Advances in the assessment of cognitive skills using computer-based measurement

Joseph C. Kush · Michael B. Spring ·
Jonathan Barkand

Abstract The impact of technology in the field of intellectual assessment has, for the most part, been limited to computerized administration and scoring. Anderson's (2001) theory of intelligence suggests that performance on traditional IQ measures is acquired via two main routes, thinking and dedicated processing systems known as modules. Empirical data used to support this relationship between the speed of basic processing and intellectual functioning have been evidenced primarily by correlations between measures of general intelligence and measures of inspection time (IT). These IT measures allow individuals to make a forced choice discrimination task without a motor component. Because only the time used to cognitively solve the problem is recorded, these responses typically occur in milliseconds. Many theorists (e.g., Burns & Nettelbeck, 2002; Deary, 2000; Jensen, 2006) consider IT to be a more "pure" measure of intelligence, because the influences of verbal skills, memory, and socialization are minimized and results are therefore considered to be more culture-fair. Until relatively recently, IT measurement was restricted to complex and expensive specialized laboratory equipment. This article describes the theoretical background and developmental process of a computer-based IT measure that is easily adaptable to accommodate the needs of the researcher.

J. C. Kush (✉) · J. Barkand
Duquesne University, 327 Fisher Hall,
600 Forbes Avenue,
Pittsburgh, PA 15282, USA
e-mail: kush@duq.edu

M. B. Spring
University of Pittsburgh,
Pittsburgh, PAUSA

Introduction

The importance of test fairness and test bias continues to be a matter of paramount importance to psychologists and special educators. Each year, between 1.5 and 1.8 million intelligence tests are administered to children in the United States as part of special education eligibility evaluations (Gresham & Witt, 1997). The use of traditional standardized tests of cognitive ability has a long and controversial history. Critics point to test questions that are often characterized as culturally biased or heavily verbally loaded or that overlap with measures of achievement (Gardner, 1983; Gould, 1996; Murdoch, 2007). Despite these criticisms, cognitive ability tests are among the most commonly used measures of assessment used by psychologists in clinical and educational settings.

Psychologists have debated how best to characterize and quantify intelligence since the beginning of the last century (Fasko, 2000; Sattler, 2001). As early as 1883, Galton used sensory motor tasks to look for the simplest properties of the nervous system that would explain individual differences in intellectual ability. Galton's work led directly to one of the earliest theoretical arguments involving the nature of human intelligence—specifically, how many cognitive abilities exist? Spearman's (1904) two-factor model of intelligence described specific, or s, factors (primary factors), as well as a second-order general factor, or g. Expanding upon this idea, Thorndike (1924) characterized intelligence as consisting of several unique factors, while, later, Thurstone (1938) expanded the model to include seven uncorrelated factors. Finally, Guilford (1967) hypothesized three broad

intelligence factors (operations, content, and products) defined by an unworkable 120+ specific or unique factors.

In contrast to the theoretically based models of intelligence proposed by the London school, Alfred Binet was driven by a more pragmatic notion of using cognitive tasks for predicting school or occupational success, often referred to in current times as the *specificity doctrine*. The earliest version of his tests included questions that solely measured specific pieces of knowledge and learned skills. Later Binet and Simon, in 1916, defined intelligence to include "judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence" (as cited in Sattler, 2001, p. 136). Binet's definition suggested that intelligence was not fixed and that the provision of appropriate or remedial schooling could modify intelligence. From a slightly different perspective, David Wechsler, in 1958, defined intelligence as "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" (Wechsler, 1958, p. 7). Wechsler and Binet shared the view that such a thing as general intelligence (*g*) does exist and believed that it operates across a wide spectrum of cognitive functions and human behaviors. In this regard, Binet's original aspiration in creating the first widely available intelligence test was to assess a student's likelihood of succeeding in school, a goal that continues to the present time (Kamphaus, 2001).

The publication of the bell curve in 1994 (Herrnstein & Murray, 1994) transformed the discussion about what constitutes intelligence from academic circles into public controversy. At the time, most of the public, as well as reviews of the book, were highly critical, frequently challenging the scientific rigor of the conclusions (Devlin, Fienberg, Resnick & Roeder 1997; Gould, 1996; Kincheloe, Steinberg, & Greeson, 1997). However, in a subsequent *Wall Street Journal* editorial statement later that year, 52 prominent researchers agreed that the conclusions of the bell curve were basically accurate. This sentiment was subsequently echoed 2 years later in an official task force report prepared by the American Psychological Association (Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci and Urbina 1996).

Intelligence and IQ tests

Many currently available, commercial tests of intelligence have been increasingly criticized for their lack of a strong theoretical foundation (Kush, 1996). Successive revisions of many IQ tests have evolved to measure increasing numbers of "intelligences." For example, the newest editions of both the Stanford Binet (Stanford Binet V;

Roid, 2003) and the Wechsler Intelligence Scale for Children–Fourth Edition (WISC–IV; Wechsler, 2003) now claim to measure multiple intelligence "factors," with only implicit corresponding changes in the theories underlying the tests. Because it remains unclear how many "types" of intelligence are being measured by these scales, there currently exists marked disagreement among psychologists regarding their level of diagnostic interpretability (Canivez, 2008; DiStefano & Dombrowski, 2006; Kush, 1996; Watkins, 2006; Watkins, Lei, & Canivez, 2007). While several theories of intelligence suggest that Spearman's (1904, 1927) general, *g* factor underlies all aspects of the construct (e.g., Brand, 1996; Jensen, 1998), the most widely accepted current views suggest that multiple independent mental abilities are nested under this general factor (Carroll, 1993; Gustafsson, 1994; Horn, 1988).

