Testing attention: Comparing the ANT with TVA-based assessment

Thomas Habekost · Anders Petersen · Signe Vangkilde

Abstract Posner's attention network model and Bundesen's theory of visual attention (TVA) are two influential accounts of attention. Each model has led to the development of a test method: the attention network test (ANT) and TVA-based assessment, respectively. Both tests have been widely used to investigate attentional function in normal and clinical populations. Here we report on the first direct comparison of the ANT to TVA-based assessment. A group of 68 young healthy participants were tested in three consecutive sessions that each contained standard versions of the two tests. The parameters derived from TVA-based assessment had better internal reliability and retest reliability than did those of the standard version of the ANT, where only the executive network score reached comparable levels. However, when corrected for differences in test length, the retest reliability of the orienting network score equaled the least reliable TVA parameters. Both tests were susceptible to practice effects, which improved performance for some parameters while leaving others constant. All pairwise correlations between the eight attention parameters measured by the two tests were small and nonsignificant, with one exception: A strong correlation ($r=0.72$) was found between two parameters of TVA-based assessment, visual processing speed and the capacity of visual short-term memory. We conclude that TVA-based assessment and the ANT measure complementary aspects of attention, but the scores derived from TVA-based assessment are more reliable.

Keywords Attention network · Parameter independence · Practice effects · Reliability

T. Habekost (✉) · A. Petersen · S. Vangkilde
Department of Psychology, University of Copenhagen,
Øster Farimagsgade 2A,
1353 Copenhagen, Denmark
e-mail: Thomas.Habekost@psy.ku.dk

Selectivity is arguably the central defining quality of attention (Bundesen & Habekost, 2008). At the same time, attention is often regarded as a complex function that involves several different mental processes and abilities (Duncan, 2006). For example, Parasuraman (2000) suggested that attention has three main aspects: selection, vigilance, and control. Selection prioritizes the currently most relevant information, whereas vigilance ensures that the organism remains sufficiently alert to maintain selectivity over time. The control aspect of attention relates to higher-order processes in which goals are flexibly activated, maintained, or deactivated under competition from alternative thoughts or response tendencies. Under slightly varying headings (e.g., selectivity, intensity, and executive attention; van Zomeren & Brouwer, 1994) the tripartite taxonomy of attention is found with many other authors in the literature.

This functional division of attention is also reflected in one of the most influential neurocognitive models in the field, the anatomical network model of Posner and colleagues. According to this model attentional effects in sensory processing areas of the brain are influenced by three supramodal anatomical networks: the orienting, alerting, and executive networks (Posner & Petersen, 1990; Posner & Rothbart, 2007). These three brain networks perform different but interrelated cognitive functions. The orienting network consists of the frontal eye fields, the superior parietal lobe, the temporo-parietal junction, and the superior colliculus. Its function is to enhance information processing at particular spatial locations by focusing attention on one or a few items in the visual field, either overtly or covertly (i.e., with or without foveation). The alerting network encompasses the locus coeruleus, parts of the thalamus, and areas of the frontal and parietal cortex, particularly in the right hemisphere. Alerting is necessary to achieve and maintain a high state of sensitivity to sensory information. The executive network involves the anterior cingulate cortex and the

lateral prefrontal cortex. The function of this anatomical network is to inhibit automatic responses and to monitor and resolve conflicts related to decision making and error detection. The network model of attention is supported by much empirical evidence. In particular, neuroimaging studies have shown selective activation of each network in matched attentional tasks (e.g., Bush, Luu, & Posner, 2000; Corbetta & Shulman, 2002; Coull, 1998; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). However, the model has been criticized for focusing too narrowly on spatially defined selection (Desimone & Duncan, 1995). Also, the notion that visual attention is a serially moving spotlight seems incompatible with the observed performance in some experimental tasks (e.g., whole report; Bundesen, 1999), which can better be explained by parallel-processing models of attention.

Individuals differ in their ability to attend effectively, and the network model of attention has led to the development of a test method to evaluate these differences. To provide simultaneous measures of the efficiency of each network, Fan, McCandliss, Sommer, Raz, and Posner (2002) presented the Attentional Network Test (ANT). The ANT is essentially a combination of two classical attentional paradigms: a flanker task (Eriksen & Eriksen, 1974) and a spatial cueing task (Posner, 1980). In a standard trial, a row of five horizontal arrows is presented and the task is to report as quickly as possible (by keypress) the pointing direction of the center arrow. Two task-irrelevant arrows flank the target arrow at each side. These distractor arrows either point in the same direction as the target (congruent) or in the opposite direction (incongruent). The display is preceded by one of several possible cues, which can either be spatially specific or generally alerting. The efficiency of each network is evaluated by differences in reaction time between cueing conditions. For example, the efficiency of the executive network is estimated by the difference in reaction time between the incongruent and the congruent flanker conditions. The test is easy to use in populations that are otherwise difficult to test (e.g., children, brain damaged persons, and even monkeys) and the validity of its two component tasks, spatial cueing and flanker interference, is very well established in the cognitive literature. An fMRI study has also provided support for the assumption that the hypothesized anatomical networks are differentially activated by the ANT task (Fan et al., 2005). The ANT has now been used in various forms in more than 60 studies of special populations (e.g., those with ADHD, Adólfsdottir, Sørensen, & Lundervold, 2008; dyslexia, Bednarek et al., 2004; or schizophrenia, Wang et al., 2005), within the field of behavioral genetics (Fan, Fossella, Sommer, Wu, & Posner, 2003), and to chart effects of cognitive training (Tang et al., 2007). Whereas the validity of the ANT seems well supported, several studies report that the reliability of the test is questionable, apart from the executive network score. This has been found

for both test–retest reliability (Fan et al., 2002; Fan, Wu, Fossella, & Posner, 2001) and for the internal reliability of scores obtained from a single test session (measured by split-half reliability; Macleod et al., 2010). Test–retest reliabilities for orienting and alerting typically fall in the range from 0.35 to 0.60, with similar scores being reported for internal reliability, whereas the reliability of the executive network score typically falls in the range from 0.65 to 0.80. The modest reliability of the ANT may reduce its statistical power as a research tool. For research purposes one normally requires a test reliability of at least 0.7, and even higher (0.8–0.9) for investigations of single cases or clinical decisions (Guilford & Fruchter, 1978). Apart from the executive network score, the ANT seems to fall short of this criterion.

A different perspective on attention is taken by another well-established model in the field, Bundesen's (1990) theory of visual attention (TVA). Instead of conceptualizing attention as a mental spotlight that moves from location to location in the visual field, TVA describes visual attention as a parallel processing race in which all stimuli compete simultaneously for representation in a short-term memory store of limited capacity. Processing capacity is distributed according to attentional weights, so that objects of high importance are processed faster and thereby have a higher probability of capturing one of the few slots in visual short-term memory. The processing race is modeled in a set of simple equations (see Method section for details). Bundesen (1990) used the equations of TVA to account for a wide range of classical findings on focused and divided attention, including results from whole report, partial report, single stimulus recognition, cued detection, and visual search (see Bundesen & Habekost, 2008, for an updated account of the literature). Bundesen, Habekost, and Kyllingsbæk (2005) went on to show that the same equations can account for many attentional effects at the level of single neurons in the brain. Bundesen et al. also suggested a possible anatomical basis for the computations of the TVA model that involves a dynamic interplay between thalamic and cortical areas of the brain.

