EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
2006, 18 (1), 90–108

Psychology Press
Taylor & Francis Group

Individual differences in the ability to avoid distracting sounds

Emily M. Elliott and Katie M. Barrilleaux
Louisiana State University, Baton Rouge, USA

Nelson Cowan
University of Missouri-Columbia, USA

The present work aims to establish a greater understanding of the cognitive mechanisms involved in avoiding distraction from speech and nonspeech sounds. Although mixed results have been presented by research investigating the hypothesis that individuals with superior working memory abilities are better able to avoid acoustic distraction, we found that working memory correlated with some aspects of performance during distraction when carefully examined. This is consistent with the view that working memory involves resisting interference. In a large sample, we examined two different tasks accompanied by acoustic distraction—serial recall and rapid colour naming—as well as two different measures of working memory (operation span and running span). We show that the previous inability to find relations between working memory and avoidance of distraction may stem from the use of inadequate correlational techniques. Additionally, the level of difficulty of the serial recall task may be an important factor. The results illustrate that commonly used statistical techniques can be misleading and furthermore that the ability to avoid distraction from irrelevant items may not be a unitary construct.

The question of how much cognitive control we can exert in the face of distraction from stimuli in the environment is one that has generated a considerable amount of interest. Early research in this topic examined the capabilities of the selective attention system by examining the performance of participants in dichotic-listening tasks, often with the requirement that one channel of input be shadowed to control the direction of attention (Treisman, 1960). This type of research helped to define the limits of the cognitive system. More recently, researchers have examined the role of attention in cognition by asking

Correspondence should be addressed to Emily Elliott, Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803, USA. Email: eelliott@lsu.edu

This work was supported by a grant from the Louisiana Board of Regents, awarded to the first author, and NIH Grant R01 HD-21338 awarded to the third author.

http://www.psypress.com/ecp DOI:10.1080/09541440500216044

participants to perform complex memory span tasks in which some elements of both storage and processing are required (Turner & Engle, 1989). Current theoretical views describe attention as an important component of working memory function (Cowan, 1995; Engle, Kane, & Tuholski, 1999). This study is an attempt to integrate findings from different tasks with the goal of understanding the ability to avoid distraction from irrelevant sounds. Two measures of working memory will be utilised, along with two different measures of performance in the presence of irrelevant sounds. Past work with complex span measures and distraction-based tasks will be discussed, and then the design of the current study will be described.

Uniting the literature from the dichotic-listening and complex span paradigms, Conway, Cowan, and Bunting (2001) examined the relationship between individual differences in working memory performance and the ability to detect one's name in the unattended channel (the task of Moray, 1959; Wood & Cowan, 1995). Working memory span was measured using the operation span task (Turner & Engle, 1989), in which participants were asked to complete arithmetic problems between items presented for later serial recall. (This task will be discussed in more detail in the Method section.) Their findings revealed that those participants with operation spans in the lowest quartile were much *more* likely to report hearing their name in the unattended channel than those participants in the highest quartile (65% vs. 20%). The results clearly indicated that the ability to avoid distraction from irrelevant auditory material (in this case, the participant's own name) was related to working memory, a finding that is consistent with other evidence that high span individuals better control attention (e.g., Kane, Bleckley, Conway, & Engle, 2001).

However, recent work with a different type of task using irrelevant auditory materials has been somewhat of a puzzle, as the expected pattern of results has not been obtained. Beaman (2004) examined the role of working memory in individual differences in the ability to avoid distraction from irrelevant sounds while engaging in a serial recall task. This work, known as the irrelevant-speech effect (Colle & Welsh, 1976; Salamé & Baddeley, 1982), refers to the finding that performance on a serial recall task decreases in the presence of irrelevant speech that participants are instructed to ignore. Because it can occur also with nonspeech distracting sounds (Jones & Macken, 1993), it has more recently been termed the irrelevant-sound effect (ISE). Beaman also used the operation span task as a measure of working memory, and he did not find evidence for a relationship between the magnitude of the ISE and working memory span. Two studies found no relationship even between serial recall and the ISE (Ellermeier & Zimmer, 1997; Neath, Farley, & Surprenant, 2003).

This issue was examined once more by Elliott and Cowan (2005) in their study of individual and developmental differences in the ISE using both speech and tone stimuli. Although a few correlations were found between measures of memory span and the ISE, the correlations were in the positive direction (larger irrelevant-sound effects for higher span individuals), which was unexpected. The findings were interpreted in light of the role of rehearsal in the performance of both memory span tasks and the serial recall task used to measure the ISE. Prominent theories of the ISE discuss the importance of rehearsal of the relevant items for disruption from the irrelevant items to occur (e.g., Beaman & Jones, 1997, 1998). In those theories, it is assumed that the irrelevant sounds, even if they are nonspeech sounds, interfere with the rehearsal process and the episodic record that it creates in working memory. Those participants who are likely to be successful in a complex memory span task are likely to make heavy use of rehearsal of the to-be-remembered items. These same participants, as a consequence, by these assumptions would be likely to show large detrimental effects of irrelevant speech on their serial recall performance. Thus, a positive correlation could result if participants with a higher score on a working-memory measure, by engaging in more rehearsal, created more opportunities for corruption of the episodic record by irrelevant sounds.

