs that are relevant to computing a participant's IT.

2.1 Method

Participants. The IT task was performed with the assistance of 35 participants that were paid a $10 honorarium for their participation. Of the 35 participants, 22 were male and 13 were female; 16 were undergraduate students, 17 were graduate students, and 2 were faculty at the University of Massachusetts Boston. The median age was 28 and ranged from ages 18 to 41. All of the participants had intact vision and some used corrective lenses.

Apparatus. Stimuli were presented on a 21-inch Dell P1130 monitor using the resolution 1024×768 and a refresh rate of 120 Hz. Participants sat approximately 60 cm from the screen, resulting in a horizontal and vertical viewing angle of 31.5° and 24.6°, respectively. Participants' responses were recorded using a standard PC mouse.

Materials. The IT task stimulus, referred to as the 'pi-figure', consisted of two vertical, parallel lines connected to a horizontal line at their tops (Figure 1b). The pi-figure comes in two forms, one with the left vertical line slightly longer and one with the right vertical line slightly longer; Figure 1b represents the latter. The line lengths for the target-stimulus were 3.4°, 5.1°, and 6.8° for the horizontal, short-vertical, and long-vertical lines respectively. To focus participants' attention, a simple cue in the form of a cross was presented immediately prior to stimulus onset (Figure 1a). Immediately following presentation of the stimulus, a backward-mask was presented to disrupt any processing of an iconic image. The mask was composed of five pi-figures randomly placed in the immediate area of the previous stimulus (Figure 1c).

Procedure. Prior to the start of the experiments, each participant was given instructions about their task. To accustom participants with the task, six initial practice trials were performed starting with extremely large SOAs and leading to moderately low SOAs. During the first practice trial, the IT stimulus was presented to the participant for 750 ms. The following five practice trials presented the target-stimulus for 750 ms, 750 ms, 525 ms, 300 ms, and 150 ms, respectively; the first experimental trial was then presented for 75 ms. Prior to every stimulus presentation, the cue figure was presented for a random period of time between 500 ms and 1000 ms. Immediately following the presentation of the stimulus, the backward-mask was presented for 300 ms.

Participants were instructed to focus on the accuracy of their response and to take as long as they needed to make their response. As they were instructed, the participants pressed the left or right mouse button to indicate they believed the left or right vertical line was longer, respectively. Following a participant's response, the next cue-stimulus-mask triplet was presented. The stimulus presentation period varied with the accuracy of a participant's responses; adjustments were made after every other triplet (trial). SOA was increased by 8.3 ms if the participant responded incorrectly to one or both trials and was decreased by 8.3 ms if the participant responded correctly to both. This process continued until two SOA periods could be identified; one in which the participant responded correctly $\geq 75\%$ of the trials, and one in which the participant responded correctly $\leq 75\%$ of the trials with at least 36 trials for each. The participant's IT was then operationally defined by using linear interpolation to estimate the time at which the participant responded correctly for exactly 75% of the trials.

2.2 Results

Inspection times were obtained for all but one participant, whose response accuracy fluctuated too greatly due to not fully understanding the task; this participant was excluded from the visual search task and the comparative visual search task. Recorded ITs varied from 33.3 ms to 158.3 ms with a mean of 80.1 ± 23.4 ms.

To analyze differences between participants with low and high ITs, two groups of 15 were formed (the middle four were excluded from analysis to allow for a small gap between IT groups). The low-IT group had ITs that ranged from 33.3 ms to 73.9 ms with a mean of 62.0 ± 9.0 ms. The high-IT group had ITs that ranged from 81.9 ms to 158.3 ms with a mean of 98.0 ± 21.4 ms.

3 Experiment 1B: Visual Search Task

Visual searches are a large part of everyday life for most of us; in fact, we perform visual searches so often that most of the time we do not even realize that we are doing so. Visual search tasks have consequently become a prominent paradigm used to gain insight into our visual attention system [31,32]; see Wolfe [32] for a review of visual search. In the visual search task presented, participants searched a display for a black, horizontal bar among a set of white, vertical distracters and white, horizontal distracters (color feature-search), or white, vertical distracters and black, vertical distracters (orientation feature-search), or white, horizontal distracters and black, vertical distracters (conjunctive-search). Loosely speaking, in the feature-search conditions, the target object seems to “pop out” from the distracter objects. The target object in the conjunctive-search condition, on the other hand, is not as readily locatable as the one in the feature-search conditions. We consequently predict that participants will, on average, be significantly better at locating the target object in the feature-search conditions after a single fixation than in the conjunctive-search condition. We will verify this by measuring the distance from the gaze-position to the target after a single saccade has been made.

3.1 Method

Participants & Apparatus. The 34 participants that finished the IT task participated in the visual search task. Stimuli were presented on the same monitor that was used for the IT task with the same resolution and refresh rate. Participants were seated the same distance from the monitor as in the IT task yielding the same viewing angles. Eye movements were recorded using the SR Research Eye-Link II eye-tracker system. The average error of visual angle in this system is 0.5°, and its sampling frequency is 500 Hz. During the visual search task, responses were recorded using a handset (often referred to as a game-pad).

Materials. The stimulus displays used in the visual search task consisted of oriented bars measuring 2.1° in length and 0.7° in width, with features varying in two dimensions, color and orientation; leaving four possible object types, horizontal, black bars; horizontal, white bars; vertical, black bars; and vertical, white bars. Each stimulus display contained 40 of these objects, of which, one random object was chosen to be the target object and was swapped for a horizontal, black object. Stimulus displays were divided into three categories: color-search (Figure2a), orientation-search (Figure 2b), and conjunctive-search (Figure 2c). Objects in the color-search displays consisted of an equal mix of objects that differed from the target object in their color or in both their color and orientation. Orientation-search displays were identical to color-search displays except that the distracters differed from the target object in their orientation or in both their orientation and color. Conjunctive-search displays were composed of an equal mix of objects that always differed from the target object in a single dimension. Objects were randomly placed in a screen-centered display area which had a length and width of 20.7°; the minimum distance between object centers was 2.6°. All stimulus displays were generated prior to starting the experiment so that each participant was subject to the same set of displays.