To a large extent, the current unitary versus multidimensional intelligence debate appears to be best represented by Carroll's (1993, 2003) three-stratum theory. Specifically, Carroll (2003) theorized that human cognitive abilities exist at three levels, or strata, that include a first, lower-order stratum comprising some 50– 60 narrow abilities, a second stratum comprising approximately 8–10 or more broad abilities, and a third still higher stratum containing a single, general intellectual ability commonly represented as *g*. Subsequent research has, however, shown that these strata do a poor job capturing first-order variance, while the *g* factor captures the greatest share of the common and total variance when apportioned with Schmid-Leiman's (1957) procedure (Watkins, 2006). Recent alternatives to Carroll's work are the models proposed by Gustafsson (Gustafsson & Undheim, 1996), wherein G (General) and Gf (G-Fluid) are held to be identical, and by Gignac (2005, 2006, 2008), who has proposed a bi-factor model where psychometric *g* directly relates to subtests, rather than being fully mediated through first-order factors. As a result, the modern day question may now be how well intelligence tests measure general intelligence *as well as* broad ability factors.

While research in intelligence conducted during the first half of the 20th century was heavily influenced by factor analysis, multifaceted descriptions of intelligence have more recently been offered that include noncognitive abilities as indices of intelligence. These theories have been influenced by the disciplines of cognitive psychology, neuropsychology, biopsychology, and cultural psychology. These include Sternberg's (1997) triarchic theory (how people solve problems, how we adapt to our environment, and how we use past experiences to solve problems), Gardner's (1983) theory of multiple intelligences, and Goleman's (1995) theory of emotional intelligence (EI). Each of these theories attempts to expand intelligence theory beyond the prediction of academic achievement to include possible psychological, personality, or environmental

influences in exploring the manner in which intelligence develops. For Gardner, intelligence includes different domains of activities, while in contrast, Sternberg attempts to define intelligence with underlying psychological processes. Relatedly, two slightly different forms of EI have been posited, with EI defined either as a type of intelligence connected to the cognitive processing of emotions or, in contrast, as grouping intelligence with personality and motivational factors. While these newer, alternative models have stimulated a renewed debate over the nature of intelligence these theories have a limited history of from 15 to 20 years, as compared with the 100 years of research examining traditional psychometric models of IQ and intelligence. Rather than a replacement of traditional cognitive models, it is possible that these newer theories may reflect something of a threshold level of cognitive ability, with minimal incremental improvement beyond a level that defines adequate competence (Nettelbeck & Wilson, 2005a). That is, some minimum level of EI, for example, may be necessary for adequate interpersonal communication and social exchanges; however, advanced levels of the construct may not correspond to any increased competence or noticeable behavioral changes.

The search for a more precise definition

The move toward a more parsimonious definition of intelligence has been termed *reductionism*. Typically, reductionist research can be understood through two broad models. In the first model, which emphasizes neural efficiency, researchers have examined psychometric intelligence functions related to the brain. Established physiological correlates with intelligence include electroencephalograph estimates, brain size and weight, nerve conduction velocity, and brain glucose metabolism (Deary & Stough, 1996). Additionally, Kail (2000) has argued that reaction time improvements with age are, in part, due to myelination and changes in the number of synaptic connections, both of which change over childhood and adolescence. Relatedly, cognitive aging involves slowing of processing speed (Salthouse, 1996), which has been found to be associated with decreases in myelination in older adults (Bartzokis, 2004).

Research examining intelligence as a property of the brain has shown that smarter people work more quickly and efficiently because they transmit electrical nerve impulses faster and use less glucose (Gottfredson, 2000). Historically, the speed at which a problem can be solved (strategy free) has commonly been thought of as an index related to overall cognitive ability (Nettelbeck, 1998, 2001). Early attempts to measure learning speed failed, however, to distinguish between the time the individual took to "mentally" solve the problem and the amount of "physical"

time it took for the individual to provide the correct answer. Only when scientists began to consider how technology could be integrated into the solution was a new paradigm for assessment created. For example, the Hick paradigm utilizes the measurement of simple or choice reaction time. Simple reaction time reflects the amount of time it takes a subject to lift their finger from a "home" button following the presentation of a single light. Choice reaction time is similar; however, the subject does not know which of several lights will go on. When prompted, the subject still lifts his finger as quickly as possible off the home button (reaction time), but he or she is now required to press a second button that is paired with the light (movement time). Research utilizing the Hick paradigm has consistently confirmed what has come to be known as Hick's law: Reaction time increases linearly as a function of the number of bits of information (lights to choose from) made available to the subject (Bates & Rock, 2004; Deary & Stough, 1996; Jensen, 1982; Nettelbeck, 1987; Vernon, 1987).

A second area of chronometric research has focused on the cognitive processes utilized by subjects and attempts to identify the relationship between these basic psychological tasks and scores from traditional IQ measures. Recently, researchers have begun to identify the particular psychological processes that are involved in the performance of speeded cognitive tasks (Deary, 2000). The selection of these elementary tasks is typically based on the assumption that they are so simple or elemental in nature that any subject could solve the tasks, given unlimited time. While Vickers is typically credited with the original theoretical development of cognitive speed or inspection time (IT) research, one of his doctoral students, Nettelbeck, conducted some of the earliest empirical studies (Nettelbeck & Lally 1976). Although the theoretical rationale underlying this field of research is almost 30 years old (Vickers, 1970, 1979), the sophistication and accuracy of empirical studies has paralleled advances in technology and has been born of cross-disciplinary discussion and collaboration.