One reason for the absence of a more consistent pattern of correlations is that they depend on individual ISE scores obtained by subtracting a speech condition from a no-speech control condition. In order for two measures to correlate, both must be sufficiently reliable. Elliott and Cowan (2005) showed that, even when the raw measures are quite reliable, the difference scores tend to be much less reliable. The highest reliability alpha coefficient obtained to date for an ISE may be that of Ellermeier and Zimmer (1997), and that coefficient was only .55. To illustrate the effect of this unreliability, according to the formula for correction of the attenuation of correlations caused by unreliability of the measures, $r_{corrected} = r_{xy} / (r_{xx} * r_{yy})$, if the true population correlation between the ISE and a working memory measure were .40 and the reliability of the two measures to be correlated were .55 and .90, respectively, then the observed correlation would be only .28. Elliott and Cowan (2005) solved this unreliability problem by using stepwise regressions on the raw measures, rather than calculating a subtraction for the ISE. If the irrelevant-sound condition accounts for significant variance in the working memory measure even after the variance from the silent control condition has been taken out, then it can be said that the irrelevant-sound effect is related to working memory. Elliott and Cowan still found inconsistent results with this regression method, but the question will be reexamined in the present paper using a different set of ISE conditions.

Another area of the literature that has shown relationships between working memory and performance in the presence of distracting stimuli is the traditional Stroop task, in which participants are asked to name the ink colour of an incongruent colour word (Stroop, 1935). Kane and Engle (2003) used the operation span task and a version of the Stroop task that also included some congruent trials (e.g., the word *red* written in red ink). When the proportion of congruent trials included in the experiment was high, participants with a low score on the working memory task were more likely to show interference effects from the incongruent stimuli (as revealed by either errors or changes in their response times to name the ink colour, depending on the experiment in the series). This was interpreted with respect to the goal maintenance necessary in the face of distraction. With a greater number of congruent trials, the goal of naming the ink colour became more difficult because the same result could be achieved by reading the word instead.

One limitation of the procedure of Kane and Engle (2003) is that it reflects a situation in which it is exceedingly difficult to ignore the distracting stimuli (printed colour words often matching the colours). Under such circumstances, participants may instead have to process them fully and actively suppress them, or prevent them from affecting the response. What is not yet clear is whether working memory performance is related to the effects of distracting colour words when the identity of the colour word is not so often tied to the identity of the printed colour. Also, it is not yet clear whether working memory performance is related to performance on a version of Stroop-type test in which the distracting words are presented in the auditory modality (Cowan & Barron, 1987; Elliott & Cowan, 2001; Elliott, Cowan, & Valle-Inclan, 1998), providing a closer analogue to the ISE.

Finally, in addition to the operation span task (Turner & Engle, 1989) described above, which is a measure of working memory commonly used in recent literature, we also used an older task termed running memory span (Pollack, Johnson, & Knaff, 1959). In this task, participants are asked to attend to a string of items without knowing when the items will stop (i.e., with a variable and unpredictable number of stimuli), and are then asked to remember a set amount of items from the end of the list (e.g., the last six or eight items). This type of task, like operation span, correlates with mental aptitudes more highly than does the typical digit span test, and correlates well with other working memory tasks (Cowan, Elliott, & Saults, 2002; Mukunda & Hall, 1992).

The running span task, as we administered it, had several attributes that made it very different from the operation span task. First, whereas operation span was administered visually, running span was administered with a list of spoken digits to be recalled, at a rapid rate of four digits per second. The acoustic presentation is important because, if correlations between memory span and acoustic distraction are driven by individual differences in individuals' tendency to attend to visual versus auditory stimuli, operation span and running span should not yield similar results. The rapid presentation also is important because it allows a clean interpretation of processes involved in running span. In a particularly elegant investigation of running span, Hockey (1973) varied the rate of presentation and also varied the processing instructions. With slow presentations, participants did best when instructed to try to rehearse; they presumably used a method in which the series to be remembered was continually updated in working memory. However, with fast presentations, participants did better when instructed simply to listen passively until the list ended. At that point, they presumably had to focus attention on an auditory sensory memory stream in order to extract the last few digits into working memory (see Cowan et al., 2002). Thus, unlike the operation span task, in our running span task attention is used at a specific, well-defined point in time. Correlations between operation span and running span, despite these substantial differences in procedure, presumably occur because there is a core ability common to them, such as the scope or agility of attention (Cowan, 2001).

In sum, the present study represents an attempt to integrate the findings discussed above in an investigation of individual differences in the ability to avoid distraction from irrelevant sounds. It is perhaps the first study to use multiple measures of auditory distraction and multiple measures of working memory to examine individual differences. Several different irrelevant-sound tasks were used: three conditions in a serial recall task (silence, irrelevant speech, and irrelevant tones) and several in a speeded colour-naming task (silence, noncolour words, and colour words, with the words at two timing intervals). Working memory was measured in two tasks (operation span and running memory span). We will attempt to answer four specific questions: (1) whether the ISE, measured using both speech and tones as distracting auditory stimuli, is related to working memory; (2) whether the cross-modal Stroop effect is related to working memory; (3) whether the two irrelevant-sound tasks are related to each other, and related to working memory in the same way; and (4) whether the two measures of working memory tap into the same abilities and therefore yield the same patterns of correlations. To foreshadow our results a bit, it appears that the ability to avoid distraction from irrelevant sounds may not be a unitary construct, as different elements of working memory abilities may be called upon as task demands change. Findings from this research can help to inform our understanding of the factors that contribute to individual differences in performance in the presence of irrelevant sounds.