According to TVA the course of the visual encoding process depends on five distinct mathematical parameters: the perception threshold, t_0; visual processing speed, C; the storage capacity of visual short-term memory, K; visual distractibility, α; and the relative attentional weight (i.e., proportion of the total available processing capacity) of each visual object, w. The w parameter can be used to compute the relative balance between attentional weights in the left and right visual field, w_{index} (defined as the ratio between left-side weights and weights summed across the whole visual field). Duncan et al. (1999) showed that these five parameters can be measured using a combination of two simple tasks, whole and partial report. In whole report, the participant must try to identify as many stimuli (typically letters) as possible from a briefly shown display. In partial report, the task is to identify only

stimuli with certain features (e.g., a particular color) while ignoring other distracting stimuli. The dependent measure is accuracy, and the TVA parameters can be extracted from the performance data by a maximum likelihood fitting procedure (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011; Kyllingsbæk, 2006). Like the ANT, TVA-based assessment has been used for many studies of special populations. This includes studies of stroke patients (Habekost & Rostrup, 2006; Peers et al., 2005), neurodegenerative diseases (Bublak et al., 2011; Finke et al., 2005), and ADHD (Finke et al., 2011). TVA-based assessment has also been used to chart the effect of pharmacological substances on healthy participants (Finke et al., 2010; Vangkilde, Bundesen, & Coull, 2011) and the effects of meditation (Jensen, Vangkilde, Frøkjær, & Hasselbalch, 2012). In total, around 35 studies have so far used TVA-based assessment for these different purposes. Contrary to the ANT, several studies have found that TVA-based assessment has good internal reliability, except for the α parameter (Finke et al., 2005; Habekost & Bundesen, 2003; Habekost & Rostrup, 2006). However the test–retest reliability of TVA-based assessment has not been systematically investigated.

The serial attention model of Posner and colleagues appears to be fundamentally at odds with the parallel-processing model proposed by Bundesen (1990). However, the two models may to a large extent describe different aspects of attention. Posner's model was developed to describe performance in speeded attentional tasks (e.g., spatial cueing, visual search), where processes related to the selection and execution of manual motor responses influence the results significantly. Data from these paradigms can also be accounted for by TVA (Bundesen, 1990), but this model relates most directly to cognitive processes that occur within a single fixation of the eyes (as measured by recognition accuracy) and have no significant response component. Another difference is that TVA does not describe the effects of alertness or arousal on attention (although work is underway to include this aspect in the model; Vangkilde, Coull, & Bundesen, 2012). On the other hand, the network model includes no description of attentional processing capacity, which is explicitly modeled in TVA (by the C and K parameters). There are however also similarities between the two models. TVA does not include a description of executive control processes (in its basic version; see Logan & Gordon, 2001, for a possible extension of the model into this cognitive domain), but it does have a measure of top-down attentional control, the α parameter, which to some extent is analogous to the ANT's executive network score.

Studies using the ANT or TVA-based assessment have typically aimed to demonstrate specific attentional deficits (or cognitive training effects), which requires that the tests can reliably dissociate different parameters of attention. Although this has worked out well in many studies, there are indications that the parameters measured by each test may not be independent from one another. In the ANT small but significant correlations between alerting and executive network scores (Fossella et al., 2002) and between alerting and orienting (Lehtonen, 2008) have been reported. This has raised concerns that the assumption of independent networks may not hold (Macleod et al., 2010). In TVA-based assessment significant correlations between the C and K parameter have consistently been reported (e.g., Finke et al., 2005; Habekost & Rostrup, 2006; Vangkilde et al., 2011). It is possible that significant correlations between parameters across the two tests may also exist, but so far no direct comparison has been made between the ANT and TVA-based assessment.

Besides functional specificity, another important methodological issue is reliability. The existing evidence indicates that in the ANT only the executive network score has good reliability, both in terms of test–retest reliability and internal (split-half) reliability (Fan et al., 2002; Fan et al., 2001; Macleod et al., 2010). The reliability of TVA-based assessment has been investigated using bootstrap statistics (Finke et al., 2005; Habekost & Bundesen, 2003; Habekost & Rostrup, 2006), which provides estimates of the measurement error related to each attention parameter within a single session (i.e., internal reliability). These analyses have shown good internal reliability except for the α parameter, but no study has so far investigated the test–retest reliability of TVA-based assessment. Practice effects are also important to consider, especially in view of the growing interest in using the two tests for cognitive training studies, where participants are tested repeatedly. Ishigami and Klein (2010) reported significant practice effects for the executive network score over multiple administrations of the ANT, but no similar study has been done for TVA-based assessment.

The aim of the present study was to make the first empirical comparison between the ANT and TVA-based assessment with regards to (1) internal and test–retest reliability, (2) susceptibility to practice effects, and (3) correlations between individual parameters. In line with previous studies, we hypothesized that the reliability of parameters from TVA-based assessment would generally be higher than those of the ANT. For the ANT we expected to see practice effects in the executive network score, but for TVA-based assessment we had no specific predictions on practice effects. Finally, we expected that the attention parameters measured in the two tests would be largely independent from each other, possibly apart from those related to attentional control (i.e., α and the executive network score).

Method

Participants

A group of 68 young healthy participants were recruited through advertisement and paid for participation (mean age= 24.5 years, SD=2.8; 41 women, 27 men). All were university

students from the Copenhagen area, and all had normal or corrected-to-normal vision. Of these participants, 62 were right-handed. All provided written informed consent, and the study was approved by the Committee for Science Ethics at the Department of Psychology, University of Copenhagen (ref: 2012/01).

Procedure

A standard version of the ANT (following Fan et al., 2002) was used (see Fig. 1). In all trials participants had to determine the direction of a central arrow, which could appear above or below a fixation cross. The target arrow was either accompanied by neutral, congruent, or incongruent flankers, and the arrow could be preceded by alerting or spatial cues, or no cue. The ANT comprised three blocks of 96 trials each that varied systematically with regard to *cue* type (no cue, central cue, double cue, or single cue) and *flanker* type (neutral, congruent, or incongruent flankers). The test blocks were preceded by 24 practice trials in each session.

A standard version of TVA-based assessment (see Vangkilde et al., 2011) with a 432-trial combination of whole and partial report was used (see Fig. 2). The test was preceded by 36 practice trials in each session. Participants were instructed to fixate a central cross throughout each trial. After a 1,000-ms delay, a stimulus display consisting of six red target letters, two red target letters, or two red target letters and four blue distractor letters was presented briefly. The display was terminated by pattern masks presented for

500 ms on all possible stimulus positions. When the screen went blank after the masks, the participants made an unspeeded report of all the red target letters they were "fairly certain" of having seen. They responded by typing the letters in any order on a standard keyboard. The stimulus duration of the 216 trials with six-target displays was varied systematically between 10 and 200 ms to cover the full performance range of each participant from threshold levels to near-ceiling scores. The two other display types (108 trials for each condition, respectively) were always presented for 80 ms, an exposure duration that was also selected to avoid both ceiling and floor effects in performance. The different trial types were presented in a randomized fashion in 12 separate blocks, each of which consisted of 36 trials. After each block, subjects were informed about the accuracy of their responses. Subjects were instructed to aim for a response accuracy between 80 % and 90 %.

Each participant was tested in three consecutive sessions of about one hour's length, with one week between. For 48 of the participants, the sequence of testing in each session was TVA-based assessment followed by ANT (plus another detection task; not reported in this article). To control for ordering effects, 20 other participants in the study received the two tests in reversed order.
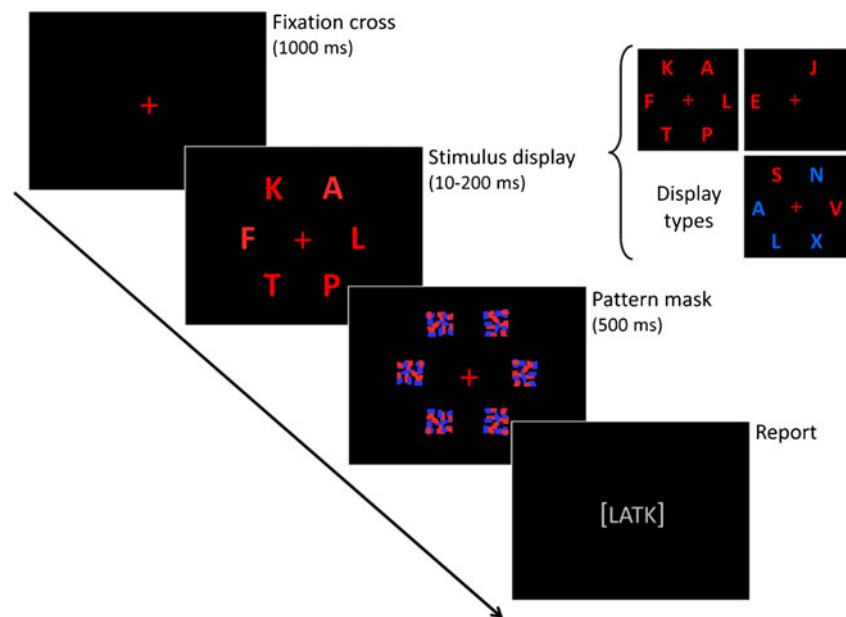
Data analysis and modeling