Inspection time origins

The distinguishing factor of the IT paradigm is that the cognitive or mental time needed to solve the problem is recorded separately from the time needed to physically indicate the correct answer. Anderson and Miller (1998), have defined IT as "the stimulus exposure duration required by a subject to make a simple perceptual judgment, for example, the relative length of two lines" (p. 239). That is, IT is the minimum exposure time required to make an accurate determination concerning some highly evident feature of a stimulus under backward-masking conditions. Because of the elemental nature and simplicity associated

with IT measures, they reflect a stronger connection to biological processes than do psychometric test scores (Jensen, 2006). The interval of time between that point and when the answer is provided is characterized as the movement time. The first interval, IT, reflects the amount of time that the subject needs to cognitively solve the problem. This subsequent interval is considered more important to understanding the speed of problem solving than is the second interval, which merely reflects the amount of time the subject needs to motorically indicate their answer.

The genesis of research that began to examine cognitively based IT has paralleled technological advances that have been incorporated into research designs. In the majority of IT studies, the stimulus has consisted of a pi figure (two vertical lines, one long and one short, connected by a horizontal line at the top), where subjects were instructed to select the side of the figure that contained the longer leg. To prevent afterimages or storage in iconic memory, a backward mask consisting of longer and wider bars has been placed over the vertical lines immediately after the presentation. Examples of these stimuli are shown in Fig. 1.

Typically, to begin a trial, subjects are instructed to hold down a button that will randomly present one of the stimuli. When the subject determines which leg is longer, he or she is instructed to press one of two other buttons (on either side of the first button) that corresponds to the side of the figure that contains the longer line. As indicated previously, given unlimited time, even very young children can make this discrimination with nearly perfect accuracy. In IT research in the early 1980s, these measurements were made using a tachistoscope, and research has begun to include computer-based IT measures. Computers have subsequently been used both to display the stimuli (Anderson, 1986; Deary, 1999; Deary, Caryl, Egan, & Wight, 1989) and to determine stimulus exposure thresholds, using adaptive staircase procedures (e.g., Egan, 1994).

In a recent meta-analysis of over 90 studies and 4,200 subjects, Grudnik and Kranzler (2001) showed that IT has consistently been found to correlate with standardized measures of intelligence—most often, performance IQ.

This finding is consistent with an abundance of evidence from other researchers (Burns & Nettelbeck, 2002; Grudnik & Kranzler, 2001; Kranzler & Jensen, 1989; Nettelbeck, 1987, 2001). The best estimates of this relationship suggest that the correlation coefficient is in the region of -.50, although Deary (2000) has proposed that the strength of the relationship may be as high as -.75. Most cognitive theorists posit that the relationship between IT and IQ is causal, with individual differences in IT producing individual differences in IQ. Specifically, speed of processing is believed to be the attribute underlying this causal relationship (Anderson, 1992; Neubauer, 1997).

Most attempts to place IT into a broad theoretical framework have connected the construct with the Horn–Cattell or Cattell, Horn, Carroll models of intelligence (Carroll, 1993, 2003; Horn, 1988); Gf (fluid)–Gc (crystallized) theory has described IT as a factor called *correct decision speed*—the speed at which people provide correct answers to a variety of tests (Carroll, 1993, 1995). However, it has long been recognized (Eysenck, 1987; Vigil-Colet & Codorniu-Raga, 2002) that not all measures of processing speed are equally correlated with general intelligence. Researchers such as Mackintosh (1998) and Burns, Nettelbeck, and Cooper (1999) have suggested that the relationship between IT and IQ exists not because IT correlates with a general factor of intelligence but, rather, because it measures a unique component characterized as cognitive speed. IT has been linked to a range of cognitive abilities, including fluid reasoning, visualization, and short-term memory (Grudnik & Kranzler, 2001).

While a growing body of research (e.g., Nettelbeck & Wilson, 1985, 2004) has suggested that IT improves throughout childhood and declines in the elderly (Nettelbeck, 1987; Nettelbeck & Wilson, 2004, 2005b; Nettelbeck & Young, 1989, 1990), Anderson (1986, 1992, 2001) has argued that the speed of processing actually remains constant across the lifespan. While Nettelbeck and Wilson (1985) suggested that longitudinal changes in IT are due to maturation, Anderson argued that young children perform poorly on cognitive measures because they are more affected by task demands and task experience; reducing the load on attention is thought to improve the performance of younger children (Anderson, 1986, 1992). Additionally, Anderson, Reid, and Nelson (2001) demonstrated that a single exposure to an IT task resulted in an improvement in IT scores 1 year later that was greater than the effect of 1 year's aging. Anderson posited that the development of age-related learning strategies, rather than maturation, underlies this observed developmental trend. As Brody (2001) has suggested, IT research may "provide the intellectual foundation for the development of a method for experimental interventions designed to increase intelligence" (p. 540). This suggestion has been supported by several additional studies

Fig. 1 Geometric pi figures

that propose, yet not conclusively, that IT may be malleable (Bors, Stokes, Forrin, & Hodder, 1999).

Methodological problems associated with inspection time

While the goal of including a backward mask is to prevent storage of the initial stimuli in iconic memory, the introduction of the mask has actually been found by some researchers to produce an unfortunate new problem—the addition of apparent movement. Specifically, some participants have reported that they have been able to use the apparent movement of the mask (covering the shorter leg of the test stimuli) as a cue or strategy to facilitate their performance (Chaiken & Young, 1993; Egan, 1986, 1994; Nettelbeck, 1982). Because the goal of IT research is to eliminate influences other then mental speed, it is critical that measures attempt to eliminate this movement effect. That is, IT should be as pure a measure as possible of the speed of transmission of the central nervous system—specifically, the rate at which oscillation between excitation and refractoriness occurs (Jensen, 1982).