METHOD

Participants

One hundred and twenty undergraduate students from Louisiana State University participated in exchange for extra credit in Psychology courses. Of these, two participants reported hearing loss, and four were not native speakers of English. Additionally, four participants were lost due to computer error (two in the operation span task and two in the running memory span task). With those exclusions, 110 were eligible to be included in the data analyses. From that sample, initial screening of the data was done based on the level of arithmetic accuracy in the operation span task. Participants had to achieve at least 80% accuracy on the arithmetic problems to be included in subsequent analyses. This criterion resulted in the removal of nine more participants. These deletions left a total sample of 101 participants (19 male) with data in all tasks. Although this is skewed towards more females in the sample, past research has not shown any gender differences in the ISE (Ellermeier & Zimmer, 1997).

Design, apparatus, and stimuli

A total of six tasks were administered in individual sessions lasting approximately 1 hour and 45 min. Two of the tasks were measures of test anxiety, which will not be discussed here. There were two task orders, with the goal of balancing the order of the two irrelevant-sound tasks. The running memory span task was always first, followed by the operation span task second. In the first task order, the irrelevant-sounds serial recall task was third, whereas, in the second task order, the cross-modal colour naming task was third. The fourth and fifth tasks were the test anxiety measures, and the sixth and final task was the other irrelevant-sounds task, depending on whether a participant was randomly assigned to the first or second task order.

Participants were run on Dell Dimension desktop computers equipped with 17-inch monitors and all programs were presented with E-Prime programming software (Schneider, Eschman, & Zuccolotto, 2002). Each task will be described separately.

Running memory span task. Participants were presented with a fixation cross on the centre of the screen to orient them to the beginning of each trial. Following this, a random list of the digits 1–9 was presented through headphones at the rate of four per second. The spoken digits were recorded in a male voice and were digitally compressed. Although the rate was very rapid, the digits were easy to comprehend. The sound level was measured to be 64 dB(A) with an earphone coupler and a Quest sound-level meter, using the fast setting. The list length was randomly varied to avoid participants' prediction of the end of the list, and five list lengths were presented (12, 14, 16, 18, and 20 digits per list). Participants were instructed to listen carefully to the digits and then try to recall just the last six digits. They had use of a placeholder when typing their response. Strict serial order was emphasised and participants were given practice trials to be sure they understood the instructions. The experimenter made sure the participants understood that they were to recall only the last six digits before continuing on to the test trials. There were three practice trials followed by 40 test trials, including eight repetitions of each of the five list lengths. Each block of five trials included all five list lengths in a random order.

Operation span task. This task was taken from the Conway and Engle (1996) version of the operation span (Ospan) task that is posted on the website for the System of Teaching Experimental Psychology (http://step.psy.cmu.edu). This version of the program presented a two-step arithmetic problem of the type [(2/2) + 2 = 3?], which participants were to assess as true or false. Immediately after their response to each arithmetic problem, a word was presented. Participants were instructed to remember the words for later recall, and to make a concentrated effort to solve the arithmetic problems correctly. All material in this task was read aloud by the participants, so that the experimenter could monitor their compliance with instructions. After receiving the items for a given trial, recall of the words was cued and participants typed in their recall response. There were three sets of items randomly presented, ranging from two to six trials in length, for a total of 60 recall responses across the 15 trials.

Irrelevant-sounds serial recall task. Participants were asked to recall eight-item lists of the consonants *f*, *k*, *l*, *m*, *q*, *r*, *s*, *t*, while ignoring any sounds heard through the headphones. Each trial began with a fixation cross for 750 ms and was followed by presentation of the to-be-remembered (TBR) items at the rate of 1 per second. This was followed by an 8 s retention interval before the cue to recall was given. Participants were asked to type in their response, and strict serial recall was emphasised. The keys on the keyboard that were not part of the current response were disabled, so that participants could only type in letters that were part of the presented list. Sounds were presented either simultaneously with the TBR items, or afterwards in the retention interval.

There were three types of auditory distractor conditions, silence as a baseline measure of serial recall performance, tones, and speech. The tones were a random selection of 500 ms sine waves of 87, 174, 266, 348, 529, 696, 788, 880, and 972 Hz, and were the same tones used in Elliott and Cowan (2005). The speech tokens were one-syllable adjectives that did not begin with the same consonants as the TBR items: *bad*, *cold*, *due*, *hot*, *nice*, *plus*, *vast*, *warm*. The speech sounds were digitised in a female voice and were 310–500 ms in duration. The rate of presentation was equivalent to the visual stimuli at 1 per second, and when presented together the auditory and visual stimuli had simultaneous onsets. The sounds were measured with a Quest sound-level meter and earphone coupler, using the fast setting, and were determined to be within the range of 66–72 dB(A). The sounds also were judged to be subjectively equal in intensity.