Procedure. Participants were informed of the categories of the stimulus displays and of the identity of the target object prior to starting the experiment. They were instructed to find the target object in each trial as quickly and accurately as possible, and to then press a button on the game-pad while fixating on the target object. Prior to starting the experimental trials, participants were fitted with the eye-tracker headset, which was followed by the calibration of the eye-tracker system. Participants were then tested under three conditions: a color-search condition, an orientation-search condition, and a conjunctive-search condition; each condition was composed solely of displays from their respective display category. Trial conditions were presented in blocks of ten trials plus two training trials that were presented the first time a trial condition was presented. Participants were shown eight blocks of trials that were broken down into two color-search blocks, two orientation-search blocks, and four conjunctive-search blocks. The ordering of blocks and stimulus displays within each block were completely randomized, except for the two training displays, which were always presented at the start of the blocks they were present in. Prior to each trial, a simple drift correction was performed in which participants were instructed to fixate on a dot shown in the center of the screen and press a button to start the trial. Trials ended only after the button press, indicating that participants were fixating on the target object.

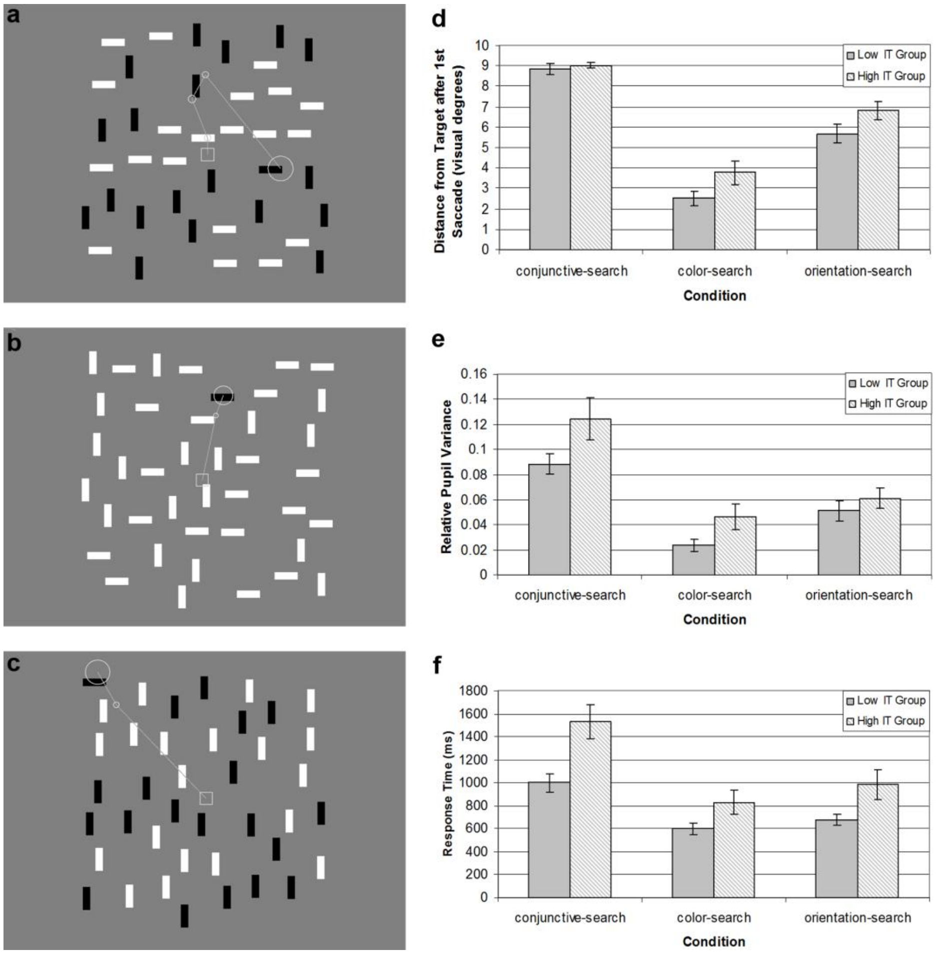

Fig. 2 (a-c) Visual search displays with a participant's sample eye movements superimposed on each: (a) conjunctive-search stimulus display, (b) color-search stimulus display, and (c) orientation-search stimulus display. (d-f) Visual search results: (d) distance to the target after the first saccade, (e) relative pupil variance, and (f) response time for the low- and high-IT groups.