Attempts to eliminate or reduce apparent movement artifacts have proven inconsistent (Chaiken & Young, 1993; Nettelbeck, 1982; Vickers, Nettelbeck, & Wilson, 1972). For example, some researchers have devised a form of mask that uses randomly selected patterns or visual noise (e.g., LED lights) to counter the use of the visual cue. Additionally, some researchers have employed a dynamic mask that minimizes, if not entirely removes, movement artifacts (Evans & Nettelbeck, 1993; Knibb, 1992; Stough, Bates, Mangan, & Colrain, 2001). Bors et al. (1999) have pointed out that most of these studies used adult populations with widely different ages, and older participants have been shown to be less likely to employ strategies in other, similar cognitive tasks (Craik, 1977). Furthermore, when movement artifacts are found, they may actually represent preconscious, rather than meta-cognitive, processes (Egan & Deary, 1992).

Additional procedural limitations have resulted in very few studies of IT with young children. Although presumed to be simple enough even for young children, the standard pi figure stimulus has proven to be difficult for many young children. As a result, a number of modifications to the stimuli have been recommended, including coloring the lines or presenting the stimuli in a game-type format such as alien antenna of different lengths or the Benny Bee IT task (Williams, Turley, Nettelbeck, & Burns, 2009). Additionally, research with young children has shown that many display left and right confusion and, in some cases, demonstrate considerable difficulty remaining on task. Expanding what is known regarding the relationship between IT and intelligence to include children is very

important, since it may be possible to identify early learning difficulties or strengths in specific areas of intelligence, such as information-processing speed (Deary, 2000; Nettelbeck, 2001). However, to date, IT studies with young children remain sparse.

Finally, as was noted by McCrory and Cooper (2007), very few empirical studies have examined whether different presumed measures of IT actually measure the same construct. The most commonly held current belief is that performance on IT measures reflects some basic biological process in addition to some task-specific individual differences, skills, or strategies. Because current extant IT research has included widely varying modes of presentation (each reflective of different cognitive demands), it remains difficult to determine how specific skills associated with IT fit within the larger paradigm of processing speed.

Architecture of the solution

This article describes a prototype software program that utilizes a PC to assess reaction time and IT. The current program was written in Java programming language and is designed to run on any platform that supports Java Runtime Environment (JRE; i.e., Mac, Unix, Windows, etc.). Jensen (2006) has cogently pointed out the advantages of a standardized computer program to assess the IT paradigm particularly in the elimination of method variance. Specifically, computer usage allows for a uniform stimulus display screen and subject response console (keyboard), both of which are critical for IT standardization. Additionally, the portability of the program allows researchers to select computers connected to monitors with specific vertical refresh rates, a factor that has been found to influence minimal exposure durations.

Following traditional chronometric research, the current IT program utilizes the following sequence. Subjects are first presented a short introduction to the task (direction slides can be easily modified for the developmental age of the participants), provide basic demographic information, and complete practice trials. Video directions can also be inserted so that the influence of reading as a confounding variable is minimized or eliminated. At the beginning of the assessment phase, subjects are briefly exposed to an orienting dot in the middle of the screen to alert them that the trial is about to begin. Subjects are then instructed to hold down the space bar, which, in turn, presents either of the two stimuli. Finally, subjects are told to release the space bar and press the correct Alt key on either side of the space bar, as soon they know the answer. IT is the interval of time between when the stimulus is presented on the screen and the time at which the space bar is released by the subject. The length of time it takes the subject to move his/ her finger from the space bar to the answer key is

considered movement time and is not included in the measurement of IT. Adults typically make this decision in milliseconds, and their IT is the average speed of all correctly completed trials. The recording of responses in milliseconds reflects measurement on a ratio scale, a feature that offers numerous psychometric advantages (Jensen, 2006).

The program allows the researcher to easily modify both the stimuli and the test conditions to accommodate a wide range of subjects and research questions. Specifically, the program allows the researcher to individualize the instructions to the subjects and create unique stimuli slides, as well as unique blank, focus, and mask slides. Additionally, the length of time the focus slide and blank slides are presented can be easily adjusted. The ability to provide standardized instructions is one of the most basic steps in the psychometric standardization of a psychological test. This feature allows the researcher to select testing conditions that are task specific or that are based on the developmental age of the subject. Researchers can make use of a widely available program, such as PowerPoint, to create and import these individualized images and can then import their results in a variety of image formats—for example, TIFF, PNG, GIF, JPEG, and so forth. The availability of this feature will allow researchers to better examine the use of specific cognitive strategies by systematically manipulating the complexity of the stimulus (see Alexander & Mackenzie, 1992; Bowling & Mackenzie, 1996; Frings & Neubauer, 2005). At this point, the program allows for the modification only of visual stimuli, although it is anticipated that a future version will allow for auditory stimuli as well.

Because the stimuli slides can be easily modified, researchers will also be able to directly compare subject performance on the standard pi stimuli with the speed at which differences between other stimuli, such as letters, colors, or geometric shapes, are processed. The presentation of a variety of IT stimuli will allow researchers to better understand the underlying relationships between IT and specific cognitive processes. This feature may prove particularly valuable for the assessment of IT in young children, since both the directions and the complexity of the stimuli can be easily modified.

An additional unique feature of the program is the ability to easily modify the stimuli mask to minimize the influence of aforementioned movement cues. For example, in one condition, subjects could be exposed to the standard pi-shaped geometric figures that are, in turn, followed by the larger mask patterns. In an alternate condition, subjects would then be exposed to two different stimuli—that is, an uppercase letter A, and a lowercase letter a. The position of these letters would be randomized so that the uppercase letter appears equally often on the right or left side of the screen. A pair of uppercase S letters could represent the subsequent mask. The letters A and S could be specifically

chosen so that one letter was composed mainly of straight lines while the other letter consisted of curves, thereby producing enough variability that the initial letters would not be stored in iconic memory, yet would minimize movement strategies. The ability to manipulate the stimulus mask (see Fig. 2) will greatly assist researchers attempting to eliminate the movement cues often associated with earlier IT research.