Five trial types were included (silence, speech during or after the list, and tones during or after the list) and the task contained 65 trials. The first five trials were practice, and included one of each type of trial. The test trials were broken up into two blocks of 30 trials each, with the opportunity for a break between blocks. The total number of test trials consisted of 12 repetitions of each possible trial type (6 in the first block and 6 in the second).

Cross-modal Stroop task. This colour-naming task was modelled after the task used by Elliott et al. (1998), with the main change being the inclusion of congruent trials. Participants were first presented with a 500 ms fixation cross and were asked to name the colour square that appeared on the screen as quickly as possible. Responses were spoken into a headset microphone connected to a response box that recorded the vocal reaction time. The experimenter keyed in the spoken response, and also keyed in special notation if the participant made a false start (such as saying ‘‘gree. . .blue’’ or if the experimenter made an error. The colours *red*, *green*, *blue*, and *yellow* were used to make up the 4.4 cm $\times$ 4.4 cm colour squares. There were three auditory conditions, silence as a baseline measure of reaction time to name the colour squares, colour words, and non-colour words. These were digitised words in a female voice, 210–500 ms long. However, on different trials of the present experiment, there could be either a match or a mismatch between the colour square and the spoken colour word. The colour words came from the same set as the visually presented colour squares, and the noncolour words came from the set of *tall*, *long*, *short*, and *big*. These words were used in Elliott and Cowan (2001) and were equated for frequency with the colour words.

In addition to the manipulation of type of auditory distractor, there was a manipulation of the presentation timing. The onset of the auditory distractor was either simultaneous with the presentation of the colour square (0 ms SOA) or the onset of the auditory distractor preceded the presentation of the colour square by 500 ms (500 ms SOA). The manipulation of SOA and the type of auditory distractor created six unique trial types. Participants began with 12 practice trials in which colour squares were presented in silence for practice with the microphone. There were two blocks of 104 test trials, each of which included 24 incongruent- and 16 congruent-colour word trials, 32 noncolour-word trials, and 32 trials with silence. Within each such condition, half of the trials occurred at a 0 ms SOA and half occurred at a 500 ms SOA. For the silent condition, a 500 ms SOA meant that 500 ms elapsed between the end of the fixation cross and the beginning of the colour square.

RESULTS

Descriptive statistics are presented in Table 1. Analyses for each task will be discussed individually, to be followed by correlation and regression analyses across tasks.

Working memory measures

Operation span task. The operation span task was scored in two different ways. One method of scoring produced a span score and the other produced a measure of the proportion of items correctly recalled (see Kane et al., 2004, for a discussion of scoring). Both methods rely on strict serial recall. In the span score method, lists that are recalled correctly are given one point and this is multiplied by the list length. The points for each list are then summed to provide a total score of the number of items correctly recalled across the task. For example, a person correctly responding to all three lists at length 2 would receive six points,

TABLE 1
Descriptive statistics for memory measures and auditory distraction tasks

Task	*Measure*	*Mean*	*Range*	*Standard error*
Operational span	Span score	19.35	2.00–50.00	1.07
	Proportion score	0.57	0.24–0.94	0.02
	Math accuracy	0.94	0.80–1.00	0.01
Running span	LL average	0.52	0.22–0.86	0.01
Serial recall	Silence	0.57	0.22–0.98	0.02
	Speech during	0.51	0.18–0.96	0.02
	Speech after	0.51	0.22–0.92	0.02
	Tone during	0.55	0.20–0.94	0.02
	Tone after	0.55	0.19–0.96	0.02
Colour naming	Colour 0 ms inc	570	421–750	6.41
	Colour 0 ms con	537	404–733	6.72
	Noncolour 0 ms	553	406–758	6.83
	Silence 0 ms	506	385–703	5.93
	Colour 500 ms inc	482	323–656	5.22
	Colour 500 ms con	458	357–600	4.76
	Noncolour 500 ms	474	376–612	4.65
	Silence 500 ms	489	378–620	4.75

N = 101. The units for the means are the proportions recalled, except for span score (number recalled) and colour naming conditions (reaction time in ms). ''During'' and ''after''' refer to whether the sound was presented during or after the list to be recalled. ''Inc'' and ''con'' refer to colour words that were inconsistent or consistent with the colours on the computer screen.

and this would be added to the points from any additional lists answered correctly. The proportion correct measure does not depend upon the entire list being answered correctly for the participant to earn credit for that list. The proportion of items correctly recalled in each list is averaged across lists, to produce a total proportion correct score.

Running memory span task. The running span task contained five list lengths, and the proportion correct scores for all list lengths were entered into a one-way, repeated-measures ANOVA. The results of this analysis indicated that performance did not differ among list lengths, and in the remaining results the average of the list lengths will be used.

Auditory distraction tasks

Cross-modal Stroop task. Data from the cross-modal Stroop task were examined to determine the number of responses coded as incorrect (the wrong colour was named), false starts (''blu...green''), or experimenter errors. Less than 1% of all trials were answered incorrectly. False starts were made on 1.5%, and experimenter errors were made on less than 0.5% of all trials. After removing those trials, the remaining trials were used to calculate the median response time for each condition. These were then entered into analyses to determine if the Stroop interference effect occurred.