Computation of ANT scores The reaction time data from the different cueing conditions of the ANT test were used to calculate an alerting network score (= $RT_{no\ cue} - RT_{ctr\ cue}$),
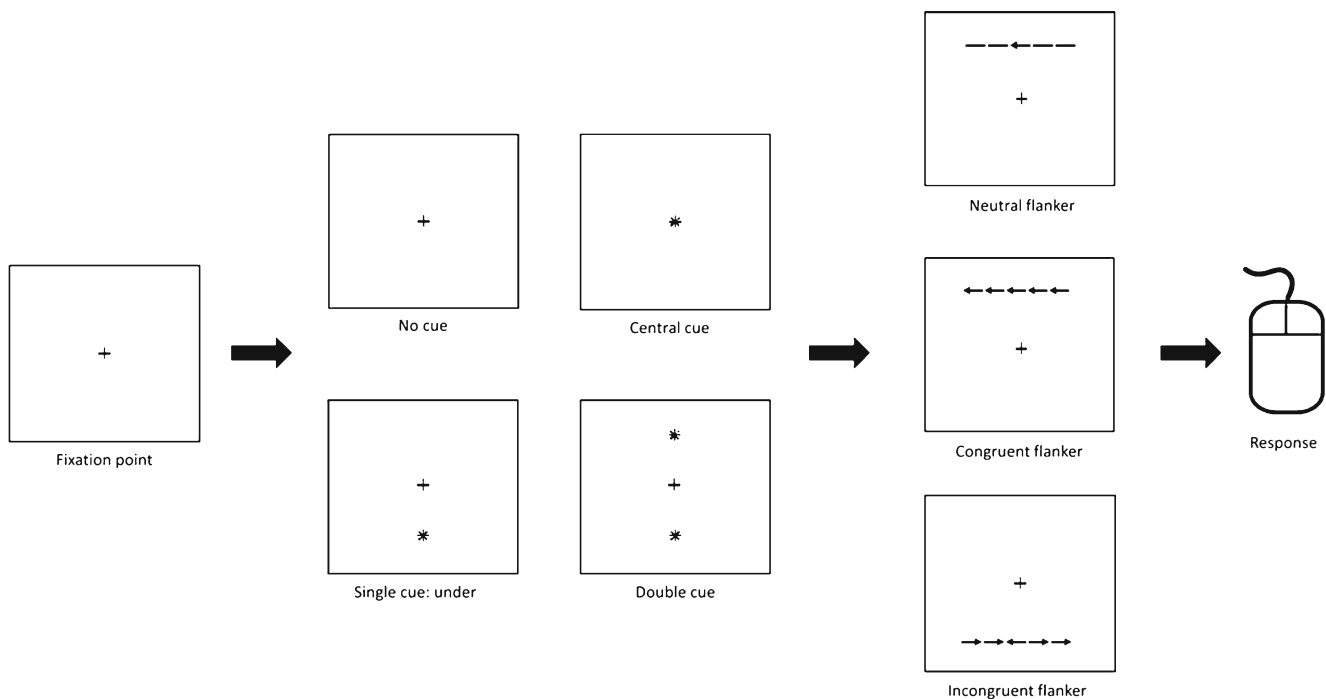
Fig. 1 Outline of a single trial of the attention network test (ANT)

Fig. 2 Outline of a single trial of theory of visual attention (TVA)-based assessment

orienting network score (= $RT_{ctr\ cue} - RT_{ornt\ cue}$), and executive network score (= $RT_{incon} - RT_{con}$) in each session for each participant.

Basic TVA equations The TVA model assumes that recognition and selection of objects in the visual field consist in making *visual categorizations*. A visual categorization has the form "object x belongs to category i." The categorization is selected when the categorization is encoded into visual short-term memory (VSTM). When a visual categorization of an object completes processing, the categorization enters VSTM if memory space for the categorization is available in this system. The capacity of VSTM is assumed to be limited to K different objects. K varies from individual to individual and is one of the main parameters of TVA. By the *rate equation* of TVA, the rate, $v(x, i)$, at which a particular visual categorization, "x belongs to i," is encoded into VSTM is given by a product of three terms:

$$v(x,i) = \eta(x,i)\beta_i\left\{\frac{w_x}{\sum_{z\in S} w_z}\right\}. \tag{1}$$

The first term, $\eta(x, i)$, is the strength of the sensory evidence that x belongs to category i. The second term, β_i, is a perceptual decision bias associated with category i ($0 \leq \beta_i \leq 1$). The third term is the relative attentional weight of object x—that is, the weight of object x, w_x, divided by the sum of weights across all objects in the visual field, S. The total visual processing speed, C, is defined as the sum of all v values across all perceptual categorizations of all elements in the visual field:

$$C = \sum_{x\in S} v(x) = \sum_{x\in S}\sum_{i\in R} v(x,i). \tag{2}$$

The attentional weights in the rate equation of TVA are derived from *pertinence* values. Every visual category j is supposed to have a pertinence π_j, which is a nonnegative real number. The pertinence of category j is a measure of the momentary importance of attending to objects that belong to category j. The attentional weight of an object x in the visual field is given by the *weight equation* of TVA,

$$w_x = \sum_{j\in R}\eta(x,j)\pi_j, \tag{3}$$

where R is the set of all visual categories, $\eta(x, j)$ is the strength of the sensory evidence that object x belongs to category j, and π_j is the pertinence of category j. By the weight equation, the attentional weight of an object is a weighted sum of pertinence values. The pertinence of a given category enters the sum with a weight equal to the strength of the sensory evidence that the object belongs to the category.

Modeling whole- and partial-report data using TVA According to TVA each object (letter) in a whole report display is processed independently and in parallel with the other items. The object is processed at a constant rate (given by its v value) during the effective exposure duration, which equals the physical exposure time, t, minus the perception threshold, t_0. During this time period, provided there are available slots in the VSTM system (i.e., less than K objects have completed processing), the object has a constant probability of being encoded at each instant. This corresponds to an exponential distribution of the encoding time. In partial report, the stimuli fall into two categories: targets (red letters) and distractors (blue letters). It is assumed that every target has approximately the same attentional weight as any other target, and every distractor has approximately the same

attentional weight as any other distractor. In such cases, the efficiency of top-down selection can be defined as the ratio, α, between the attentional weights of a distractor and a target:

$$\alpha = w_{\text{distractor}} / w_{\text{target}}. \tag{4}$$

Given these general assumptions, and assuming a particular set of parameter values, one can calculate the probability of any possible outcome of a whole or partial report trial (e.g., the probability that stimuli are encoded at Positions 1, 2, and 4, but not at Positions 3, 5, and 6; see Kyllingsbæk, 2006, and Dyrholm et al., 2011, for details). This implies that for a given individual, one can estimate the set of parameter values that maximize the probability of obtaining the total set of observed data. In our study, the individual data from TVA-based assessment were fitted by an improved maximum likelihood fitting procedure using the LibTVA toolbox for MATLAB by Dyrholm et al. We used a model with 14 free parameters given by a number of assumptions: First, it was assumed that trial by trial the storage capacity of visual short term memory (K) was drawn from a probability distribution characterized by five free parameters (i.e., the probabilities that $K=1, 2, \ldots, 5$), where these five probabilities summed to a value between 0 and 1, and the remaining probability up to a value of 1 was accounted for by the probability that $K=6$. Hence the K value reported in the results section is the expected K given a particular probability distribution. In a similar fashion, the perceptual threshold, t_0, was assumed to be drawn trial by trial from a normal distribution with a given mean and standard deviation (two free parameters). In contrast, the speed of visual processing, C, was a constant (one free parameter). In addition, attentional weights (w values) were estimated individually for targets at each of the six stimulus locations (five free parameters, as the sum of the six attentional weights was fixed at a value of 1). Finally, top-down controlled selectivity was modeled by a single free parameter, α, that was defined as the ratio of the attentional weight of a distractor on a certain spatial location to the attentional weight of a target on the same location.
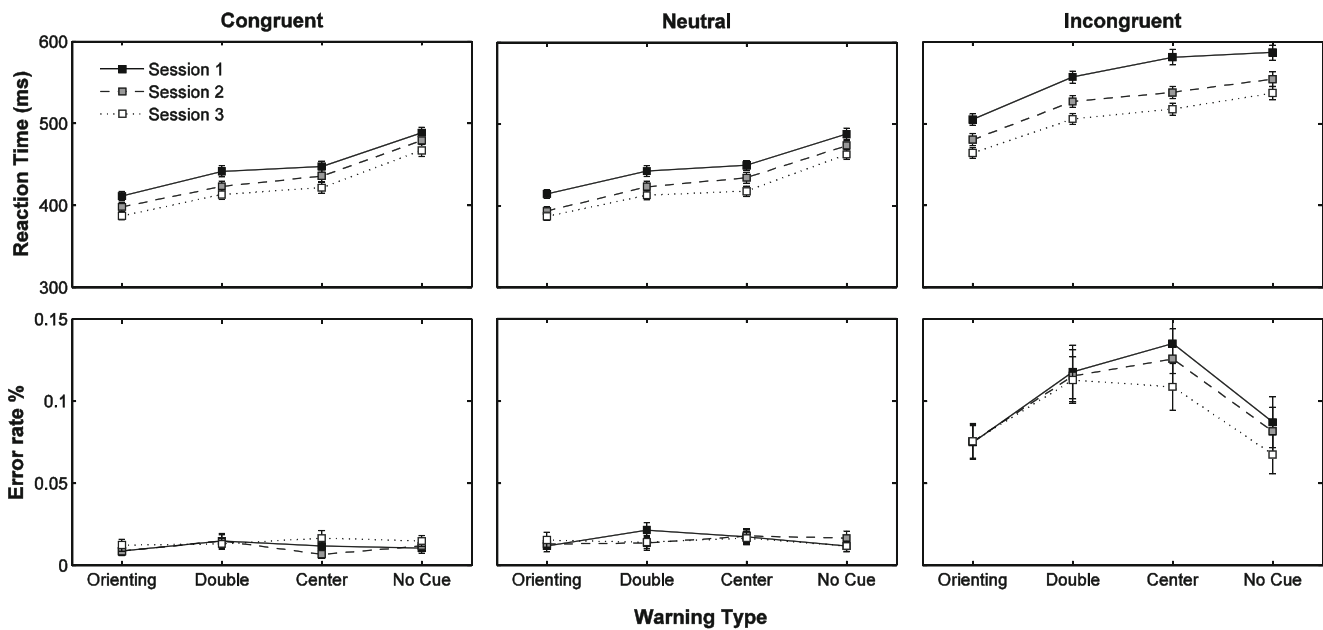
Results

ANT

Reaction times Reaction times were analyzed by conducting a mixed analysis of variance (ANOVA) with Cue Type (orienting, double, central, or no cue), Congruency (congruent, neutral, or incongruent), and Session (1–3) as within-subjects factors, and Test Order (TVA–ANT or ANT–TVA) as a between-subjects factor. Only correctly reported trials and trials with a reaction time within three standard deviations of the mean in each condition were used in the analysis of reaction times. The ANOVA showed no main effect of

test order, $F(1, 66)<0.01$, n.s., but revealed a significant main effect of session, $F(2, 132)=45.70$, $p<0.001$, $\eta_p^2=0.409$, reflecting the overall improvement in reaction times across the three sessions. Moreover, we found significant main effects of cue type, $F(3, 198)=415.88$, $p<0.001$, $\eta_p^2=0.863$, and congruency, $F(2, 132)=556.14$, $p<0.001$, $\eta_p^2=0.894$, indicating (see Fig. 3) that responses were slower to incongruent than to congruent stimuli, and that orienting cues led to the fastest responses, whereas the no-cue condition led to the slowest responses of the four warning types. The ANOVA also revealed a significant Congruency×Cue Type interaction, $F(6, 396)=23.60$, $p<0.001$, $\eta_p^2=0.263$, reflecting that alerting cues (center or double cues) without spatial information resulted in a further increase in reaction times in the incongruent condition compared with the effect of an orienting cue or when no cue was presented in the same condition. Most interestingly, in the context of the present study, the ANOVA showed a significant Session×Congruency interaction, $F(4, 264)=25.10$, $p<0.001$, $\eta_p^2=0.276$, reflecting a larger drop in reaction time across session in the incongruent condition compared with the congruent and neural condition. Finally, we found a significant Session×Cue Type interaction, $F(6, 396)=2.32$, $p=0.049$, $\eta_p^2=0.034$, and a significant Session×Congruency×Cue type interaction, $F(12, 792)=2.40$, $p=0.013$, $\eta_p^2=0.035$, however, both effect sizes were low and no clear interpretable patterns were found in the data.

It has been suggested that effects of fatigue due to prolonged testing may increase reaction times and that general inattentiveness may be seen as an increased variability in reaction times (Dockree et al., 2006; O'Connell et al., 2008). Thus, being tested on a long TVA-based paradigm before running the ANT test could potentially lead to reaction time disadvantages for the TVA–ANT group. However the main effect of test order on the reaction times across sessions was insignificant in the omnibus ANOVA. We also controlled whether the coefficient of variation of the reaction time (the standard deviation of the reaction times divided by the mean reaction time) for each participant in each session was affected by session number and test order. The coefficient of variation of the reaction time was very similar for the two groups over the three sessions (range: 0.21–0.22). This was also confirmed in a mixed ANOVA with Session (1–3) as a within-subjects factor and Order (TVA–ANT, ANT–TVA) as a between-subjects factor, which showed neither significant main effects nor a significant interaction (all ps>0.30, all η_p^2s<0.02).

Error rates Error rates were analyzed by conducting a repeated measures ANOVA with a similar setup to the one used in the analysis of the reaction times. The ANOVA showed no main effects of test order, $F(1, 66)=0.001$, $p=0.982$, $\eta_p^2<0.001$, or of session, $F(2, 132)=1.38$ $p=0.256$, $\eta_p^2=0.020$, but did show significant main effects of cue type,

Fig. 3 Results from the ANT task. Error bars indicate standard errors of the means

$F(3, 198)=15.570$, $p<0.001$, $\eta_p^2=0.191$, and congruency, $F(2, 132)=82.82$, $p<0.001$, $\eta_p^2=0.557$, reflecting that responses to incongruent stimuli were less accurate than congruent and neutral stimuli and that alerting cues (center or double cues) without spatial information resulted in more errors compared with orienting cues and the no-cue condition. However, a significant Congruency×Cue Type interaction, $F(6, 396)=12.68$, $p<0.001$, $\eta_p^2=0.161$, suggested that alerting cues only resulted in more errors in the incongruent condition and not in the congruent and neutral conditions.