3.2 Results

The visual search task and comparative visual search task produced a large amount of data, of which only the most relevant will be presented. Unless otherwise noted, all variables for the visual search task were analyzed using a 2-way ANOVA with IT group (low-IT vs. high-IT) as a between-subject factor and task condition (color-search vs. orientation-search vs. conjunctive-search) as a within-subject factor. The distance between participants' fixation location at the button press and the target object, referred to as trial accuracy from here on, was 2.6° ± 0.7°. To provide a baseline to compare trial accuracies against, rotated trial accuracies were also computed by rotating the display 180° around its center and computing the distance between participants' non-rotated final-fixation location and the rotated target-object; rotated trial accuracy was 8.1° ± 2.5°. The significant difference between these two accuracy measures, $t(29) = 11.59$, $p < 0.001$, demonstrates that participants were in fact performing their task. Furthermore, while trial accuracies did differ significantly across task condition (color-search: 1.5°; orientation-search: 2.7°; conjunctive-search: 3.5°), $F(2;56) = 38.44$, $p < 0.001$, they did not vary significantly across IT group (low-IT: 2.4°; high-IT: 2.7°), $F(1;28) = 1.12$, $p > 0.1$. The distance between the location of the second fixation (after the first saccade) and the target object (see Figure 2d) was significantly lower for low-IT participants (5.7°) than for high-IT participants (6.5°), $F(1;28) = 4.24$, $p < 0.05$. As predicted, the accuracy after the first saccade also varied significantly between task conditions (color-search: 3.1°; orientation search: 6.2°; conjunctive-search: 8.9°), $F(2;56) = 148.27$, $p < 0.001$. Relative pupil variance (see Figure 2e) varied significantly across task condition (color-search: 0.04; orientation-search: 0.06; conjunctive-search: 0.11), $F(2;56) = 92.16$, $p < 0.001$, and tended to be smaller for participants in the low-IT group (0.05) than in the high-IT group (0.08), $F(1;28) = 3.23$, $p < 0.1$, which implies that low-IT participants required a smaller cognitive load to complete the visual search task trials than high-IT participants.

Response times (see Figure 2f) were found to be significantly shorter for low-IT participants (755.3 ms) than for high-IT participants (1114.1 ms), $F(1;28) = 7.90$, $p < 0.01$. Response times also differed significantly across task condition (color-search: 709.6 ms; orientation-search: 828.5 ms; conjunctive-search: 1266.0 ms), $F(2;56) = 65.15$, $p < 0.001$. Interestingly, the interaction between IT group and task condition was significant as well, $F(2;56) = 9.04$, $p < 0.001$, which is likely due to the greater separation between response times during the conjunctive-search condition than either of the feature-search conditions. Neither fixation duration nor saccade length can account for the difference in response times between IT groups. Fixation durations were not significantly lower for low-IT participants (241.2 ms) than they were for high-IT participants (238.2 ms), F(1;28)

< 1.0. Saccade lengths were also found to not vary significantly across IT group (low-IT: 5.2°; high-IT: 4.8°), $F(1;28) < 1.0$. The number of fixations was significantly lower for low-IT participants (2.3 fixations) than for high-IT participants (3.3 fixations), $F(1;28) = 11.51$, $p < 0.005$, which is a direct consequence of the statistically identical fixation durations between the two IT groups. The number of fixations also varied significantly across the task condition (color-search: 1.9 fixations; orientation-search: 2.4 fixations; conjunctive-search: 4.2 fixations), $F(2;56) = 104.00$, $p < 0.001$. A significant interaction between IT group and task condition that corresponds to the one for response time was also found for the number of fixations, $F(2;56) = 9.04$, $p < 0.001$.

3.3 Discussion

In the *Introduction* we identified three potential hypotheses that could explain performance differences in the visual tasks from Experiment 1. The results of the first task strongly suggest that the third hypothesis is correct, as no significant difference was found between low-IT and high-IT participants' fixation durations or saccade lengths, whereas we did find significant differences in participants' response times and the number of fixations made per trial. Another piece of evidence that strongly supports the adoption of the third hypothesis is that low-IT participants' first saccade was found to be significantly closer to the target than high-IT participants' first saccade, which suggests that low-IT participants were able to recognize the target object significantly more often during their first fixation.

The adoption of one of the hypotheses regarding the relationship between IT and the visual tasks also has implications for the nature of what IT measures. That is, if either of the first two hypotheses is correct, IT most likely relates to performance differences in the visual tasks through sensory abilities, such as the speed of sensory intake; however, if the third hypothesis is correct, it would indicate that IT relates to performance differences through a factor that is, at least partially, cognitive in nature. Consequently, the results of the first visual task suggest that IT measures participants' cognitive processing speed. This conclusion is supported by the fact that low-IT participants demonstrated significantly less pupil variance within the trials, which indicates that – besides being able to complete trials more quickly - they were also less cognitively taxed when doing so. Furthermore, there was an interaction between IT-group and task condition, which indicates that low-IT participants demonstrate significantly greater abilities in one or two of the task conditions. If the difference between IT-groups lied solely in the speed of information intake, we would expect a constant factor of response time differences;

instead we find a factor that seems to increase with greater processing difficulty in the experimental conditions.

4 Experiment 1C: Comparative Visual Search Task

As with visual search tasks, comparative visual search tasks rely on stringent visual-attentional control [33]. However, unlike visual search tasks, comparative visual search tasks require the effective use of visual working memory for task completion. In this way, comparative visual search tasks have been shown to yield valuable insight into our use of visual working memory and visual-attentional control. For instance, Inamdar and Pomplun [34] demonstrated that participants would increase the use of their visual working memory to compensate for more costly eye movements up to their visual working memory capacity.

In the comparative visual search task presented here, participants were shown two nearly identical sets of objects (one object was dissimilar between the two sets) positioned on the left and right sides of the monitor. The two sets of objects were composed of the same oriented bars that were used in the visual search task. It was the participants' task to locate the single difference between the two sets of objects. Since these two sets were set up so that participants could not simultaneously attend to both of them at the same time, participants had to first "load" their visual working memory with objects from one side of the display and then "retrieve" what they had loaded to make a comparison against the objects on the opposite side of the display. Given the nature of IT, it is possible, or even likely, that participants with a low IT can load their working memory more quickly than participants with a high IT, which would allow them to load more objects into memory during the same amount of time. Consequently, we predict that participants with a low IT will be significantly faster (i.e., exhibit shorter response times) at locating the difference between the two hemi-fields while exhibiting evidence of larger visual working memory loads than participants with a high IT.