First a 2 (SOA) × 4 (auditory condition: colour incongruent, colour congruent, noncolour, or silence) ANOVA of the eight different conditions in the cross-modal Stroop task was conducted. There was a significant main effect of SOA, $F(1, 100) = 428.07$, $MSE = 2054$, $p < .01$, reflecting faster overall response times in the 500 ms condition ($M = 476$ ms) than in the 0 ms condition ($M = 542$ ms). The main effect of SOA is consistent with previous findings in this task, and has been interpreted as a warning effect (Elliott et al., 1998). The main effect of auditory condition was also significant, $F(3, 300) = 73.90$, $MSE = 534$, $p < .01$. Response times were fastest in the silence and congruent colour word conditions ($M = 497$ ms for both), followed by the noncolour word condition ($M = 513$ ms). The slowest responding occurred in the incongruent colour word condition ($M = 526$ ms). Finally, the Auditory condition × SOA interaction was significant as well, $F(3, 300) = 126.01$, $MSE = 428$, $p < .01$.

To investigate this interaction, the analysis was followed by separate one-way analyses of each SOA condition. The analysis of the 0 ms SOA conditions was significant, $F(3, 300) = 122.30$, $MSE = 617$, $p < .01$, and Bonferroni post hoc testing was done to determine the differences among the conditions. All four of the 0 ms SOA conditions differed significantly (all ps < .01). The patterns of response times indicated a significant interference effect in the presence of incongruent colour words and a significant facilitation effect in the presence of congruent colour words (relative to the performance in the noncolour word condition). The analysis of the 500 ms SOA conditions was also significant, $F(3, 300) = 52.12$, $MSE = 345$, $p < .01$, and was also followed by Bonferroni post hoc testing. The results indicated that all comparisons were significantly different, with the exception of the comparison of the incongruent colour word condition and the silent condition. Consistent with the interpretation of a warning effect, response times were slowest in the silent condition. The fastest response times were seen in the congruent colour word condition, but the interpretation of a facilitation effect is confounded with the presence of the warning effect. To summarise, for the 0 ms SOA, silent < consistent < noncolour < inconsistent, whereas, for the 500 ms SOA, consistent < noncolour < inconsistent = silent.

Serial recall task. The dependent variable in the serial recall task with irrelevant sounds was the proportion of correct responses. The five auditory conditions were entered into a one-way, repeated-measures ANOVA to determine if an ISE occurred. This analysis was significant, $F(4, 400) = 11.85$, $MSE = 0.0059$, $p < .01$, and was followed by a Bonferroni post hoc test. The pattern of performance indicated by this test was that performance in silence was best and did not differ significantly from performance in the two tone conditions (tone during and tone after); however, performance in the two speech conditions was significantly lower than in both silence and tones. The two speech conditions did not differ significantly (speech during and speech after).

Due to the absence of an effect of the presentation timing of the distractors, the conditions were collapsed across trials in which the irrelevant sounds occurred during versus after the list. The remaining three conditions were then analysed to determine if the silent control condition differed significantly from the average of the speech and tone conditions. This analysis produced a significant result, $F(2, 200) = 23.69$, $MSE = 0.0038$, $p < .01$, which was then followed by a Bonferroni post hoc test that followed the same pattern as the analyses above. Tones and silence did not differ significantly, although performance in speech was significantly lower than the other two conditions. This absence of an irrelevant tone effect is inconsistent with previous work (see individual experiments within Elliott & Cowan, 2005) and is puzzling due to other published reports of irrelevant tone effects (Jones & Macken, 1993; LeCompte, Neely, & Wilson, 1997). However, as we will see, one explanation is that individuals differed in their ability to overcome the effects of irrelevant tones, in a systematic rather than random manner.

Correlations and regression analyses

Correlations. Correlations among the working memory measures and the irrelevant-sound tasks are presented in Table 2. Inspection of this table reveals the clear relationships among the working memory measures and the conditions of the serial recall task, but no relationships among the working memory measures and the conditions of the colour-naming task. The colour and noncolour conditions in the Stroop task produced significant negative correlations with serial recall; the correlations were negative because faster performance (lower RTs) in the colour-naming task went with higher serial recall scores.

The absence of any significant relationship between working memory measures and colour-naming speeds in Table 2 is difficult to interpret because the attentional control factor in colour naming should be found only in the subtraction of the colour conditions from the corresponding noncolour conditions, or possibly in the subtraction of the noncolour conditions from the silent condition. We have noted that these subtractions do not result in very reliable measures. In fact, Cronbach's alpha measures of reliability for the effect of speech during or after the list and tones during or after the list, as subtractions from the silent control condition, were all .23 or lower, despite the high reliabilities for individual conditions shown in Table 2; and those for subtractive measures in the Stroop task were lower still. Indeed, there were no significant correlations between working memory measures and any of the difference scores, in either the serial recall task or in the colour-naming task.