Attention network scores Figure 4 shows the attention network scores for each of the three sessions. The scores were analyzed by conducting a mixed ANOVA with Session (1–3) as a within-subjects factor and Test Order (TVA–ANT, ANT–TVA) as a between-subjects factor for each of the three network scores. The ANOVAs showed no significant difference between the alerting network scores in the three sessions, $F(2, 132)=0.28$, $p=0.757$, $\eta^2=0.004$, whereas attenuations of both the orienting network scores, $F(2, 132)=8.00$, $p=0.001$, $\eta^2=0.108$, and executive network scores, $F(2, 132)=37.18$, $p<0.001$, $\eta^2=0.360$, were supported by the ANOVAs. No main effects of test order were found, all Fs$(1, 66)<3.91$, all ps>0.05.

Internal reliability The data files were split into two halves by assigning trials in each condition to one of two groups in an alternating fashion. Split-half correlations were computed for each network score in each of the three testing sessions. All correlations were significantly different from zero at $p<0.001$, apart from the orienting score in Session 1 and the

alerting scores in all sessions, all of which were, however, still significant at $p<0.05$. The correlations were then corrected to full session length by the Spearman–Brown prediction formula[1] (see Table 1). The mean correlation across the three testing sessions for the executive network score was 0.82, and it was markedly lower for both the alerting and orienting network scores (0.48 and 0.61, respectively).

Test–retest reliability To obtain measures of test–retest reliability, we computed pairwise Pearson correlations of performance in different sessions for each ANT parameter (see Table 2). All correlations were highly significant ($p<0.001$), apart from the correlation between the alerting score in the first and second session ($r=0.14$, $p=0.24$) and between the second and third session ($r=0.35$, $p=0.004$). The executive network score had a retest reliability of 0.79 on average, the reliability of the orienting score was considerably lower (0.59 on average), and the reliability of the alerting score was only 0.25 on average. It should however be noted that the reliability of the alerting score improved from the first to the second retesting.

TVA

Basic performance pattern The mean values of the attentional parameters estimated from the performance in

---

[1] The predicted reliability, ρ^*, is estimated as $\rho^*=N\rho/[1+(N-1)\rho]$, where N is the number of tests combined, and ρ is the reliability of the current test (Brown, 1910; Spearman, 1910).

Fig. 4 Network scores based
on results from the ANT task.
Error bars indicate standard
errors of the means
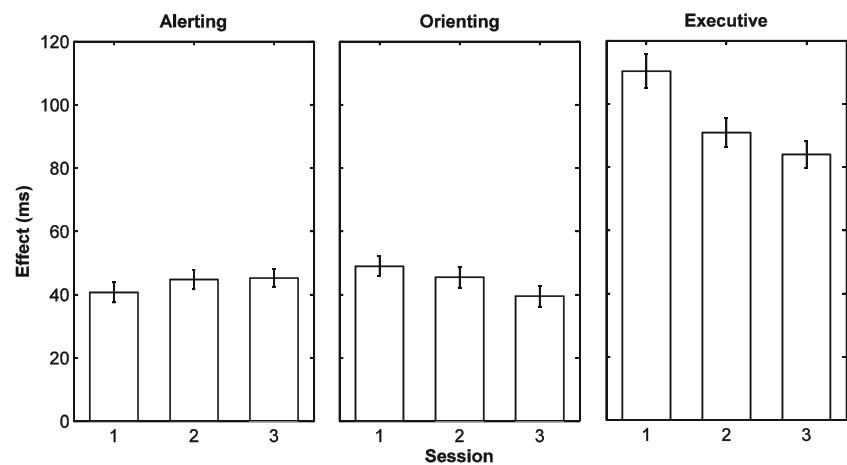
TVA-based assessment are summarized for the three individual test sessions in Table 3. The mean error rate and the goodness of fit (percentage of variance explained) are also indicated.

Mixed ANOVAs with Session (1–3) as a within-subjects factor and Test Order (ANT–TVA, TVA–ANT) as a between-subjects factor revealed a significant increase (main effect of session) over the course of the three test sessions in K, $F(2, 132)=7.92$, $p=0.001$, $\eta_p^2=$

Table 1 Internal reliability for attention network test (ANT) scores and theory of visual attention (TVA) parameters

	Session 1	Session 2	Session 3	Mean
ANT (full version: 288 trials)				
Alerting	.47 (.57)	.56 (.65)	.42 (.52)	.48 (.58)
Orienting	.50 (.60)	.65 (.74)	.67 (.75)	.61 (.70)
Executive	.87 (.91)	.82 (.87)	.77 (.83)	.82 (.87)
TVA (full version: 432 trials)				
K	.96	.96	.94	.95
t_0	.90	.91	.91	.91
C	.86	.88	.90	.88
α	.88	.91	.90	.90
w_{index}	.96	.95	.93	.94
TVA (half version: 216 trials)				
K	.94 (.97)	.93 (.96)	.90 (.95)	.92 (.96)
t_0	.77 (.87)	.82 (.90)	.86 (.92)	.82 (.90)
C	.72 (.84)	.70 (.82)	.62 (.77)	.68 (.81)
α	.58 (.73)	.78 (.87)	.80 (.89)	.72 (.83)
w_{index}	.92 (.96)	.89 (.94)	.92 (.96)	.91 (.95)

For the full version of TVA-based assessment, one participant's data from Session 3 could not be modeled. For the half version of TVA-based assessment, one participant's data from Session 2 and four participants' data from Session 3 could not be modeled. The figures reported in parentheses represent the Spearman–Brown predicted correlations if the test length had been 432 trials.

0.107, and C, $F(2, 132)=58.12$, $p<0.001$, $\eta_p^2=0.468$, and significant reductions of t_0, $F(2, 132)=7.36$, $p=0.002$, $\eta_p^2=0.100$, α, $F(2, 132)=25.72$, $p<0.001$, $\eta_p^2=0.280$, and the error rate, $F(2, 132)=16.22$, $p<0.001$, $\eta_p^2=0.197$, but no change in w_{index}, $F(2, 132)=1.19$, $p=0.304$, $\eta_p^2=0.018$ (cf. Fig. 5). The order in which the tests were run (main effect of test order) did not affect any of the parameters (all ps>0.32, all η_p^2s< 0.02), except for a small but significant effect on t_0, $F(1, 66)= 8.63$, $p=0.005$, $\eta_p^2=0.116$. Post hoc t tests revealed that this was driven by a higher t_0 value in the ANT–TVA group in the first session only (Session 1, $t(66)=-3.89$, $p<0.001$, $d=-0.96$; Session 2, $t(66)=-1.50$, $p=0.477$, $d=-0.37$; Session 3, $t(66)=-2.14$, $p=0.108$, $d=-0.52$, Bonferroni-corrected p values). Accordingly, this resulted in a small Session×Order interaction for t_0, $F(2, 132)=5.21$, $p=0.010$, $\eta_p^2=0.073$. The error rates also showed a negligible interaction effect, $F(2, 132)=4.27$, $p=0.028$, $\eta_p^2=0.061$, but no other parameters were affected differentially by the test order across sessions (all ps>0.22, all η_p^2s<0.03).

Internal reliability Split-half correlations were computed for each parameter and in each of the three testing sessions (see Table 1). All correlations were significantly different from zero at $p<0.001$ (apart from α, which was significant at $p=0.001$). The split-half correlations were corrected to the full session length using the Spearman–Brown prediction formula, and the results showed high internal reliability for all of the five parameters measured by TVA-based assessment. As measured by the mean correlations across the three testing sessions, the most reliable parameters were w_{index} (0.94) and K (0.95), whereas the correlations for t_0, C, and α were in the range 0.88–0.91, also very high.