4.1 Method

Participants & Apparatus. The 34 participants that finished the IT task participated in the comparative visual search task. The apparatus from the visual search task served as the apparatus in the comparative visual search task as well.

Materials. The two sets of objects were separated by 5.5° and a single black line down the center of the display. Each side contained 20 objects that were composed of an equal number of the four different object types (i.e., five of each type). Objects were placed on each side such that the minimum distance between the centers of any two objects was at least 2.6°. The two sides of the display were identical except for a single discrepancy. A discrepancy occurred when an object from either side had its color or orientation swapped. The stimulus display categories were therefore defined by the dimension of the target object that was swapped, resulting in two categories of displays: color-discrepant displays (Figure 3a); and orientation-discrepant displays (Figure 3b). All stimulus displays were generated prior to starting the experiment so that every participant was subject to the same set of displays.

Procedure. Participants were given initial instructions about the nature of the experiment and their role in the task. They were instructed to locate the discrepancy between the two sides as quickly and accurately as possible, and to press a button on the game-pad while fixating on either object of the discrepancy. Participants were also instructed to serially search the display for the discrepancy starting at the top, and that if they had not located the discrepancy by the time they hit the bottom, to then search as they saw fit. Participants were tested under three trial conditions which were presented in blocks of 20 trials plus two training trials shown the first time a trial condition was presented. Prior to starting a block of trials, participants were shown a string of text that was used to indicate that the dimension of the discrepancy was either color, orientation, or unspecified for the following block, leading to the following three trial conditions: the color-discrepant condition, the orientation-discrepant condition, and the uninformed condition, respectively. The two informed conditions were composed solely of displays from their respective stimulus display categories. The uninformed condition was composed of an equal number of displays taken from each of the two stimulus display categories. Participants were shown four blocks of stimulus displays, of which, two blocks were uninformed, one was color-discrepant, and one was orientation-discrepant. The ordering of blocks and stimulus displays within each block was completely randomized except for the two training trials which were presented at the start of the blocks they were present in. Prior to the start of each trial, a drift correction similar to the one used in the visual search task was performed, except that the fixation point was presented at the top-center of the screen. Trials ended only after the button press indicating that participants were fixating on the target object.

4.2 Results

Unless otherwise noted, variables for the comparative visual search task were analyzed using a 3-way ANOVA with IT group (low-IT vs. high-IT) as a between-subject factor and discrepancy information (uninformed vs. informed) along with discrepancy dimension (color vs. orientation) as within-subject factors. Trial accuracy for all participants was 3.2° ± 1.5°, and the rotated trial accuracy was 17.6° ± 1.2°. The significant difference between the trial accuracies, $t(29) = 32.91$, $p < 0.001$, demonstrates that participants were performing the correct task. Trial accuracy did not vary between IT group (low-IT: 3.2°; high-IT: 3.3°), $F(1;28) < 1.0$, and neither did relative pupil variance (low-IT: 0.29; high-IT: 0.32), $F(1;28) < 1.0$. We assume that the jump in relative pupil variance between the visual search task and the comparative visual search task, as well as the invariance between IT groups in the comparative visual search, is due to the required extensive use of working memory to complete the comparative visual search task, since pupil size has been shown to be influenced by working memory load [26]. Consequently, the lack of a pupil size difference between IT groups in comparative visual search would be due to a ceiling effect with regard to working memory load.

Comparative visual search response times (see Figure 3c) were significantly shorter for low-IT participants (6355 ms) than high-IT participants (9272 ms), $F(1;28) = 15.23$, $p < 0.005$. As with the visual search task, fixation duration and saccade length could not be held accountable for the difference in response times between the two IT groups. Fixation durations did not vary significantly across IT group (low-IT: 216.0 ms; high-IT: 214.6 ms), $F(1;28) < 1.0$, and neither did saccade lengths (low-IT: 8.8°; high-IT: 8.2°), $F(1;28) = 1.05$, $p > 0.1$. Since it is possible that the saccade length measure could have been confounded by inter-hemifield saccades, the intra-hemifield saccade length was also computed, which also did not vary significantly across IT group (low-IT: 4.3°; high-IT: 4.4°), $F(1;28) < 1.0$.

As with the visual search task, low-IT participants also made significantly fewer fixations per trial in the comparative visual search task (low-IT: 24.1 fixations; high-IT: 35.0 fixations), $F(1;28) = 18.62$, $p < 0.001$. To get a better understanding of how this difference arose, the number of fixations per trial was divided into those made prior to reaching the target object and those made after. Since participants were instructed to compare objects in series from top to bottom, the number of prior-fixations was calculated by counting all the fixations that occurred prior to the participant reaching the target object's vertical area, which is defined as the center of the target object ± 50 pixel-rows; the after-fixations were then defined as all remaining fixations.