TABLE 2

Correlations with working memory measures and irrelevant-sounds tasks

Task	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
Working memory														
OspanScore (1)	*.73*													
OspanProp (2)	.90**	*.71*												
RunSpan (3)	.31**	.34**	*.90*											
Serial recall														
Silence (4)	.36**	.36**	.59**	*.85*										
Speech average (5)	.35**	.36**	.54**	.86**	*.91*									
Tone average (6)	.43**	.42**	.58**	.88**	.92**	*.90*								
Colour naming														
Colour0Ine (7)	−.14	−.11	−.16	−.26*	−.29*	−.29*	*.83*							
Colour0Con (8)	−.11	−.05	−.19	−.18	−.21*	−.20*	.84**	*.67*						
Noncolour (9)	−.06	−.03	−.17	−.25*	−.25*	−,21*	.89**	.83**	*.90*					
Silence0 (10)	−.13	−.10	−.10	−.19	−.23*	−.19	.89**	.81**	.89**	*.88*				
Colour500Inc (11)	−.09	−.05	−.15	−.23*	−.26*	−.26*	.78**	.72**	.74**	.79**	*.83*			
Colour500Con (12)	−.10	−.02	−.19	−.24*	−.31*	−.28*	.73**	.70**	.70**	.75**	.82**	*.44*		
Noncolour500 (13)	−.11	−.05	−.10	−.19	−.24*	−.23*	.79**	.72**	.79**	.81**	.87**	.83**	*.87*	
Silence500 (14)	−.07	−.02	−.16	−.21*	−.25*	−.22*	.83**	.78**	.83**	.87**	.89**	.82**	.90**	*.89*

$N = 101$. Ospan = operation span, Prop = proportion, RunSpan = running span. Colour0Inc = colour naming with a 0 ms interval and inconsistency between the colour and spoken word; Con = consistent; and so on. Numbers in italics are chronbach's alpha measure of reliability. In the colour naming task, reliability measures were based on mean reaction times with the exclusion of data points below 100 ms or above 2 s; these were highly (> .93) correlated with the median reaction times, which were used for the reported correlations.

$^{*}p < .05$, $^{**}p < .001$.

Regressions. Past research in the area of individual differences in the ISE has used difference scores to measure the magnitude of the disruption, typically subtracting performance in the presence of speech from performance in silence. However, the low reliability of difference scores can be problematic (Elliott & Cowan, 2005). Given this difficulty, regression analyses were used to partition the variance from the silent condition separately from the conditions with irrelevant sounds, thus removing the variance due to the baseline measure.

In the first regression analysis, operation span scores were used as the dependent variable with the silent control condition entered in first, producing a significant $R^2 = .13$, $p < .001$. When the average of the irrelevant-tone condition was added in second, the result was a significant $\Delta R^2 = .06$, $p < .01$. This same analysis was done in the reverse order and the average of the tone condition accounted for 19% of the variance, $p < .001$. When the silent control condition was added in second it did not account for any additional variance. To ascertain the direction of this relationship, partial correlations were calculated. The correlation between the operation span score and the silent control condition, with the tone average condition partialled out, was $r_p = -.05$, n.s. However, the correlation between the operation span score and the irrelevant-tone condition average, with the silent control condition partialled out, was $r_p = .26$, $p < .01$. This indicates an effect of the tones in the serial recall task that is larger for low span individuals than for high span individuals; the high span individuals do better in the tone condition than one would predict on the basis of their silent control condition scores. This appears to be one consequence of the better attentional-control capabilities of high span individuals.

In order to clarify the significant regression effects, results were examined separately in the 26 individuals with the highest and 26 individuals with the lowest operation spans (approximately the top and bottom quartiles). For high spans, the proportions correct in the silence and tone conditions were nearly identical at .69 and .68, respectively. In the low spans, they were also nearly identical, at .48 and .47, respectively. However, the relation between silence and tone conditions appeared to differ between these groups. For high spans, the regression equation was: ‘‘tones = .72 (silence) + .18’’. For low spans, however, it was ‘‘tones = .78 (silence) + .10’’. It was almost identical to this for the middle-span subjects: ‘‘tones = .77 (silence) + .10’’. An informal description of these results is that, in high span individuals, performance in the presence of tones was more protected than it was in other participants, even when performance in silence was relatively low.

Regression analyses comparable to those for the tone conditions were conducted with the average of the irrelevant-speech conditions. Speech consistently accounted for nonsignificant amounts of variance beyond what was accounted for by the silent control condition. The implications are that the dual-task nature of the serial recall task with irrelevant sounds may differ depending on whether the sounds are speech or tones. The tone condition accounts for extra variance.

Very similar patterns of results were obtained when the operation span proportion score was used. Perhaps, unlike tones, irrelevant speech cannot be blocked by the mechanisms of attention, even in high span individuals.

Also, comparable analyses were conducted with running memory span, as opposed to operation span, as the criterion variable, but the results were not the same. The silent control condition accounted for a great deal of the variance in running memory span, $R^2 = .35$, $p < .001$. When the irrelevant-tone condition was added in second it only accounted for a nonsignificant 2% of additional variance. When the irrelevant-tone condition was entered into the equation first, it produced $R^2 = .34$, $p < .001$, and the silent control condition produced $\Delta R^2 = .03$, $p < .05$. Similarly, the irrelevant-speech condition added no variance when entered after the silent control condition. The speech condition entered first accounted for $R^2 = .28$, $p < .001$, and the silent condition added afterward contributed $\Delta R^2 = .06$, $p < .01$. Thus, for running span, the irrelevant-sound conditions accounted for less variance than the silent control condition. One possible theoretical interpretation of these results is consistent with the idea that the running span task is performed in a very different manner than the operation span task, thus leading to the differences between the regression results for these two working memory tasks. The running span task relies more heavily on participants drawing the responses from the focus of attention, as compared to the operation span task, which involves reactivating items that are no longer in the focus of attention. We will return to this matter of the differences between the two working memory tasks in the final discussion.