Test–retest reliability To obtain measures of test–retest reliability, we computed pairwise Pearson correlations of

Table 2 Test–retest correlations for attention network test (ANT) scores and theory of visual attention (TVA) parameters

	Session 1: Session 2	Session 2: Session 3	Mean
ANT (full version: 288 trials)			
Alerting	.14 (.20)	.35 (.45)	.25 (.32)
Orienting	.58 (.67)	.59 (.69)	.59 (.68)
Executive	.84 (.89)	.74 (.81)	.79 (.85)
TVA (full version: 432 trials)			
K	.84	.82	.83
t_0	.58	.77	.68
C	.63	.75	.69
α	.59	.85	.72
w_{index}	.89	.90	.89
TVA (half version: 216 trials)			
K	.76 (.86)	.86 (.92)	.81 (.89)
t_0	.43 (.61)	.67 (.80)	.55 (.70)
C	.60 (.73)	.60 (.75)	.59 (.74)
α	.44 (.61)	.68 (.81)	.56 (.71)
w_{index}	.79 (.88)	.82 (.90)	.80 (.89)

The figures reported in parentheses represent the Spearman–Brown predicted correlations if the test length had been 432 trials.

performance in the different sessions for each TVA parameter (see Table 2). All of the correlations were highly significant ($p<0.001$). Between Sessions 1 and 2, K and w_{index} again had the highest reliability, whereas t_0, C, and α were in the range 0.58–0.63. At the second retesting, the reliability of these three parameters improved considerably (now ranging from 0.75 to 0.85), whereas K and w_{index} maintained their high levels.

Effects of test length To investigate how the higher reliability of the TVA test would hold up when the test was reduced to a shorter version than the 288 trials of

Table 3 Attentional performance in the three sessions of TVA-based assessment

	Session 1		Session 2		Session 3	
Parameter	M	SD	M	SD	M	SD
t_0	19.9	10.6	17.9	8.7	17.3	9.3
C	53.5	18.1	72.8	23.4	83.7	29.4
K	3.19	0.70	3.32	0.68	3.47	0.71
α	.79	.41	.58	.37	.48	.31
w_{index}	.50	.11	.49	.10	.49	.10
Error rate	.23	.08	.19	.06	.18	.07
Prop. var	.94	.04	.94	.03	.94	.03

The units for the individual parameters are t_0 (ms), C (letters/s), and K (letters). α ranges from perfect selection, at 0, to nonselectivity, at 1.0. w_{index} ranges from extreme leftward (1.0) to extreme rightward (0) bias of attentional weighting, with .5 representing lateral symmetry. "Prop. var" describes the proportion of the variance explained by TVA modeling.
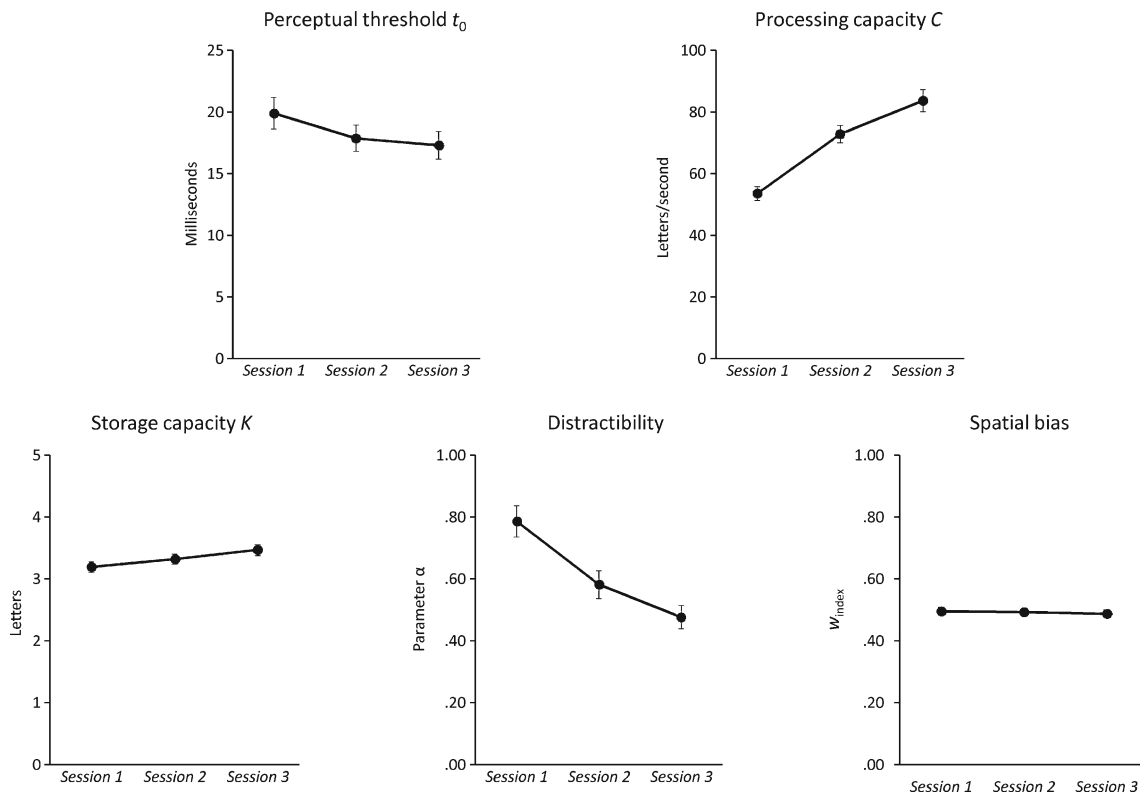
the ANT, we computed internal reliability correlations for data from only the first half of each test session (i.e., 216 trials; see Table 1). Even with fewer trials the internal reliability of TVA-based assessment remained superior to that of the ANT test. Only the reliability of the executive network score was at a comparable level—better than the least reliable two TVA parameters, C and α, but inferior to K and w_{index}. For retest reliability, we made the same comparison for data from only the first half of each TVA test session (see Table 2). In this case, the reliability of the executive network score was in the same range (about 0.8) as the two most reliable TVA parameters, K and w_{index}. However, the retest reliability of the orienting network score was now comparable to those of the three least reliable TVA parameters—t_0, C, and α. The reliability of the alerting network score remained much poorer than that of all TVA parameters.

Another way of controlling for the different number of trials in the two tests would be to use the Spearman–Brown formula to estimate what the reliability correlations would have been if each test had contained 432 trials. These estimations are reported for the ANT in parentheses in Tables 1 and 2. To make a direct comparison for this analysis, we also used the Spearman–Brown formula to correct the half version of TVA to full length (also reported in parentheses). The pattern of the results is the same one found above: For the internal reliability, only the executive network score was comparable to TVA-based assessment, but for retest reliability, the orienting network score also reached approximately the same level as the least reliable TVA parameters.

Fig. 5 Results from TVA-based assessment. Error bars indicate standard errors of the means

Pairwise correlations between parameters

Pearson correlations were computed in order to examine the relation between each pair of parameters from the two tests (see Table 4). Significant interparameter correlations were found between C and K ($r=0.72$, $p<0.001$). All of the other correlations were small and nonsignificant when Bonferroni corrected for multiple comparisons. Notably, we found no significant correlation between α and the executive network score, even before Bonferroni correction ($r=0.24$, $p=0.053$).