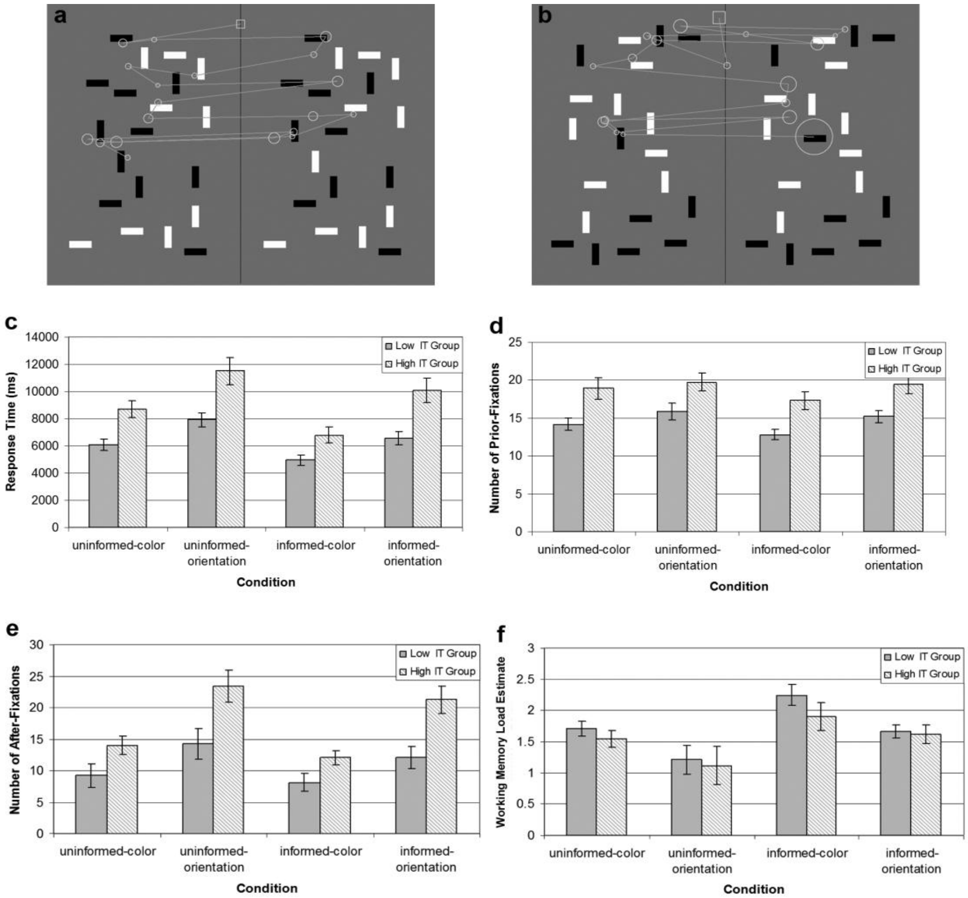

Fig. 3 (a-b) Comparative visual search displays with a participant's sample eye movements superimposed on each: (a) example color-discrepant stimulus display and (b) example orientation-discrepant stimulus display. (c-f) Comparative visual search task results: (c) response times, (d) number of prior-fixations, (e) number of after-fixations, and (f) estimate of visual working memory load for the low- and high-IT groups.

The number of prior-fixations (see Figure 3d) made per trial was significantly lower for low-IT participants (14.0 fixations) than for high-IT participants (18.3 fixations), $F(1;28) = 10.86$, $p < 0.005$. The fact that low-IT participants made significantly fewer prior-fixations demonstrates that participants in the low-IT group are more efficient at loading objects into memory, retrieving objects from memory, or both. As with the prior-fixations, the number of after-fixations (see Figure 3e) was also significantly lower for low-IT participants (10.2 fixations) than for high-IT participants (16.6 fixations), F(1;28) = 11.42, $p < 0.005$, which clearly reveals that participants in the low-IT group either miss the target less often, recognize the target more quickly, or both.

Interestingly, variables that are related to visual working memory varied significantly between the IT groups, but the actual estimation of visual working memory load did not. The estimate of participants' visual working memory load was computed by dividing the number of objects above the target by the number of inter-hemifield saccades made prior to a participant reaching the target object. While the visual working memory load estimate (see Figure 3f) did not vary significantly across IT group (low-IT: 1.8 objects; high-IT: 1.7 objects), F(1;28) < 1.0, it did vary significantly across discrepancy dimension (color: 1.9 objects; orientation: 1.7 objects), F(1;28) = 7.20, $p < 0.05$, and discrepancy information (uninformed: 1.7 objects; informed: 1.9 objects), F(1;28) = 6.34, $p < 0.05$. A significant interaction between discrepancy dimension and discrepancy information was also found for the estimate of visual working memory load, F(1;28) = 15.50, $p < 0.001$. The significant interaction obtained from the estimate of visual working memory load is in line with our expectations and is likely due to participants' ability to perceptually group objects better when they knew the discrepant dimension was color.

4.3 Discussion

As with the visual search task, low-IT participants were able to complete the comparative visual search task more quickly and with fewer fixations than high-IT participants across all conditions. Furthermore, and also in line with the visual search task, the greater performance of the low-IT group could not be accounted for by either fixation duration or saccade length. This result provides further support for the third hypothesis stating that low-IT participants subject the incoming visual information to more comprehensive cognitive processing which allows them to more effectively direct their eye movements. However, the lack of significant difference between the fixation durations and saccade lengths of the low- and high-IT groups could be accounted for by more effective visual working memory use. The support for this possibility is weak, as the working memory load estimate did not turn out to be significantly different between IT-groups. The conclusion that low-IT participants subject items to more comprehensive processing is also supported by the fact that low-IT participants make fewer after-fixations than high-IT participants, which indicates that they either locate the target object faster, miss the target object less often, or both.

The results from the comparative visual search task also provide further evidence against a purely sensory interpretation of performance differences between the IT-groups. In particular, as with the visual search task, the difference factor between response times follows the perceived difficulty of the task conditions.

5 General Discussion of Experiment 1

In Experiment 1 we attempted to gain some insight into the nature of IT by first measuring the ITs of a group of participants and then dividing the participants into two groups: those with a low IT (i.e., those hypothesized to have higher processing speed or intake speed) and those with a high IT. We then tested participants using two visual tasks while measuring various oculomotor and performance variables. In both of these tasks, we found that participants in the low-IT group completed trials significantly more quickly than those in the high-IT group during all task conditions. This result tells us –as expected– that low-IT participants possess some form of greater visual ability, but what abilities in particular are greater? Prior to the experiment, we identified three potential hypotheses regarding the way that the greater visual ability could be related to task performance. The results from both visual tasks overwhelmingly support the third identified hypothesis, which states that low-IT participants' greater visual abilities are the result of more comprehensive processing of visual items which allows them to more efficiently direct their eye movements during the task. Support for this hypothesis came in the form of low-IT participants making significantly fewer fixations per trial during both tasks, a tendency toward smaller relative pupil-size variance during the visual search task, a significantly more accurate first saccade during the visual search task, and significantly fewer after-fixations during the comparative search task. Furthermore, if low-IT participants' greater ability were only due to a low-level perceptual advantage, as suggested by the first two hypotheses, we would expect a consistent factor of performance and fixation count differences during the visual search task. On the contrary, performance and fixation count differences between IT groups during both tasks appear to increase with greater processing difficulty.