After the researcher has identified and loaded the two stimuli slides, he or she then specifies the number of trials. The program records both IT and movement time for each presented stimuli, as well as identifying the stimulus presented and the answer provided. The program incorporates an adaptive staircase procedure such that the researcher can easily modify the number of iterations in which stimuli are presented, as well as the percentage of correct responses representing a successful set of responses. After displaying the stimuli for a fixed period of time, the system waits for the subject to indicate the correct answer before presenting the next stimulus. The timed format allows the researcher to define the experiment with multiple exposure times. For example, the researcher could define the number of stimuli to be presented in each cycle—for example, 10. Second, an initial exposure time is set—for example, 500 ms. Next, the exposure time decrement is set–for example, e.g. 100 ms–and the number of iteration sets at that decrement amount. Finally, a second decrement time can be set—for example, decrement by 25 ms rather than 100 ms—along with the number of iteration sets at that decrement amount. It is also possible to set an experiment termination if the success rate falls below a researcher-defined minimum. The program automatically provides integrated record keeping and summarization reports.

An advantage of computer-based measures of cognitive processing, such as the current IT program, is that the
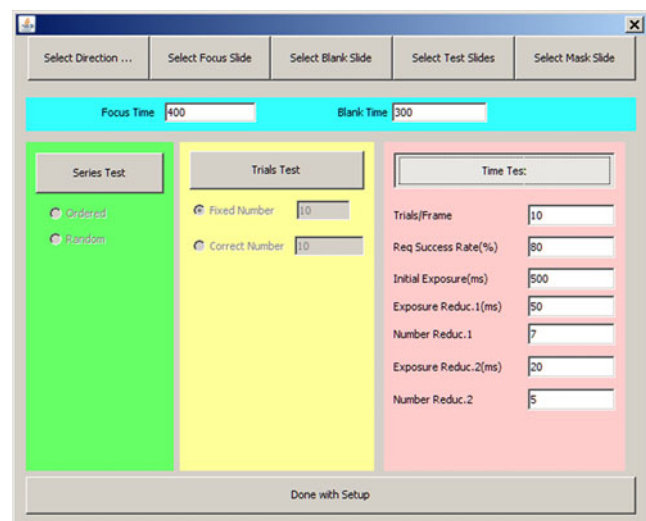

Fig. 2 Experimental setup menu

delivery format matches recent increases in technological usage by the general population. Specifically, there is growing evidence that increased usage of computer applications (including computer games) is associated with higher levels of visual and spatial tasks (Greenfield, 1998). Most computer applications include content that transfers the typical mode of information processing from verbal to visual (Subrahmanyam, Greenfield, Kraut, & Gross, 2001). The increase of nonverbal intelligence that has been demonstrated in recent years (Flynn, 1999) may ultimately prove to be related to increased usage of computer technologies during the same time.

The ease with which the program can be modified addresses many of the desiderata for a standardized measure of IT outlined by Jensen (2006). The standardization of IT apparatus and testing procedures is essential in eliminating method variance and allows for more accurate comparisons of results gathered across different settings. With regard to procedures, as indicated previously, the current program allows for the creation of explicit, uniform instructions and a set number of practice trials. Because the program is platform independent, it will allow researchers to easily compare results across a variety of display screens and response consoles (computer keyboards). In the standardization of the IT apparatus, screen size, background color, luminosity, and refresh rate are all critical factors in increasing the assessment of true task accuracy (Jensen, 2006). The current program also eliminates clerical errors, requires very little storage capacity, and allows for comparisons across several chronometric paradigms, all features Jensen (2006) has highlighted as characteristics necessary for the advancement of the paradigm.

Final thoughts

Unlike previous IT applications, the present IT program may offer users the ability to assess a facet of cognitive skills across a common dimension. At a theoretical level, it remains unclear whether IT is a stable construct across the lifespan or whether the construct parallels the gradual decline in mental performance that results with increasing age (Der & Deary, 2006). Unlike traditional tests of cognitive ability, measures of IT have been shown to be stable among children, and the correlation between IT and intelligence appears to be higher among children who demonstrate sustained attention during the IT task (Nettelbeck & Young, 1989, 1990). However, it is unclear whether process- or task-specific, age-related differences exist in very young children, since the development of many cognitive processes in young children is not linear (e.g., Huttenlocher, Haight, & Bryk, 1991). There is research suggesting that age-associated improvements in IT are mediated by developmental improvements in the ability to engage in

sustained attention (Hutton, Wilding, & Hudson, 1997), and as a result, poorer IT performance in children may reflect not only the obvious influence of distractibility, but also the more subtle aspects of controlling attention, all of which influence cognitive performance in general.

Relatedly, the presence of possible gender differences is equally uncertain (Burns & Nettelbeck, 2005; Gregory, Nettelbeck, Burns, Danthiir, Wilson and Wittert 2010). It also remains possible that IT deficits may be associated with particular developmental impairments, such as learning disabilities (Deary, 2000; Nettelbeck, 2001). A tool that can be used on a standard personal computer and that can be easily modifiable represents an important advancement over previous methods that have required a specialized apparatus, not easily available or affordable to most researchers. Finally, if IT proves to be modifiable over time, psychologists and educators may be able to respond to Brody's (2001) notion that a method for experimental interventions designed to increase intelligence can be designed. This finding would truly be revolutionary.