All eight Stroop-task measures (0 ms and 500 ms presentations of inconsistent colour words, consistent colour words, noncolour words, and silence) also were examined, but even the combination of all of them did not account for a significant amount of variance in regressions on operation span or running span.

DISCUSSION

This study represents the first attempt to compare different types of measures of distraction by irrelevant sounds, both to each other and to individual difference measures of working memory. Several previous studies have looked for positive relations between working memory and the ability to overcome effects of irrelevant sounds, which would be manifest as negative correlations between working memory task performance and the magnitude of the ISE (Beaman, 2004; Ellermeier & Zimmer, 1997; Elliott & Cowan, 2005; Neath et al., 2003). This relation would be expected on the grounds that high working memory indicates the ability to control attention and filter out distraction (e.g., Conway et al., 2001; Kane et al., 2001). Nevertheless, an effect of that nature still has never been obtained for the irrelevant-speech effect of Salamé and Baddeley (1982), and the present study is the first one to find such a relation in the context of the irrelevant-sound effect of Jones and Macken (1993), with tone distractors. There is convergent evidence in developmental work; Elliott (2002) found that young children had larger ISEs than adults, for both speech and tone distractors.

The significant relation between working memory and the irrelevant-tone effect on serial recall observed in the present study was relatively weak, but it still is of considerable theoretical importance that the expectation was confirmed. One account of the present results is that attention can be used to filter out distracting tones, though only by individuals with high working memory span. In contrast, it may not be possible even for them to filter out distracting speech sounds, which may automatically enter the phonological loop and contaminate memory for printed verbal items, as suggested by Salamé and Baddeley (1982). In a very different experimental procedure that does not involve serial recall, Conway et al. (2001) found a relation between working memory and the ability to filter out irrelevant speech, but that was when the relevant and irrelevant channels were speech streams to opposite ears and the relevant channel only had to be repeated, not retained.

There are several reasons why previous studies may not have obtained the expected relations between working memory and irrelevant-sound effects. The first is the poor reliability of difference scores that are taken as the proper measures of irrelevant-sound effects. That problem can be overcome through regressions using combinations of the raw scores instead of correlations between difference scores, as Elliott and Cowan (2005) have already pointed out. It was that regression technique that was successful here. Second, though, Elliott and Cowan noted psychometric properties that could tend to cancel out the desired relation. Individuals who perform better in the silent condition of serial recall have further for their scores to drop in the presence of irrelevant sounds. Third, most previous studies of individual differences did not include irrelevant-tone effects, though Elliott and Cowan did.

However, the present study is the first one in which, despite any such measurement difficulties, a positive relation was observed between working memory and the ability to filter out irrelevant sounds. One of the main methodological differences in this study, as compared to previous work, was the level of difficulty of the task. This ISE task used a series of consonants as the to-be-remembered items, whereas previous studies have used the digits 1–9. Although the only keys that participants could type on the keyboard were those in the response set, the keys were not marked in any way and it appeared to affect the difficulty level of the task. Mean levels of performance were lower than reported data, such as Elliott and Cowan (2005). It is unlikely that participants were able to memorise the items in the response set and, as result, did not perform as well. This finding of lower performance with consonants as compared to digits was also shown by Hughes, Vachon, and Jones (2005), and they discussed the role of stimulus familiarity in the task. Future research should address this issue because increasing the level of difficulty of the task may be responsible for the correlations we obtained here.

Mixed results were obtained on the question of whether the process of filtering out irrelevant sounds is similar or different in serial recall and colour naming tasks. No relation was observed between working memory and the ability to filter our irrelevant stimuli in the Stroop procedure, but the work of Kane and Engle (2003) suggests that such a relation should be easier to observe when the proportion of trials with consistency between the relevant colour and the irrelevant colour word is made much higher, so that attending to the spoken word would yield the correct result on most trials. In the present study, because only 15.4% of all trials included a colour word consistent with the printed colour, note that these consistent-colour words were not truly helpful to the subjects. With a 0 ms SOA they resulted in reaction times faster than noncolour or inconsistent-colour words, but still slower than in the silent control condition. We predict that, in a procedure in which the majority of trials were consistent-colour trials (cf. Kane & Engle, 2003), silence would produce slower responding than the consistent-word condition, so that attending to the consistent word would become truly helpful on most trials. (That information cannot be obtained from Kane and Engle as they did not use a cross-modal Stroop-like procedure and did not include stimuli with no irrelevant verbal component.)

However, another alternative is that the ability to avoid distraction from irrelevant sounds is not a unitary construct. The lack of correlations between performance in the cross-modal Stroop task and the measures of working memory could have been driven by the lack of a memory component in the Stroop task. As mentioned above, without the goal maintenance required by a high level of congruent trials, the cross-modal Stroop task did not share in the task requirements of the three other tasks in this experiment. The cognitive mechanisms required when avoiding distraction in a memory-demanding situation may be different from the cognitive mechanisms required in a situation where quick response times are instead the priority.