The results discussed above strongly imply that the observed greater visual abilities stem, at least partially, from post-sensory processing differences. This conclusion has strong repercussions for the nature of what IT is measuring. In particular, any purely sensory interpretation of IT would need to be rejected if the variance in IT is found to be accounted for by differences in processing speed, which is what we found during Experiment 1. Nevertheless, it is possible –perhaps even likely– that IT is a measure that transcends the pre- and post-sensory division, but our data only demonstrate a clear difference in post-sensory abilities.

The results of the previous tasks also have implications for the current prevailing theory regarding the nature of IT, the integration theory. Specifically, the integration theory states that the presentation of the backward-mask following the IT stimulus prevents participants from correctly discriminating which leg is longer because at sufficiently short SOAs, the stimulus and mask are *integrated* into a single sensory observation. IT under this theory is then defined as a measure of participants' temporal resolution; if the resolution is fine enough, the two images will not be integrated and can be perceived separately. This theory of IT, however, is incapable of accounting for the results obtained from Experiment 1. Instead, we present a different idea of what IT measures.

Based on the results of the two visual tasks, we propose that IT is actually a measure of the speed at which information propagates across neural pathways, which transcends the pre- and post-sensory boundary. Consider the following crude, but effective analogy. In this analogy, imagine that our brain is represented by a tree (not that this is realistic, but for the sake of the analogy) and that higher branches represent higher levels of cognition; information in this analogy is represented as water and comprehension as a budding leaf. To test for comprehension, we can simply draw a line through a branch of the tree (define the task), apply water (present the stimulus), and check for a bud (analyze the response). In the case of the IT task, we would draw a line through a very low branch and apply water for different periods of time to determine the amount of water needed to usually produce a bud. In this case, IT is indicative of the efficiency of the tree to transfer a sufficient amount of water to allow discrimination. If the water is removed before enough of it can reach the line, no bud will be produced and the discrimination will be made at chance levels. Since, in this analogy, higher levels of cognition are represented by higher branches, the water needed to produce buds on these branches must travel a greater distance, which results in larger differences in the tree's ability to transfer the water (comprehension or task performance). In other words, we expect greater performance differences between low- and high-IT participants for cognitively more demanding tasks – a pattern of results that we have consistently found in both visual tasks reported above, which the integration theory cannot explain.

This view of IT, which we shall refer to as the watered-tree theory, does not come without its own implications. In addition to resolving what IT is actually a measure of, the causal direction of IT is also resolved if the watered-tree model is correct; that is, in the watered-tree model, both IT and IQ are the result of a participant's neural efficiency; those participants with greater neural efficiency will not only demonstrate a lower IT but also significantly greater intellectual abilities. However, it should be noted that IT only accounts for approximately 25% of the variance in IQ with other factors also contributing significantly, which almost certainly plays a part in the elusive nature of intelligence.

The watered-tree model makes a number of predictions about the nature of the IT task that can be tested to simultaneously examine its and the integration theory's validity. To design such a task, one could simply draw lines further up the branches or on different branches; that is to say, we could define other tasks similar to the IT task that test participants using higher cognitive (e.g., which of three lines differs from the other two?) and diverse cognitive (e.g., which of two objects is brighter?) tasks, respectively. If the watered-tree model is correct, the IT results for each task should be distinct while still correlating significantly with each other across participants.

One very simple test that could be performed to examine the validity of the theories would be to simply increase the line thickness of the IT stimulus and backward-mask, which, if the integration theory is correct, should not affect the results of the IT task. Conversely, if the watered-tree model is correct, the resulting ITs should be significantly less, because more water is provided to the tree - i.e., more signal is applied to the relevant neural structures. Moreover, the watered-tree theory predicts a strong correlation between thin- and thick-line IT across participants. While future work will be directed towards performing some of the tests described above, we have already performed a simple test using the thick-line IT, the results of which are presented in the following section.

6 Experiment 2: Hypothesis Test

The results of the first experiment led to the identification of a new theory regarding the nature of IT and the rejection of the current prevailing theory of IT, the integration theory. Specifically, the proposed theory posits IT as a measure of information propagation across the various neural pathways. In opposition, the integration theory posits IT as a measure of temporal resolution where the backward-mask in the standard IT task prevents successful discrimination of the stimulus figure at short SOAs because the two become integrated into a single figure. We also identified a simple experiment that could be performed to test the implications of both theories. This test involves a simple modification of the standard IT task in which the lines of each figure in the task are thickened by some moderate amount. In the integration theory, the results of the IT task should be identical to those of the thick-line task, since in either case, an integrated figure would prevent discrimination of the stimulus figure. On the other hand, if the theory we propose is correct, the results of the thick-line IT-task should provide slightly shorter IT times while still preserving the inter-individual differences (i.e., the times from the two tasks should be highly correlated across participants). In an attempt to test the merits of both theories regarding the nature of IT, we tested a series of participants using both the standard IT task and the thick-line IT task.