However, advances in IT research also produce a catch-22 for psychologists. When used as predictors of school success or academic achievement, current commercial IQ tests represent the industry standard (Freberg, Vandiver, Watkins, & Canivez, 2008; Parker & Benedict, 2002; Sattler, 2001), and the practice of comparing IQ and achievement scores represents one of the most common methods for diagnosing learning disabilities (Yen, Konold, & McDermott, 2004). Critics of commercially available IQ tests (Gould, 1996; Murdoch, 2007) have argued that they include content that is influenced by nonintellectual factors, including but not limited to reading ability, socioeconomic status, test-taking strategies, and cultural familiarity. The inclusion of these components improves the predictive power of the instruments because intelligence *and* other factors are being assessed (Watkins et al., 2007). Clearly the knowledge of an individual's cognitive ability, as well as other factors related to academic achievement, offers an advantage over the sole knowledge of intellectual skills. However, as Jensen (1979) pointed out over 30 years ago, intelligence must be distinguished from learning, memory, and achievement. An IQ test that boasts of assessing multiple types of intelligence or IQ factors will also have more "cash validity" than an instrument that measures only a single factor, and most commercial IQ tests are increasingly designed to measure the multiple latent traits of factor analysis. Theories of multiple intelligences have evolved or, perhaps, deevolved to the point where everything (e.g., memory, vocabulary, knowledge of social etiquette, reading ability) is considered by some to reflect a type of intelligence. For example, approximately two thirds of the information contained on the Stanford–Binet (5th edition) and tests of academic achievement reflects shared variance or overlapping

content, a figure that is too high for instruments thought to be measuring related yet discrete constructs (Kush, 2005). Intelligence is related to, but not the same as, academic achievement, and as Naglieri has cogently pointed out (Naglieri & Das, 1997; Naglieri & Rojahn, 2004), most current IQ tests are contaminated with achievement content that confounds their interpretability.

IT measures offer an interesting alternative. The elemental nature of IT (e.g., only one very simple task is included) is an appealing metric for intelligence. However, because cognitive measures that focus solely on how quickly an individual can mentally solve a very simple problem are much more narrowly focused then commercial IQ tests that include tasks of memory and learned knowledge and information, their predictive power is also greatly reduced, since only one task is being assessed. The final resolution may not be an either/or solution but, rather, a combination of the two. The utilization of commercial IQ tests can and should be continued because they assess so many diverse factors. The supplemental inclusion of an IT measure, which may more accurately reflect a specific cognitive ability, could then offer the best of both worlds.

References

Alexander, J. R., & Mackenzie, B. D. (1992). Variations of the 2-line inspection time stimulus. *Personality and Individual Differences, 13,* 1201–1211. doi:10.1016/0191-8869(92)90256-O

Anderson, M. (1986). Inspection time and IQ in young children. *Personality and Individual Differences, 7,* 677–686. doi:10.1016/0191-8869(86)90037-1

Anderson, M. (1992). *Intelligence and development: A cognitive theory.* Oxford: Blackwell.

Anderson, M. (2001). Annotation: Conceptions of intelligence. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 287–298. doi:10.1017/S0021963001007016

Anderson, M., & Miller, K. L. (1998). Modularity, mental retardation, and speed of processing. *Developmental Science, 1,* 239–245. doi:10.1111/1467-7687.00037

Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: What a difference a year makes. *Intelligence, 29,* 475–486. doi:10.1016/S0160-2896(01)00073-3

Bartzokis, G. (2004). Age-related myelin breakdown: A developmental model of cognitive decline and Alzheimer's disease. *Neurobiology of Aging, 25,* 5–18.

Bates, T. C., & Rock, A. (2004). Personality and information processing speed: Independent influences on intelligent performance. *Intelligence, 32,* 33–46. doi:10.1016/j.intell.2003.08.002

Bors, D. A., Stokes, T. L., Forrin, B., & Hodder, S. L. (1999). Inspection time and intelligence: Practice, strategies and attention. *Intelligence, 27,* 111–129. doi:10.1016/S0160-2896(99)00010-0

Bowling, A. C., & Mackenzie, B. D. (1996). The relationship between speed of information processing and cognitive ability. *Personality and Individual Differences, 20,* 775–800. doi:10.1016/0191-8869(96)00022-0

Brand, C. (1996). *The g factor: General intelligence and its implications.* Chichester, U.K.: Wiley.

Brody, N. (2001). Inspection time. Past, present and future. *Intelligence, 29,* 537–541. doi:10.1016/S0160-2896(01)00075-7

Burns, N. R., & Nettelbeck, T. (2002). Inspection time in the structure of cognitive abilities: Where does IT fit? *Intelligence, 31,* 237–255. doi:10.1016/S0160-2896(02)00120-4

Burns, N. R., & Nettelbeck, T. (2005). Inspection time and speed of processing: Sex differences on perceptual speed but not IT. *Personality and Individual Differences, 39,* 439–446.

Burns, N. R., Nettelbeck, T., & Cooper, C. J. (1999). Inspection time correlates with general speed of processing but not with fluid ability. *Intelligence, 27,* 37–44.

Canivez, G. L. (2008). Hierarchical factor structure of the Stanford–Binet Intelligence Scales–Fifth edition. *School Psychology Quarterly, 23,* 533–541. doi:10.1037/a0012884

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* New York: Cambridge University Press.

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research, 30,* 429–452. doi:10.1207/s15327906mbr3003_6

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5–21). New York: Pergamon.

Chaiken, S. R., & Young, R. K. (1993). Inspection time and intelligence: Attempts to eliminate the apparent movement strategy. *The American Journal of Psychology, 106,* 191–210.