Discussion

We report on the first direct comparison between the ANT and TVA-based assessment, two widely used tests of attention. Based on a study of 68 young healthy participants, we aimed to compare the two tests with regards to reliability, practice effects, and correlations between the individual attention parameters. The internal reliability (as measured by split-half analysis) was high for all parameters of TVA-based assessment. In the ANT only the reliability of the executive network score approached these levels, whereas

Table 4 Correlations between attentional parameters

ANT Parameters	TVA Parameters						
	Orienting	Conflict	K	C	t_0	α	w_{index}
Alerting	0.15	0.00	−0.28	−0.14	0.21	0.04	−0.08
Orienting		−0.07	−0.07	−0.18	−0.14	−0.01	−0.05
Executive			0.04	−0.05	0.02	0.24	−0.01
K				**0.72**[**]	−0.11	−0.21	−0.01
C					0.14	−0.18	−0.17
t_0						0.25	−0.07
α							+0.08

[**]$p<0.001$ (bold; Bonferroni corrected)

the reliability of the alerting and orienting scores were much lower. A similar pattern was found for test–retest reliability, measured over three test sessions, although the differences between the ANT and TVA-based assessment were not as pronounced for this measure. Also, after correcting for the difference in test length, the retest reliability of the orienting network score equaled that of the three least reliable TVA parameters. Our study also revealed significant practice effects in the two tests. In TVA-based assessment, a strong practice effect was found for the C parameter (visual processing speed), and the K (storage capacity), t_0 (perception threshold), and α (distractibility) parameters also improved significantly over the course of the three testing sessions. In contrast, the w_{index} (spatial attentional bias) remained stable throughout testing. In the ANT, the executive network score was strongly improved by retesting, a slight but significant practice effect was observed for orienting, and the alerting score did not change significantly by retesting. Finally, our results showed that the parameters measured by each test are largely independent, as measured by low and nonsignificant correlations between most test scores. There was one exception to this general pattern. Within the set of attention parameters measured by TVA assessment, the C and K parameters covaried strongly. The order in which the two tests were given did not influence these main results, apart from an effect on one parameter (t_0) in one test session.

Our results confirm previous findings that in the standard version of the ANT, only the executive network score has good reliability, in terms of both internal (split-half) and retest reliability. In contrast, all parameters of TVA-based assessment showed high internal reliability. The retest reliabilities were also in the acceptable range for research purposes, especially the w_{index} and K parameters, which is reassuring, since no previous study had charted the retest properties of TVA-based assessment. The finding shows that TVA-based testing is well-suited for studies of cognitive training and other investigations that involve repeated testing of the same participants. One possible explanation for the higher reliability of TVA-based assessment is that the standard version of the test that we used is 50 % longer than the ANT (432 vs. 288 trials). When we looked at internal reliability for data from only half of the TVA session (i.e., 216 trials), the estimates of reliability for TVA-based assessment were still superior to those from the ANT, although in case of retest reliability, the orienting network score was now at a comparable level. The same pattern emerged from an estimation based on the Spearman–Brown prediction formula. The lower reliability of the ANT may be explained by the subtractions used for calculating the network scores. As was pointed out by Macleod et al. (2010), difference scores can be associated with low reliability as a result of the inverse relation between difference score reliability and the correlation between the two variables used in the creation of

the difference score (Salthouse & Hedden, 2002). Another explanation relates to the dependent variables used by the two tests: accuracy versus reaction time. Reaction time is typically associated with significant intraindividual variation (Salthouse & Berish, 2005). Without the variability introduced by response related processes, it is perhaps not surprising that the more functionally specific TVA parameters can be measured more reliably. In previous investigations (e.g., Finke et al., 2005; Habekost & Rostrup, 2006) one notable exception to the high internal reliability of TVA measures was the α parameter. In the present study, we found that this parameter also had good internal reliability, though α (and the C parameter) was more sensitive to a reduction in the number of trials than the other TVA parameters. The improved reliability of α may be related to our experimental design and data analysis. Unlike in previous studies, we used a combined whole and partial report design, where the two tasks were interleaved during testing and the data from both conditions were analyzed together. The additional information provided by the whole report data should constrain the fitting of the α parameter more strongly, leading to more reliable estimates.

We also found significant practice effects in both tests, involving parameters C, K, t_0, and α in TVA-based assessment as well as the executive and orienting network scores of the ANT. The strongest effects were found for C and the executive network score. The increase in visual processing speed over the course of the three sessions probably reflects improvements in visual pattern recognition, that is, strengthening of the particular visual representations (letter forms) required to perceive the brief displays of the experiment. A similar argument has been made for the improvement in the K parameter that is observed with letter experience (Sørensen & Kyllingsbæk, 2012), and could be made for the practice effect on the t_0 parameter as well. The practice effect on the executive network score replicates a previous study (Ishigami & Klein, 2010), which also found that the effect was based on decreased reaction time in the incongruent trials (cf. Fig. 3). That is, participants became better at ignoring the distracting flankers with practice. Presumably a similar explanation can be applied to the gradual improvement of the α parameter. The small practice effect observed on the orienting network score went against the predicted direction and cannot readily be explained. The practice effects observed in this study are important to keep in mind for future studies and also for the interpretation of the many existing studies that have used either test. The findings show that it is essential to control for the level of test experience between participants. However, the good reliability of TVA-based assessment implies that this test can still be used to compare (equally) practiced participants: although several TVA parameters will improve with practice, the relative ordering of individuals should remain fairly

constant. The same applies to the executive network score of the ANT.

Within the set of measures provided by TVA-based assessment, the correlation between C and K was strong ($r=0.72$). This is markedly higher than in a previous study by Finke et al. (2005; $r=0.40$), but in line with recent investigations by Vangkilde et al. (2011) and Jensen et al. (2012). Using an identical test paradigm to the present study, these two studies found correlations in the range of 0.57 to 0.77. The discrepancy to Finke et al.'s (2005) findings may be explained by recent improvements in the mathematical fitting procedure for TVA analysis. Dyrholm et al. (2011) presented a new fitting procedure for TVA-based assessment, which eliminates some of the estimation biases associated with the previous analysis procedure (Duncan et al., 1999; Kyllingsbæk, 2006), especially regarding estimation of K. This new procedure was employed in the present study as well as those of Vangkilde et al. (2011) and Jensen et al. (2012), and the results suggest that previous studies may have underestimated the correlation between C and K due to more noisy parameter estimation. A strong correlation between individual levels of C and K may be surprising in view of the fact that the two parameters represent conceptually and mathematically distinct aspects of attention (Shibuya & Bundesen, 1988). As was suggested by Habekost and Starrfelt (2009), their correlation may be explained by a shared neural basis. TVA-based studies of lesions in different parts of the brain indicate that both visual processing speed and the capacity of visual short-term memory depend on largely overlapping networks of brain areas that include extrastriate, thalamic, basal ganglia, as well as parietal and frontal regions. It seems that individuals vary more in the general efficiency of this network than specifically in one or the other visual capacity parameter.

The low correlations between all of the other test parameters support our prediction that TVA-based assessment and the ANT are directed at separate aspects of attention, reflecting the differences in theoretical focus of their parent models. On the one hand, a range of attentional measures that are highly controlled for the influence of response and motor processes are provided by TVA-based assessment: visual processing speed and storage capacity (C and K), perception thresholds (t_0), visual distractibility (α), and attentional weighting (w_{index}). On the other hand, response-related measures of spatial orienting, executive response control, and general alerting can be obtained from the ANT. For both research and clinical purposes it might often be useful to combine the two tests. This can provide a comprehensive investigation across many attentional functions for a given individual or special population. TVA-based assessment can provide measures of attentional capacity (C, K, and t_0) and attentional weighting (w_{index} and α) within the visual domain. TVA's three capacity parameters

are especially relevant for real-life situations with high visual processing load (e.g., reading, traffic perception, or sports) where perceptual processing limitations can significantly impede performance (see Caspersen & Habekost, 2013, for a discussion). TVA's two weighting parameters relate to other aspects of visual function: The ability to remain focused on particular types of visual input under distraction (α), and the distribution of visual attention across the visual field (w_{index}, which is for example altered in patients with the neglect syndrome; Duncan et al., 1999). These five measures of processing efficiency in the visual system complements the ANT's assessment of the three general attention networks, whose wider functional relevance is probably more well-known than that of the TVA parameters (see, e.g., Parasuraman, 2000). From a practical viewpoint both tests can be performed within approximately 1 h, at least with healthy participants. When testing clinical populations (e.g., individuals with brain damage) who are vulnerable to test fatigue, one might consider reducing the number of trials in TVA-based assessment. As our study shows, the test retains good reliability even with half the standard number of trials, though the estimation of some TVA parameters are more affected than others by a reduction in trials.

In summary, TVA-based assessment and the ANT measure distinct aspects of attention. The two tests can therefore be considered as complementary tools for characterizing the attentional abilities of a given individual or special population. When interpreting the scores of the tests one should keep in mind their psychometric properties, in particular the low reliability of the alerting network scores of the ANT as well as the substantial practice effects that are associated with specific parameters of each test.