Additionally, this experiment allowed a comparison of two measures of working memory: running span and operation span. The positive correlation between these two measures provides evidence that they are in fact, measuring the same construct. However, the correlations were r = .31 when using the Ospan score and r = .34 when using the Ospan proportion measure. It is known that both of these measures correlate well with aptitudes (Cowan et al., 2005), and this ability to predict aptitudes may be what is driving the relationship observed in the present data set. Differences among the two working memory tasks could be responsible for the differences in predictive ability with the irrelevant-sounds serial recall task. This is also consistent with the idea that the ability to avoid distraction may not be one single ability, but instead of combination of abilities dependent upon task demands.

Previous work by other investigators has shown that the ability to overcome proactive interference (PI) is very important in the relationship between performance measures and working memory scores (Lustig, May, & Hasher, 2001). In the study by Lustig et al., measures were taken to minimise the amount of PI present in a typical working memory task. When this was done, correlations between performance on the working memory task and performance in a prose recall task dropped to zero. This work illustrates that an important component of working memory capacity is the ability to overcome PI. In the current study, the two working memory tasks that we used may induce participants to counteract the negative effects of PI in different ways. In running span, participants can use the focus of attention to perform the task successfully, while in operation span participants will have to make connections with items that have moved out of the focus of attention and reactivate these items at the time of recall. Previous evidence has demonstrated that running span, and another typical measure of working memory known as reading span, account for different patterns of variance, while a task such as counting span accounted for similar variance to running span (Cowan et al., 2003). Running span may correlate with measures of working memory that tap into this ability to hold items in the focus of attention, thus explaining the correlations with measures of intelligence while explaining the lack of correlations with the ISE.

More research is needed for the development of a model of the ability to avoid distraction from irrelevant sounds, including the idea that this ability may not be a unitary construct. The present study provides several important pieces of information regarding the relationships among working memory and the ISE, as well as the relationship between the two irrelevant-sound tasks used here. The results from the tone condition of the serial recall task are especially noteworthy. Future research will examine the relationship between operation span and the irrelevant-tone condition more closely, as well as the level of difficulty of the serial recall task.

REFERENCES

Beaman, C. P. (2004). The irrelevant sound phenomenon revisited: What role for working memory capacity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(5), 1106–1118.

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 459–471.

Beaman, C. P., & Jones, D. M. (1998). Irrelevant sound disrupts order information in free recall as in serial recall. *Quarterly Journal of Experimental Psychology*, *51A*, 615–636.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*, 17–31.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review*, *8*, 331–335.

Conway, A. R. A., & Engle, R.W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. *Memory*, *4*, 577–590.

Cowan, N. (1995). *Attention and memory: An integrated framework* (Oxford Psychology Series, No. 26). New York: Oxford University Press. (Paperback edition published 1997)

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Cowan, N., & Barron, A. (1987). Cross-modal, auditory–visual Stroop interference and possible implications for speech memory. *Perception and Psychophysics*, *41*, 393–401.

Cowan, N., Elliott, E. M., & Saults, J. S. (2002). The search for what is fundamental in the development of working memory. In R. Kail & H. Reese (Eds.), *Advances in Child Development and Behavior* (Vol. 29, pp. 1–49). San Diego, CA: Academic Press.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100.

Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliott, E. M., Lacey, J. F., et al. (2003). Children's working memory processes: A response-timing analysis. *Journal of Experimental Psychology: General*, *132*, 113–132.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the ''irrelevant speech effect''. *Journal of the Acoustical Society of America*, *102*, 2191–2199.

Elliott, E. M. (2002). The irrelevant speech effect and children: Theoretical implications of developmental change. *Memory and Cognition*, *30*, 478–487.

Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color–word interference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 654–667.

Elliott, E. M., & Cowan, N. (2005). Individual differences in memory span and in the effects of irrelevant sounds on memory performance: The coherence of the irrelevant sound effect. *Memory and Cognition*, *33*, 664–675.

Elliott, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal, color–word interference effects. *Perception and Psychophysics*, *60*, 761–767.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Hockey, R. (1973). Rate of presentation in running memory and direct manipulation of input-processing strategies. *Quarterly Journal of Experimental Psychology A*, *25*, 104–111.

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 736–749.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 369–381.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47–70.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189–217.

LeCompte, D. C., Neely, C. B., & Wilson, J. R. (1997). Irrelevant speech and irrelevant tones: The relative importance of speech to the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 472–483.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *130*, 199–207.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*, 56–60.

Mukunda, K. V., & Hall, V. C. (1992). Does performance on memory for order correlate with performance on standardized measures of ability? A meta-analysis. *Intelligence*, *16*, 81–97.

Neath, I., Farley, L. A., & Surprenant, A. M. (2003). Directly assessing the relationship between irrelevant speech and articulatory suppression. *Quarterly Journal of Experimental Psychology*, *56*(A), 1269–1278.

Pollack, I., Johnson, L. B., & Knaff, P. R. (1959). Running memory span. *Journal of Experimental Psychology*, *57*, 137–146.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, *21*, 150–164.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *17*, 242–248.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel*? Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 255–260.