Craik, F. M. I. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 384–420). New York: Van Nostrand-Reinhold.

Deary, I. J. (1999). Intelligence and visual and auditory information processing. In P. L. Ackerman & P. C. Kyllonen (Eds.), *Learning and individual differences: Process, trait, and content determinants.* Washington, DC: American Psychological Association.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain.* Oxford: Oxford University Press.

Deary, I. J., Caryl, P. G., Egan, V., & Wight, D. (1989). Visual and auditory inspection time: Their interrelationship and correlations with IQ in high ability subjects. *Personality and Individual Differences, 10,* 525–534.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects and problems. *American Psychologist, 51,* 599–608.

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and lifestyle survey. *Psychology and Aging, 21,* 62–73.

Devlin, B., Fienberg, S., Resnick, D., & Roeder, K. (Eds.). (1997). *Intelligence, genes and success: Scientists respond to the Bell curve.* New York: Springer.

DiStefano, C., & Dombrowski, S. C. (2006). Investigating the theoretical structure of the Stanford–Binet–Fifth edition. *Journal of Psychoeducational Assessment, 24,* 123–136. doi:10.1177/0734282905285244

Egan, V. (1986). Intelligence and inspection time: Do high-IQ participants use cognitive strategies? *Personality and Individual Differences, 7,* 695–700. doi:10.1016/0191-8869(86)90039-5

Egan, V. (1994). Intelligence, inspection time and cognitive strategies. *British Journal of Psychology, 85,* 305–316.

Egan, V., & Deary, I. J. (1992). Are specific inspection time strategies prevented by concurrent tasks? *Intelligence, 16,* 151–167. doi:10.1016/0160-2896(92)90002-9

Evans, G., & Nettelbeck, T. (1993). Inspection time: A flash mask to reduce apparent movement effects. *Personality and Individual Differences, 15,* 91–94.

Eysenck, H. J. (1987). Speed of information processing, RT, and the theory of intelligence. In P. A. Vernon (Ed.), *Intelligence and speed of information processing* (pp. 57–85). New York: Ablex.

Fasko, D., Jr. (2000). An analysis of multiple intelligences theory and its use with the gifted and talented. *Roeper Review, 23*, 126–130. doi:10.1080/02783190109554083

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*, 5–20. doi:10.1037/0003-066X.54.1.5

Freberg, M. E., Vandiver, B. J., Watkins, M. W., & Canivez, G. L. (2008). Significant factor score variability and the validity of the WISC–III Full Scale IQ in predicting later academic achievement. *Applied Neuropsychology, 15*, 131–139.

Frings, C., & Neubauer, A. (2005). Are masked-stimuli-discrimination-tests in masked priming studies measures of intelligence?—An alternative task for measuring inspection time. *Personality and Individual Differences, 39*, 1181–1191. doi:10.1016/j.paid.2005.02.024

Galton, F. (1883). *Inquiries into human faculty and its development*. London: MacMillan.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gignac, G. E. (2005). Revisiting the factor structure of the WAIS–R: Insights through nested factor modeling. *Assessment, 12*, 320–329.

Gignac, G. E. (2006). The WAIS–III as a nested factors model: A useful alternative to the more conventional oblique and higher-order models. *Journal of Individual Differences, 27*, 73–86.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: g as superordinate or breadth factor. *Psychology Science Quarterly, 50*, 21–43.

Goleman, D. P. (1995). *Emotional intelligence*. New York: Bantam.

Gottfredson, L. S. (2000). Pretending that intelligence doesn't matter. *Cerebrum, 2*, 75–96.

Gould, S. J. (1996). *The mismeasure of man* (2nd ed.). New York: Norton.

Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve* (pp. 81–125). Washington, DC: American Psychological Association.

Gregory, T., Nettelbeck, T., Burns, N. R., Danthiir, V., Wilson, C., & Wittert, G. (2010). Sex differences in inspection time with age. *Personality and Individual Differences, 50*, 593–596. doi:10.1016/j.paid.2010.11.034

Gresham, F. M., & Witt, J. C. (1997). Utility of intelligence tests for treatment planning, classification, and placement decisions: Recent empirical findings and future directions. *School Psychology Quarterly, 12*, 249–267. doi:10.1037/h0088961

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence, 29*, 523–535. doi:10.1016/S0160-2896(01)00078-2

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Gustafsson, J. E. (1994). Hierarchical models of intelligence and educational achievement. *Advances in Psychology, 106*, 45–73. doi:10.1016/S0166-4115(08)62752-1

Gustafsson, J. E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). New York: Macmillan.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Horn, J. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 645–685). New York: Plenum.

Huttenlocher, J., Haight, W., & Bryk, A. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236–248.

Hutton, U., Wilding, J., & Hudson, R. (1997). The role of attention in the relationship between inspection time and IQ in children. *Intelligence, 24*, 445–460.

Jensen, A. R. (1979). The nature of intelligence and its relationship to learning. *Journal of Research and Development in Education, 12*, 79–85.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer.

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

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.

Kail, R. (2000). Speed of information processing: Developmental change and links to intelligence. *Journal of School Psychology, 38*, 51–61.

Kamphaus, R. W. (2001). *Clinical assessment of child and adolescent intelligence* (2nd ed.). Boston: Allyn and Bacon.

Kincheloe, J., Steinberg, S., & Greeson, A. (Eds.). (1997). *Measured lies*. New York: St. Martin's.

Knibb, K. (1992). A dynamic mask for inspection time. *Personality and Individual Differences, 13*, 237–248.

Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence, 13*, 329–347. doi:10.1016/S0160-2896(89)80006-6

Kush, J. C. (1996). Factor structure of the WISC–III for students with learning disabilities. *Journal of Psychoeducational Assessment, 14*, 32–40.