References

Adólfsdottir, S., Sørensen, L., & Lundervold, A. J. (2008). The attention network test: A characteristic pattern of deficits in children with ADHD. *Behavioral and Brain Functions, 4,* 9. doi:10.1186/1744-9081-4-9

Bednarek, D. B., Saldaña, D., Quintero-Gallego, E., García, I., Grabowska, A., & Gómez, C. M. (2004). Attentional deficit in dyslexia: A general or specific impairment? *NeuroReport, 15,* 1787–1790.

Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 3,* 296–322. doi:10.1111/j.2044-8295.1910.tb00207.x

Bublak, P., Redel, P., Sorg, C., Kurz, A., Förstl, H., Müller, H. J., et al. (2011). Staged decline of visual processing capacity in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging, 32,* 1219–1230. doi:10.1016/j.neurobiolaging.2009.07.012

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547. doi:10.1037/0033-295X.97.4.523

Bundesen, C. (1999). A computational theory of visual attention. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action* (pp. 54–72). Oxford, UK: Oxford University Press.

Bundesen, C., & Habekost, T. (2008). *Principles of visual attention: Linking mind and brain*. Oxford, UK: Oxford University Press.

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112,* 291–328. doi:10.1037/0033-295X.112.2.291

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulated cortex. *Trends in Cognitive Sciences, 4,* 215–222.

Caspersen, I. C., & Habekost, T. (2013). Selective and sustained attention in children with spina bifida myelomeningocele. *Child Neuropsychology, 19,* 55–77.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.

Coull, J. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology, 55,* 343–361.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222. doi:10.1146/annurev.ne.18.030195.001205

Dockree, P. M., Bellgrove, M. A., O'Keeffe, F. M., Moloney, P., Aimola, L., Carton, S., et al. (2006). Sustained attention in traumatic brain injury (TBI) and health controls: Enhanced sensitivity with dual-task load. *Experimental Brain Research, 168,* 218–229.

Duncan, J. (2006). EPS Mid-Career Award 2004: Brain mechanisms of attention. *Quarterly Journal of Experimental Psychology, 59,* 2–27. doi:10.1080/17470210500260674

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General, 128,* 450–478. doi:10.1037/0096-3445.128.4.450

Dyrholm, M., Kyllingsbæk, S., Espeseth, T., & Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: The case of TVA. *Journal of Mathematical Psychology, 55,* 416–429. doi:10.1016/j.jmp.2011.08.005

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143–149. doi:10.3758/BF03203267

Fan, J., Fossella, J., Sommer, T., Wu, Y., & Posner, M. I. (2003). Mapping the genetic variation of executive attention onto brain activity. *Proceedings of the National Academy of Sciences, 100,* 7406.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26,* 471–479.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14,* 340–347. doi:10.1162/089892902317361886

Fan, J., Wu, Y., Fossella, J., & Posner, M. I. (2001). Assessing the heritability of attentional networks. *BMC Neuroscience, 2,* 14.

Finke, K., Bublak, P., Krummenacher, J., Kyllingsbæk, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: Evidence from normal subjects. *Journal of the International Neuropsychological Society, 11,* 832–842.

Finke, K., Dodds, C. M., Bublak, P., Regenthal, R., Baumann, F., Manly, T., et al. (2010). Effects of modafinil and methylphenidate on visual attention capacity: A TVA-based study. *Psychopharmacology, 210,* 317–329.

Finke, K., Schwarzkopf, W., Müller, U., Frodl, T., Müller, H. J., Schneider, W. X., et al. (2011). Disentangling the adult attention-deficit hyperactivity disorder endophenotype: Parametric measurement of attention. *Journal of Abnormal Psychology, 120,* 890–901. doi:10.1037/a0024944

Fossella, J. A., Sommer, T., Fan, J., Wu, Y., Swanson, J. M., Pfaff, D. W., et al. (2002). Assessing the molecular genetics of attention networks. *BMC Neuroscience, 3,* 14.

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill.

Habekost, T., & Bundesen, C. (2003). Patient assessment based on a theory of visual attention (TVA): Subtle deficits after a right frontal–subcortical lesion. *Neuropsychologia, 41,* 1171–1188.

Habekost, T., & Rostrup, E. (2006). Persisting asymmetries of vision after right side lesions. *Neuropsychologia, 44,* 876–895.

Habekost, T., & Starrfelt, R. (2009). Visual attention capacity: A review of TVA-based patient studies. *Scandinavian Journal of Psychology, 50,* 23–32.

Ishigami, Y., & Klein, R. M. (2010). Repeated measurement of the components of attention using two versions of the Attention Network Test (ANT): Stability, isolability, robustness, and reliability. *Journal of Neuroscience Methods, 190,* 117–128.

Jensen, C. G., Vangkilde, S., Frøkjær, V., & Hasselbalch, S. G. (2012). Mindfulness training affects attention—or is it attentional effort? *Journal of Experimental Psychology: General, 141,* 106–123.

Kyllingsbæk, S. (2006). Modeling visual attention. *Behavior Research Methods, 38,* 123–133. doi:10.3758/BF03192757

Lehtonen, S. E. (2008). Self-reported inattention and hyperactivity–impulsivitiy as predictors of attention network efficiency. *Dissertation Abstracts International: Section B, 68,* 8402.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108,* 393–434. doi:10.1037/0033-295X.108.2.393

Macleod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology, 24,* 637–651. doi:10.1037/a0019803

O'Connell, R. G. O., Bellgrove, M. A., Dockree, P. M., Lau, A., Fitzgerald, M., & Robertson, I. H. (2008). Self-alert training: Volitional modulation of autonomic arousal improves sustained attention. *Neuropsychologia, 46,* 1379–1390.

Parasuraman, R. (2000). The attentive brain: Issues and prospects. In R. Parasuraman (Ed.), *The attentive brain* (pp. 3–16). Cambridge, MA: MIT Press.

Peers, P. V., Ludwig, C. J., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., et al. (2005). Attentional functions of parietal and frontal cortex. *Cerebral Cortex, 15,* 1469–1484. doi:10.1093/cercor/bhi029

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58,* 1–23. doi:10.1146/annurev.psych.58.110405.085516

Salthouse, T. A., & Berish, D. E. (2005). Correlates of within-person (across-occasion) variability in reaction time. *Neuropsychology, 19,* 77–87.

Salthouse, T. A., & Hedden, T. (2002). Interpreting reaction time measures in between-group comparisons. *Journal of Clinical and Experimental Neuropsychology, 24,* 858–872.

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 591–600.

Sørensen, T. A., & Kyllingsbæk, S. (2012). Short-term storage capacity for visual objects depends on expertise. *Acta Psychologica, 140,* 158–163. doi:10.1016/j.actpsy.2012.04.004

Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology, 3,* 271–295. doi:10.1111/j.2044-8295.1910.tb00206.x

Tang, Y.-Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., et al. (2007). Short-term meditation training improves attention and self-regulation. *Proceedings of the National Academy of Sciences, 104,* 17152–17156. doi:10.1073/pnas.0707678104

Vangkilde, S., Bundesen, C., & Coull, J. T. (2011). Prompt but inefficient: Nicotine differentially modulates discrete components of attention. *Psychopharmacology, 218,* 667–680. doi:10.1007/s00213-011-2361-x

Vangkilde, S., Coull, J. T., & Bundesen, C. (2012). Great expectations: Temporal expectation modulates perceptual processing speed. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1183–1191. doi:10.1037/a0026343

van Zomeren, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. Oxford, UK: Oxford University Press.

Wang, K., Fan, J., Dong, Y., Wang, C., Lee, T. M. C., & Posner, M. I. (2005). Selective impairment of attentional networks of orienting and executive control in schizophrenia. *Schizophrenia Research, 78,* 235–241.