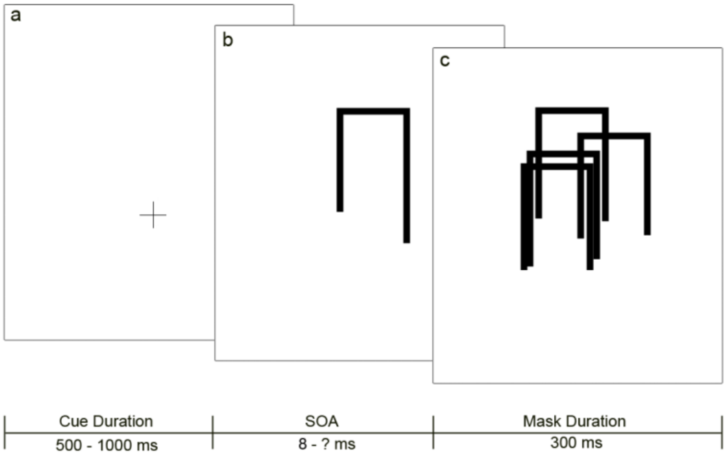

Fig. 4 Figures used for the thick-line IT task: (a) the cue figure, (b) the 'pi-figure', and (c) the backward-mask figure.

6.1 Method

Participants & Apparatus. Seven students from the University of Massachusetts Boston participated in the experiment and were paid a $10 honorarium for their participation. Of the 7 participants, 4 were male and 3 were female; 5 were undergraduate students and 2 were graduate students at the University of Massachusetts Boston. The median age was 24, and ages ranged from 18 to 32 years. All of the participants had intact vision and some used corrective lenses. The apparatus that was used for the IT task in Experiment 1 was used for this experiment as well.

Materials. The materials from the IT task in Experiment 1 were also used for the standard version of the IT task in this experiment. For the thickened version of the IT task, the line thickness was increased from one pixel to ten pixels. The difference between the figures used for both versions of the task can be seen by comparing Figure 1 to Figure 4.

Procedure. The procedure used for the IT task in Experiment 1 was also used for both versions of the IT task in this experiment. The two versions of the IT task were run in a pseudorandom order to reduce the effect of fatigue.

6.2 Results

Results were obtained for all seven subjects in both tasks. The ITs recorded during the standard version of the task ranged from 41.7 to 130.7 ms with a mean of 90.3 ± 30.8 ms (s.d.). The ITs recorded during the thickened version of the IT task ranged from 40.0 to 116.7 ms with a mean of 82.7 ± 27.6 ms. The average difference between the ITs recorded during the two versions of the task was 7.6 ± 6.7 ms. A paired-sample t-test revealed that the ITs recorded during the two versions of the task were significantly different, $t(6) = 3.0$, $p < 0.05$. In support of the introduced hypothesis, the correlation between the ITs recorded during the two versions of the IT task were found to be highly correlated with each other, $r = 0.98$, $p < 0.001$.

6.3 Discussion

The results of Experiment 2 follow the predictions of the watered-tree theory regarding the nature of IT. The use of a slightly thicker line during the IT task was able to significantly reduce the amount of time participants needed to be shown the stimulus figure in order to identify the longer leg of the pi-figure. Furthermore, the ITs recorded from the two versions of the IT task correlate almost perfectly, which is also predicted by the watered-tree hypothesis. On the other hand, the currently prevailing integration theory predicts that there should have been no change in the resulting ITs.

7 Conclusions

On the basis of the two experiments presented here, it seems justified to reject the integration theory regarding the nature of inspection time in lieu of the watered-tree theory. However, the previous experiments have not conclusively verified the watered-tree theory, but instead just provided some evidence in its favor. More experimentation will be needed to further test the watered-tree theory as a plausible explanation of individual differences in IT.

In the *General Discussion of Experiment 1* we hypothesized that the underlying cause of the differences in the speed of information propagation were related to participants' neural efficiency. As such, it should then be asked, in what way are low-IT participants' neural pathways more efficient? Perhaps the easiest conclusion to draw regarding this question is that participants with a low-IT actually have nerve impulses that travel faster than high-IT participants. Indeed, this is a hypothesis that has compelled a handful of researchers to investigate whether nerve conductance velocity (NCV) are correlated with measures of intelligence which sometimes resulted in substantial correlations [e.g., 35,36]. However, further analyses of these studies have resulted in the conclusion that, "the small differences in NCV between bright and dull subjects could not account for more than a small amount of the relatively large difference between these groups on IT and [reaction time] RT tasks." [37, pp. 368].

Alternatively, neural efficiency is sometimes discussed in relation to the individual computational properties of a single neuron [37], but a more interesting interpretation is actually based on the efficiency of the connections between the neurons [38]. In such a case, it is not the efficiency of a single neuron's computation, but the collective computation of an assembly of neurons, which may or may not constitute a functional unit. Evidence for such an interpretation has actually come from a psychophysical study performed with patients that had recently gained the ability to see: "in von Senden's (1932) reports, human beings who have been operated on for congenital cataracts may gradually learn to identify objects. In ordinary circumstances their readiness for identification appears as efficient as that of normal people. On the other hand, when subjects are given only a very short time to scrutinize objects, as when objects are exposed through a camera that opens but for a fraction of a second or through a tachistoscope, identification fails, but does not fail in the normal subject." [39, pp. 97]. Presumably, the difference between the normal subjects and those that were operated on is the lack of an efficient neural organization that can be attributed to missing the critical development period that is known to take place in the first few years of life [40]. Interestingly, the report from von Senden [41] very closely mimics the design for the standard IT task. As such, a possible explanation is that participants with a low-IT actually have a more efficient neural organization that allows them to perceive the pi-figure more quickly than high-IT participants, which in turn enables them to discriminate the longer leg using shorter presentation times before the visual mask has a chance to disrupt perception. This is a very attractive hypothesis due to its ability to unite many efforts aimed at examining the nature of intelligence, but it will take much more future work to verify that this is the correct interpretation.