Kush, J. C. (2005). Review of the Stanford–Binet Intelligence Scales: Fifth edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 979–984). Lincoln: University of Nebraska.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: Oxford University Press.

McCrory, C., & Cooper, C. (2007). Overlap between visual inspection time tasks and general intelligence. *Learning and Individual Differences, 17*, 187–192. doi:10.1016/j.lindif.2007.03.008

Murdoch, S. (2007). *IQ: A smart history of a failed idea*. Hoboken, NJ: Wiley.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System: Interpretive handbook*. Itasca, IL: Riverside.

Naglieri, J. A., & Rojahn, J. (2004). Construct validity of the PASS theory and CAS: Correlations with achievement. *Journal of Educational Psychology, 96*, 174–181.

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

Nettelbeck, T. (1982). Inspection time: An index for intelligence? *Quarterly Journal of Experimental Psychology, 34A*, 299–312. doi:10.1080/14640748208400843

Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 295–346). Norwood, NJ: Ablex.

Nettelbeck, T. (1998). Jensen's chronometric research: Neither simple for sufficient but a good place to start. *Intelligence, 26*, 233–241. doi:10.1016/S0160-2896(99)80006-3

Nettelbeck, T. (2001). Correlation between inspection time and psychometric abilities: A personal interpretation. *Intelligence, 29*, 459–474. doi:10.1016/S0160-2896(01)00072-1

Nettelbeck, T., & Lally, M. (1976). Inspection time and measured intelligence. *British Journal of Psychology, 67*, 17–22.

Nettelbeck, T., & Wilson, C. (1985). A cross-sectional analysis of developmental differences in speed of visual processing. *Journal of Experimental Child Psychology, 40*, 1–22. doi:10.1016/0022-0965(85)90063-3

Nettelbeck, T., & Wilson, C. (2004). The Flynn effect: Smarter not faster. *Intelligence, 32*, 85–93. doi:10.1016/S0160-2896(03)00060-6

Nettelbeck, T., & Wilson, C. (2005a). Uncertainty about the biology of intelligence: A role for a marker test. *Cortex, 41,* 234–235.

Nettelbeck, T., & Wilson, C. (2005b). Intelligence and IQ: What teachers should know. *Educational Psychology, 25,* 609–630. doi:10.1080/01443410500344696

Nettelbeck, T., & Young, R. (1989). Inspection time and intelligence in 6-year old children. *Personality and Individual Differences, 10,* 605–614. doi:10.1016/0191-8869(89)90220-1

Nettelbeck, T., & Young, R. (1990). Inspection time and intelligence in 7-year old children: A follow-up. *Personality and Individual Differences, 11,* 1283–1290.

Neubauer, A. C. (1997). The mental speed approach to the assessment of intelligence. In J. Kingma & W. Tomic (Eds.), *Advances in cognition and education: Reflections on the concept of intelligence* (pp. 149–174). Greenwich, Connecticut: JAI Press.

Parker, D. R., & Benedict, K. B. (2002). Assessment and intervention: Promoting successful transitions for college students with ADHD. *Assessment for Effective Intervention, 27,* 3–24. doi:10.1177/073724770202700302

Roid, G. H. (2003). *Stanford–Binet Intelligence Scales–Fifth edition.* Itasca, IL: Riverside.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Sattler, J. M. (2001). *Assessment of children* (4th ed.). San Diego, CA: Jerome Sattler.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53–61.

Spearman, C. (1904). "General intelligence": Objectively determined and measured. *The American Journal of Psychology, 15,* 201–293.

Spearman, C. (1927). *The abilities of man.* Cambridge: New York.

Sternberg, R. (1997). *Successful intelligence.* New York: Plume.

Stough, C., Bates, T. C., Mangan, G. L., & Colrain, I. (2001). Inspection time and intelligence: Further attempts to eliminate the apparent movement strategy. *Intelligence, 29,* 219–230.

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

Thorndike, E. L. (1924). Measurement of intelligence: The present status. *Psychological Review, 31,* 219.

Thurstone, L. L. (1938). *Primary mental abilities.* Chicago: University of Chicago Press.

Vernon, P. A. (1987). *Speed of information processing and general intelligence.* Norwood, NJ: Ablex.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics, 13,* 37–58. doi:10.1080/00140137008931117

Vickers, D. (1979). *Decision processes in visual perception.* London: Academic Press.

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.

Vigil-Colet, A., & Codorniu-Raga, M. J. (2002). How inspection time and paper and pencil measures of processing speed are related to intelligence. *Intelligence, 33,* 1149–1161. doi:10.1016/S0191-8869(02)00005-3

Watkins, M. W. (2006). Orthogonal higher order structure of the Wechsler Intelligence Scale for Children–Fourth edition. *Psychological Assessment, 18,* 123–125. doi:10.1037/1040-3590.18.1.123

Watkins, M. W., Lei, P.-W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence, 35,* 59–68. doi:10.1016/j.intell.2006.04.005

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Wechsler, D. (2003). *The WISC–IV technical and interpretive manual.* San Antonio, TX: Psychological Corporation.

Williams, S. E., Turley, C., Nettelbeck, T., & Burns, N. R. (2009). A measure of inspection time in 4-year-old children: The Benny Bee IT task. *British Journal of Developmental Psychology, 27,* 669–680. doi:10.1348/026151008X354573

Yen, C.-J., Konold, T. R., & McDermott, P. A. (2004). Does learning behavior augment cognitive ability as an indicator of academic achievement? *Journal of School Psychology, 42,* 157–169. doi:10.1016/j.jsp.2003.12.001