References

1. Galton, F. (1883). Inquiries into Human Faculty. London, UK: Macmillan.
2. Hick, W. E. (1952). On the rate of gain of information. Quarterly Journal of Experimental Psychology, 79, 349-360.
3. Sternberg, S. (1966). High speed scanning in human memory. Science, 153, 652-654.
4. Brody, N. (1992). Intelligence (2nd Ed.). San Diego, CA: Academic Press.
5. Detterman, D. K. (1987). What does reaction time tell us about intelligence? In P.A. Vernon (Ed.), Speed of information-processing and intelligence (pp. 177-200). Norwood, NJ: Ablex.
6. Vickers, D., Nettelbeck, T. & Wilson, R. J. (1972). Perceptual indices of performance: The measurement of 'inspection time' and 'noise' in the visual system. Perception, 1, 263-295.
7. Nettelbeck, T. (2001). Correlation between inspection time and psychometric abilities: A personal interpretation. Intelligence, 29, 459-474.
8. Nettelbeck, T. & Lally, M. (1976). Inspection time and measured intelligence. British Journal of Psychology, 67, 17-22.
9. Nettelbeck, T., Cheshire, F. & Lally, M. (1979). Intelligence, work performance and inspection time. Ergonomist, 22, 291-297.
10. Vickers, D., & Smith, P. (1986). The rationale for the inspection time index. Personality and Individual Differences, 7, 609-624.
11. Deary, I. J. (1993). Inspection time and WAIS-R IQ subtypes: a confirmatory factor analysis study. Intelligence, 17, 223-236.
12. Stough, C., Brebner, J., Nettelbeck, T., Cooper, C. J., Bates, T. & Mangan, G. L. (1996). The relationship between intelligence, personality and inspection time. British Journal of Psychology, 87, 255-268.
13. Burns, N. R. & Nettelbeck, T. (2003). Inspection time in the structure of cognitive abilities. Intelligence, 31, 237-255.
14. Grudnik, J. L. & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. Intelligence, 29, 523-535.
15. Kranzler, J. H. & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. Intelligence, 13, 329-347.
16. Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), Speed of Information Processing and Intelligence (pp. 295-346). Norwood, NJ: Ablex.
17. Brand, C. R. & Deary, I. J. (1982). Intelligence and 'inspection time'. In H. J. Eysenck (Ed.), A model for intelligence (pp. 133-148). New York, NY: Springer-Verlag.
18. Deary, I. J. & Stough, C. (1996). Intelligence and Inspection Time Achievements, Prospects, and Problems. American Psychologist, 51, 599-608.
19. Deary, I. J., (2001). Human intelligence differences: toward a combined experimental-differential approach. TRENDS in Cognitive Sciences, 5, 164-170.
20. White, M. (1993). The inspection time rationale fails to demonstrate that inspection time is a measure of the speed of post-sensory processing. Personality and Individual Differences, 15, 185-198.

21. Knibb, K. (1992). A dynamic mask for inspection time. Personality and Individual Differences, 13, 237-248.
22. Evans, G., & Nettelbeck, T. (1993). Inspection time: a flash mask to reduce apparent movement effects. Personality and Individual Differences, 15, 91-94.
23. White, M., (1996). Interpreting inspection time as a measure of the speed of sensory processing. Personal and Individual Differences, 20, 351-363.
24. Burns, N. R., Nettelbeck, T., White, M. (1998). Testing the interpretation of inspection time as a measure of speed of sensory processing. Personality and Individual Differences, 24, 25-39.
25. Burns, N. R., Nettelbeck, T., & Cooper, J. (1999). Inspection time correlates with general speed of processing but not with fluid ability. Intelligence, 27, 37-44.
26. Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. Science, 154, 1583-1585.
27. Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. Science, 157, 218-219.
28. Kahneman, D. (1973) Attention and effort. Englewood Cliffs, N.J: Prentice-Hall.
29. Beatty, J. (1982). Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. Psychological Bulletin, 91, 276-292.
30. Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. Psychological Bulletin, 124, 372-422.
31. Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12, 97-136.
32. Wolfe, J. M. (1998). Visual Search. In H. Pashler (Ed.), Attention (pp. 13-71). Hove, England UK.
33. Pomplun, M., Sichelschmidt, L., Wagner, K., Clermont, T., Rickheit, G., & Ritter, H. (2001). Comparative visual search: A difference that makes a difference. Cognitive Science, 25, 3-36.
34. Inamdar, S. & Pomplun, M. (2003). Comparative Search Reveals the Tradeoff between Eye Movements and Working Memory Use in Visual Tasks. In R. Alterman & D. Kirsh (Eds.), Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society, 2003 (pp. 599-604). Boston, MA.
35. Vernon, P. A. & Mori, M. (1989). Intelligence, reaction times, and peripheral nerve conduction velocity. Intelligence, 16, 273-288.
36. Reed, T. E. & Jensen, A. R. (1991). Arm nerve conduction velocity (NCV), brain NCV, reaction time, and intelligence. Intelligence, 15, 33-47.
37. Deary, I. J. & Caryl, P. G. (1997). Neuroscience and human intelligence differences. Trends in neuroscience, 20, 365-371.
38. Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. Psychological Review, 109, 116-136.
39. Hunt, J. M. (1961). *Intelligence and Experience*. New York, NY: Ronald Press Company.
40. Gregory, R. L. (2003). Seeing after blindness. Nature Neuroscience, 6, 909-910.
41. von Senden, M. (1932). Space and Sight: The Perception of Space and Shape in the Congenitally Blind Before and After Operation. London, UK: Methuen.